Revista Colombiana de Estadística
Junio 2010, volumen 33, no. 1, pp. 1 a 11

A Probability Model for the Child Mortality in a Family

Un modelo probabilístico para la mortalidad en la infancia en una familia

Himanshu Pandey[a], Jai Kishun[b]

Department of Mathematics and Statistics, Gorakhpur University, Gorakhpur, India

Abstract

This paper proposed, under assumptions of inflated type fixed displaced geometric model, the distribution pattern of families according to number of child deaths within the first five years of life. The proposed model involves several parameters related to child mortality in a family, which is estimated with Method of Moments and Maximum Likelihood Estimation techniques. The proposed models fitted the observed data showing a better approximation at the survey area and draw some vital conclusions.

***Key words*:** F distribution, G-estimation, M-estimation, t-distribution, Mortality, Death, Contraception, Probability model.

Resumen

Este documento presenta, bajo el supuesto de un modelo geométrico de desplazamiento fijo, patrones de distribución de las familias, de acuerdo con el número de defunciones de sus hijos menores de cinco años. El modelo emplea diferentes parámetros relacionados con la mortalidad en la infancia en una familia, estimada con el método de momentos y de máxima verosimilitud. Los modelos propuestos ajustan los datos observados, mostrando mejor aproximación a la encuesta de área y describe algunas conclusiones vitales.

***Palabras clave*:** distribución F, estimación G, estimación M, Distribución t, mortalidad, anticoncepción, modelos de probabilidad.

[a]Professor. E-mail: himanshupandey@is.iiita.ac.in
[b]Professor email

1. Introducción

Considerable interest has been shown in the past by several researchers to measure the levels of child mortality. Currently in the Developing Nations, the force of child mortality is still high at the younger ages particularly during the infancy. Infant and child mortality remain disturbingly high in developing countries despite the significant decline in most parts of the developed world. The state of the world's children indicated that about 12.9 million children die every year in the developing world (UNICEF 1987). Mortality for infants and child under the age of 5 years are expressed as the number of deaths in a given period. Infant mortality is defined as death during the first year of life and child mortality as that between the first and fifth birthdays. The deaths during childhood suffer from substantial degree of errors. Usually errors occurs due to recall laps which result in omission of events, misplacement of deaths and the distortion of reports on the duration of vital events.

The most important factors for child mortality are food shortages, contaminated water crowded and sub-standard housing, unchecked infectious diseases, the absence of day care facilities for the children of working mothers and the lack of minimally adequate and free medical care. The influences of most of these factors are absent in the present areas of these countries under study. However, some unmeasured genetic, environmental and behavioural components still remain nonnegligible.

In demography Child Mortality are useful as a Sensitive Index of a Nation's Health Conditions and as guided for the structuring of Public Health Programmes. Child Mortality is interrelated to social, cultural, economic, physiological and other factor. The high rate of infant and child mortality shows a low-level development of the health programme and also for the Nation's. Infant and Child mortality has been of interest of researchers and demographers because of its apparent relationship with fertility and indirect relationship with the acceptance of modern contraceptive methods (Kabir & Amir 1993).

Some attempts have been made to estimate the levels of child mortality by using data available from the different survey and other specific sources. Hill & Devid (1989) have suggested an approach for estimating child mortality from all births which have taken place in last five years before the survey. However, the estimate obtained through this method also suffers from the problem of under reporting (Pathak et al. 1991). In these circumstances, a number of attempts have been made to study the age pattern of mortality by using models (Goldblatt 1989, Heligman & Pollard 1980, Krishnan & Jin 1993, Ronald & Lawrence 1992, Thiele 1972). Initially, Keyfit (1977) used a hyperbolic function to study the infant and child mortality. Later, Arnold (1993) used pareto distribution; Krishnan & Jin (1993) and Chauhan (1997) applied finite range model for the same.

Brass (1995*a*, 1995*b*) is one of the proponents of indirect method of mortality estimation. He based his mortality estimate on retrospective data given by women of reproductive age on the number of children ever born and their status (either death or living). Other contributor in this line includes Preston & Palloni

(1977). However, indirect infant and child mortality estimates result from poor, inadequate and incomplete data, especially in developing countries. Most deaths outside hospital premises were not recorded and that many people do not record infant deaths because they only keep track of such occurrence as misfortunes, and when recorded, the age at death were either under or overstated.

The direct measures of mortality being not reliable, the problem may be overcome by the recently developed model building approaches which make it possible to obtain estimates from information other than vital statistics. In this connection, Srivastava (2001) has been proposed a probability model for the distribution of family according to number of child deaths (Deaths within the first five years of life). The main objective of this paper is to modify the Srivastava (2001) model by taking the finite range.

2. Probability Model

Let x denote the number of child deaths in a family at the survey point. Then the distribution of x is derived under the following assumption.

1. Only those families are considered in which at least one birth prior to the survey has occurred.
2. At the survey point, a family either has experienced a child loss or not. Let α and $(1-\alpha)$ be the respective proportions.
3. Out of α proportion of families, let β be the proportion of families in which only one child death has occurred.
4. Remaining $(1-\beta)\alpha$ proportion of families, experiencing multiple child deaths, follows a displaced geometric distribution with parameter p according to the number of child deaths.

Under these assumptions, the probability distribution of X is given by

$$\left.\begin{aligned} P[X=0] &= (1-\alpha),\ k=0 \\ P[X=1] &= \alpha\beta,\ k=1 \\ P[X=k] &= \frac{(1-\beta)\alpha p q^{k-2}}{1-q^N},\ k=2,3,\dots,N \end{aligned}\right\} \tag{1}$$

Where p denotes the success of child deaths in a family and $q=1-p$. If we put $q^N=0$, then the proposed model (1) reduced to Srivastava (2001). The proposed model (1) is an improvement over the Srivastava (2001) model by taking a concept of finite range model i.e. $k=0,1,2,3,\dots,N$.

2.1. Estimation Method of moment

The method of moments is discussed to estimate the four parameters α, β, p and N of the probability function (1). Now it is difficult to estimate all these parameters, so it is assumed that N is the maximum numbers of child deaths occurred. Let (x, f) be a frequency distribution whose parameters to be estimated from the observed distributions of families according to the number of child deaths. The following estimation technique is employed for estimating rest of the parameter.

Equating proportions of zeroth cell, first cell frequency and simple mean to their corresponding observed values respectively, which is converted into the following equations,

$$\frac{f_0}{f} = 1 - \alpha \tag{2}$$

$$\frac{f_1}{f} = \alpha\beta \tag{3}$$

$$\overline{X} = \alpha\beta + (1-\beta)\alpha \left[\frac{(1 - q^{N-1})}{p(1 - q^N)} - \frac{N q^{N-1}}{(1 - q^N)} + \frac{1}{(1 - q^N)} \right] \tag{4}$$

Where,

$f_0 =$ Observed value of zeroth cell frequency

$f_1 =$ Observed value of first cell frequency

$f =$ Total number of observations $= \sum_i f_i$

$\overline{X} =$ Sample mean of the observed values.

2.2. Method of Maximum Likelihood

The proposed model involves four parameters α, β, p and N to be estimated from the observed distribution of families according to the number of child deaths, but it cannot be possible to estimate all these simultaneously by this method, so the value of N has been taken as method of moments. Let $x_1, x_2, x_3, \ldots, x_n$ be a random sample of size N from the population (1). The likelihood function L for the given sample can be expressed as

$$L = (1-\alpha) f_0 (\alpha\beta) f_1 \times \left[\frac{(1-\beta)\alpha p}{1 - q^N} \right]^{f_2} \left[\frac{\alpha\{1 - \beta - (1-\beta)p\}}{1 - q^N} \right]^{f - f_0 - f_1 - f_2} \tag{5}$$

Taking log both sides, we get

$$\log L = f_0 \log(1-\alpha) + f_1 \log(\alpha\beta) + f_2 \log \left[\frac{(1-\beta)\alpha p}{1 - q^N} \right] + (f - f_0 - f_1 - f_2) \log \left[\alpha(1-\beta) \left\{ 1 - \frac{p}{1 - q^N} \right\} \right]$$

Now, partially differentiating w.r.t. α, β and p respectively and equating to zero.

$$\frac{\partial \log L}{\partial \alpha} = -\frac{f_0}{(1-\alpha)} + \frac{f-f_0}{\alpha} = 0 \tag{6}$$

$$\frac{\partial \log L}{\partial \beta} = \frac{f_1}{\beta} + \frac{f-f_0-f_1}{1-\beta} = 0 \tag{7}$$

$$\frac{\partial \log L}{\partial p} = \frac{f_2\left[\left(1-q^N\right) - pN(1-p)^{N-1}\right]}{p\left(1-q^N\right)} - (f-f_0-f_1-f_2)\left[\frac{\left(1-q^N\right) - pN(1-P)^{N-1}}{\left[\left(1-p^N\right)-p\right]\left(1-p^N\right)}\right] = 0 \tag{8}$$

Solving equations (6), (7) and (8), the estimate of α, β and p can easily be obtained as

$$\alpha = \frac{f-f_0}{f}$$

$$\beta = \frac{f_1}{f-f_0}$$

$$\frac{p}{\left(1-q^N\right)} = \frac{f_2}{f-f_0-f_1}$$

The second partial derivatives of $\log L$ obtained is

$$\frac{\partial^2 \log L}{\partial \alpha^2} = -\frac{f_0}{(1-\alpha)^2} - \frac{f-f_0}{\alpha^2} \tag{9}$$

$$\frac{\partial^2 \log L}{\partial \beta^2} = -\frac{f_1}{\beta^2} - \frac{f-f_0-f_1}{(1-\beta)^2} \tag{10}$$

$$\frac{\left(\partial^{\uparrow} 2 \log L\right)}{\left(\partial p^{\uparrow} 2\right)} = \frac{\left(f_{\downarrow} 2[N(1-N)(1-p)^{\uparrow}(N-2)]\right)}{\left(1-q^{\uparrow} N\right)} - \frac{\left(f-f_{\downarrow}0-f_{\downarrow}1-f_{\downarrow}2\right)\left[\left\{\left\{(1-q^{\uparrow})^{\uparrow}2p\left(1-q^{\uparrow}\right)\right\}\left\{N(N-1)p(1-p)^{\uparrow}(N-2)\right\}\right\}\right]}{\left[(1-q^{\uparrow}N)^{\uparrow}2\right]} \tag{11}$$

Now, partial derivative of $\frac{\partial \log L}{\partial \alpha}$, $\frac{\partial \log L}{\partial \beta}$ and $\frac{\partial \log L}{\partial p}$ w.r. to, β, p and α respectively we get as following,

$$\frac{\partial^2 \log L}{\partial \alpha \partial \beta} = \frac{\partial^2 \log L}{\partial \beta \partial p} = \frac{\partial^2 \log L}{\partial p \partial \alpha} = 0 \tag{12}$$

Here,

$$E(f_0) = f(1-\alpha)$$

$$E(f_1) = f\alpha\beta$$

$$E(f_2) = \frac{f(1-\beta)\alpha p}{1-q^N}$$

$$E(f - f_0 - f_1 - f_2) = f\alpha(1-\beta)\left\{1 - \frac{p}{1-q^N}\right\}$$

Using the above facts, the expected value of the second partial derivatives obtained as

$$\phi_{11} = E\frac{\left[\frac{-\partial^2 \log L}{\partial\alpha^2}\right]}{f} = \left[\frac{1}{1-\alpha} + \frac{1}{\alpha}\right] \tag{13}$$

$$\phi_{22} = E\frac{\left[\frac{-\partial^2 \log L}{\partial\beta^2}\right]}{f} = \alpha\left[\frac{1}{1-\beta} + \frac{1}{\beta}\right] \tag{14}$$

$$\phi_{22} = E\frac{\left[\frac{-\partial^2 \log L}{\partial\beta^2}\right]}{f} = \alpha\left[\frac{1}{1-\beta} + \frac{1}{\beta}\right] \tag{15}$$

the covariance between the estimators becomes zero since

$$E\left(\frac{\partial^2 \log L}{\partial\alpha\partial\beta}\right) = E\left(\frac{\partial^2 \log L}{\partial\beta\delta p}\right) = E\left(\frac{\partial^2 \log L}{\partial\alpha\partial p}\right) = 0 \tag{16}$$

Thus, the asymptotic variances of the estimator can be obtained as

$$V(\widehat{\alpha}) = \frac{1}{\phi_{11}}, V\left(\widehat{\beta}\right) = \frac{1}{\phi_{22}}, V(\widehat{p}) = \frac{1}{\phi_{33}} \tag{17}$$

3. Application

The suitability of the proposed model is examined to the study that has been conducted in North-Eastern Libya stretching from Benghazi to Emsaad. From the study area, 7 localities out of 27 have been selected by probability proportional to numbers of families in the localities. The data on fertility and mortality under age 5 along with some other demographic characteristics have been collected from 1,252 couples of childbearing ages of selected localities. About one-third (35.7 percent) of the investigated mothers have lost at least one child. The percentage of multiple child loss mothers is 11.3 and these mothers have given, one an average, 10 or more births. The differential in child loss by fertility level is highly significant. However, child mortality to mothers having lower and differential in child mortality by fertility in north-eastern Libya 325 medium (6 ever born children) fertility is similar. This study indicates that high parity and high mortality move in the same direction (Bhuyan & Deogratias 1999) and one set of data has been taken from a Household Sample Survey in Brazil in 1987. Details are given in Sastry (1997). Other two set of sample data were collected under a Survey entitled "Effect of breastfeeding on fertility in North Rural India" in 1995 and "A Demographic survey on fertility and mortality in rural Nepal: A Study of Palpa and Rupandehi Districts" in 2000. The details of these two set of data are given in Srivastava (2001).

The parameters of the proposed model have been estimated by the method of moment and method of maximum likelihood. The estimated values of different parameters are given in tables 1 to 4 for the child deaths.

The estimated value of α are 0.3753, 0.2139, 0.2683 and 0.3570 for India, Nepal, North East Brazil and North East Libya, respectively. It represents that the proportion of families experiencing a child loss was found slightly higher in India (0.3753) than North East Libya (0.3570), North East Brazil (0.2683) and Nepal (.2139). The estimate of β are 0.5857, 0.7528, 0.6560 and 0.6846, respectively, for India, Nepal, North East Brazil and North East Libya. It means that the proportion of families having only one child death was found greater for Nepal (0.7527) as compared to other countries. The estimated values for the probability of success of death p are 0.5889, 0.6257, 0.6490 and 0.6630 by the method of moment and 0.6022, 0.7110, 0.6178 and 0.6592 by the maximum likelihood, respectively, for the above mentioned countries. The average number of child death per family $\alpha\beta + (1-\beta)\alpha\left[\frac{(1-q^{N-1})}{p(1-q^N)} - \frac{Nq^{(N-1)}}{(1-q^N)} + \frac{1}{(1-q^N)}\right]$ for Eastern Uttar Pradesh (India), Nepal, North East Brazil and North East Libya were found to be 0.63, 0.30, 0.41 and 0.53 respectively. This show that, on an average, the child mortality is high in Eastern Uttar Pradesh (India). The exact variances of the estimators obtained by maximum likelihood method are also given. for Eastern Uttar Pradesh (India), Nepal, North East Brazil and North East Libya were found to be 0.63, 0.30, 0.41 and 0.53 respectively. This show that on an average the child mortality is high in Eastern Uttar Pradesh (India). The exact variances of the estimators obtained by maximum likelihood method are also given.

Changes in levels of mortality may be attributed to socioeconomic factors such as improvements in primary health care services, control of epidemics, availability of health care facilities, and with the improvement in economic condition among lower parity women, there is a downward shift in child mortality. However economic condition and mortality move in the same direction among high parity women. Differential impacts of age of female spouse at marriage are observed among mothers of different parity level.

A inflated geometric distribution for finite range provides a suitable description of child mortality at micro level, i.e. at the family level (Tables 1 to 4). The value of χ^2 are insignificant at 5 percent level of significance for all set of data. The proposed model fitted satisfactorily and described the pattern of child mortality to several sets of sample data in Indian Subcontinents.

Table 1: Distribution of Observed and Expected Number of Families, according to the Number of Child Deaths in Eastern Uttar Pradesh (India).

Number of child dead	Observed number families	Method of Moment (Expected no. of families)	Method of Maximumum Likelihood (Expected no. of families)
0	506	506.0070	506.0070
1	178	178.0490	178.0490
2	76	74.3124	75.9659
3	32	30.5520	30.2177
4	8	12.5609	12.0200
5	6	5.1633	4.7792
6	3	2.1226	1.9007
7	1	1.1628	1.0529
Total	810	810.0000	810.0000
$\widehat{\alpha}$		0.3753	0.3753
$\widehat{\beta}$		0.5857	0.5857
$\widehat{p}$		0.5889	0.6023
$V(\widehat{\alpha})$			0.000289
$V(\widehat{\beta})$			0.000798
$V(\widehat{p})$			0.003006
χ^2		2.2841	2.3983
$d.f.$		4	4

Source: Srivastava (2001)

Table 2: Distribution of Observed and Expected number of Families, According to the Number of Child Deaths in Nepal.

Number of child dead	Observed number families	Method of Moment (Expected no. of families)	Method of Maximumum Likelihood (Expected no. of families)
0	669	668.9711	668.9711
1	137	137.0314	137.0314
2	32	28.1500	31.6996
3	6	10.5365	9.2479
4	3	3.9438	2.6726
5	2	1.4762	0.7724
6	2	0.5525	0.5232
7	0	0.3385	0.0768
Total	851	851.0000	851.0000
$\widehat{\alpha}$		0.2139	0.2139
$\widehat{\beta}$		0.7528	0.7528
$\widehat{p}$		0.6257	0.7110
$V(\widehat{\alpha})$			0.000196
$V(\widehat{\beta})$			0.0001022
$V(\widehat{p})$			0.006352
χ^2		7.0227	7.3799
$d.f.$		4	4

Source: Srivastava (2001)

TABLE 3: Distribution of Observed and Expected Number of Families, according to the Number of Child Deaths in North East Brazil.

Number of child dead	Observed number families	Method of Moment (Expected no. of families)	Method of Maximumum Likelihood (Expected no. of families)
0	769	769.0167	769.0167
1	185	184.9810	184.9810
2	60	62.9986	59.9999
3	26	22.1125	22.9320
4	9	7.7615	8.7646
5	1	2.7243	3.3498
6	1	0.9562	1.2803
7	0	0.4492	0.6757
Total	1051	1051.0000	1051.0000
$\widehat{\alpha}$		0.2683	0.2683
$\widehat{\beta}$		0.6560	0.6560
$\widehat{p}$		0.6490	0.6178
$V(\widehat{\alpha})$			0.000215
$V(\widehat{\beta})$			0.000800
$V(\widehat{p})$			0.004051
χ^2		2.5663	2.8021
$d.f.$		4	4

Source: Sastry (1997)

TABLE 4: Distribution of Observed and Expected Number of Families, according to the Number of Child Deaths in North East Libya.

Number of child dead	Observed number families	Method of Moment (Expected no. of families)	Method of Maximumum Likelihood (Expected no. of families)
0	805	805.0360	805.0360
1	306	305.9916	305.9916
2	93	93.5115	92.9755
3	36	31.5134	31.6953
4	7	10.6200	10.8049
5	2	3.5789	3.6834
6	1	1.2061	1.2557
7	2	0.5425	0.5576
Total	1252	1252.0000	1252.0000
$\widehat{\alpha}$		0.3570	0.3570
$\widehat{\beta}$		0.6846	0.6846
$\widehat{p}$		0.6630	0.6592
$V(\widehat{\alpha})$			0.000172
$V(\widehat{\beta})$			0.000483
$V(\widehat{p})$			0.002457
χ^2		6.5231	6.4771
$d.f.$		4	4

Source: Bhuyan & Deogratias (1999)

4. Conclusion

From the above discussion of proposed model related with child mortality it is concluded that infant and child mortality are still higher in Developing countries like India, Nepal, N-E Brazil and N-E Libya as shown in Tables 1 to 4. Changes in the level of child mortality is directly or indirectly associated with socioeconomic factors prevailing in that countries.

To overcome the child mortality spread of education is first and foremost essential factor. By imparting education to the parents and family members they will be conscious and aware to the health of the child. To keep their child healthy and sound they will follow the advice of the doctors from the date when the mother is conceived. Parents will adhere the time schedule to vaccinate the child to control different diseases. They will also care for hygienic atmosphere and cleanness.

Keeping all viewpoints, the proposed model is an essential tool for predicting the child mortality level of any country and it can be used as an indicator of good health condition of the society.

Acknowledgement

The authors are very thankful to referees for their valuable suggestions.

[Recibido: julio de 2009 — Aceptado: diciembre de 2009]

References

Arnold, B. C. (1993), 'Pareto Distributions', *Statistical Distributions* **5**.

Bhuyan, K. C. & Deogratias, R. (1999), 'On a Probability Model for Child Mortality Pattern in North East Libya', *Turkish Journal of Population Studies* (21), 33–38.

Brass, W. (1995*a*), A Simple approximation for the time Location of Estimates of Child Mortality from Proportions dead by age of Mother, *in* 'Advances in Methods for Estimating Fertility and Mortality from Limited and Defective Data', London School of Hygiene and Tropical Medicine, London, pp. 1–16.

Brass, W. (1995*b*), The Derivative of Life Tables from Retrospective Estimates of Child and Adult Mortality, *in* 'Advances in methods for estimating fertility and mortality from limited and defective data', London School of Hygiene and Tropical Medicine, London.

Chauhan, R. K. (1997), 'Graduation of Infant Deaths by Age', *Demography India* **2**(26), 261–274.

Goldblatt, P. O. (1989), 'Mortality by Social Class, 1971-85', *Population Trends* (56), 6–15.

Heligman, L. & Pollard, J. H. (1980), 'Age Pattern of Mortality', *Journal of Institute of Actuaries* (117), 49–80.

Hill, A. G. & Devid, H. P. (1989), Measuring Child Mortality in the Third World, *in* N. Sources & N. Approaches, eds, 'IUSSP Proceeding of International Conference', New Delhi, India.

Kabir, M. & Amir, R. (1993), 'Factors Influencing Child Mortality in Bangladesh and Their Implications for the National Health Programme', *Asia-Pacific Population Journal* **8**(3), 31–46.

Keyfit, N. (1977), *Applied Mathematical Demography*, John Wiley, New York, United States.

Krishnan, P. & Jin, Y. (1993), 'A Statistical Model of Infant Mortality', *Janasamkhya* **11**(2), 67–71.

Pathak, K. B., Pandey, A. & Mishra, U. S. (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**(1), 15–24.

Preston, S. H. & Palloni, A. (1977), 'Fine-Tuning Brass Type Mortality Estimates with Data on Ages of Surviving Children', *Population Bulletin of the United Nations* (10).

Ronald, D. L. & Lawrence, R. C. (1992), 'Modeling and Forecasting U.S. Mortality', *Journal of American Statistical Associations* (87), 659–675.

Sastry, N. (1997), 'A Nested Frailty Model for Survival Data, with an Application to the Study of Child Survival in North East Brazil', *Journal of the American Statistical Association* **92**.

Srivastava, S. (2001), Some Mathematical Models in Demography and Their Application, Doctoral thesis, Banaras Hindu University, Varanasi, India.

Thiele, P. N. (1972), 'On Mathematical Formula to Express the Rate of Mortality Throughout the Whole of Life', *Journal of Institute of Actuaries* (16), 313.

UNICEF (1987), The State of the World's Children, Technical report, Oxford University Press.