THE INTERNATIONAL JOURNAL OF
SCIENCE & TECHNOLEDGE

Goodness of Fit of Finite Range Reliability Model for Testing Under-reporting of Infant Deaths in India

Niyati Joshi
Deputy Director (Statistics), Ministry of Commerce and Industry
Government of India, New Delhi, India
Mahesh Nath Singh
Senior Research Officer
Public Health Foundation of India at National Institute of Medical Statistics (ICMR)
Ansari Nagar, New Delhi

Abstract:
Developed countries are coupled with better recording and registration system of their vital events where most of the infants deaths are biological in nature. Contrastingly in countries like India, infant deaths are high due to environmental causes coupled with misconceptions due to lack of knowledge, digit preferences for age data responses and poor registration system for lack of infrastructural and resourceful manpower availability which leads to misreporting/underreporting and non-sampling errors in surveys for data on infant deaths. In this light, this paper is an attempt to fit a reliability model proposed by Mukherji and Islam (1983) for estimating misreporting/underreporting of deaths below the first year of life in two states, Kerala and Uttar Pradesh (India), diametrically opposite on the various demographic indicators especially Infant mortality rate. Data from Indian National Family Health Survey (NFHS) (1992-93) has been utilized to compute infant deaths by computing probability distribution function for Uttar Pradesh and Kerala and fitted into the model. There is a good fit for Kerala compared to Uttar Pradesh. United States Vital Statistics (1960) has also been considered for validation of the results. By examining the distribution of Expected and Observed deaths, we find that there is underreporting/misreporting at different months of age during infancy. The study reveals that there are more environmental problems along with underreporting of infant deaths, digit preference and poor registration system in Uttar Pradesh than Kerala.The study suggests that policy steps should be taken to improve physical environmental conditions and reporting and registration system for infant deaths in Uttar Pradesh.

Key words: *Infants, Misreporting, Underreporting, Finite Range, Reliability Model, India*

1. Background And Introduction

Vital events like births and deaths registration suffer from omission and underreporting. Developed countries are coupled with better recording and registration system of their vital events where most of the infant deaths are biological in nature. Contrastingly in societies like India, infant deaths are high due to environmental causes coupled with misconceptions due to lack of knowledge, digit preferences for age data responses and poor registration system for lack of infrastructural and resourceful manpower availability which leads to misreporting/underreporting and non-sampling errors in surveys for data on infant deaths. Partly with the idea to overcome this breakdown, supposed by defective data, Government of India in 1960 introduced sample registration system. Although the reliability and quality of data on births and deaths in the civil registration system have considerably improved over time, they still suffer from considerable degree of errors of mismatching and omissions. It is observed in retrospective surveys that events are misreported partly due to ignorance and partly due to the tendency of the digit preferences of the respondents. Thus the data on deaths collected through vital registration system or sample registration system or retrospective surveys suffer from one defects or another as mentioned above. To understand and overcome this problem of underreporting/misreporting, an attempt has been made to develop and fit suitable models to data on age distribution of deaths.

In this paper, an attempt has been made to apply a model (Mukherjee and Islam, 1983) which will take into account the nature of declining tempo of deaths by age during first year of life. The model is used to give a functional shape to the phenomenon on distribution of infant deaths at different months during infancy.

2. Literature Review

Very few studies related to the age patterns of mortality have developed models of mortality which are mathematical expressions for graduation of age patterns of mortality. They are mostly suitable to graduate rates by age after thirty. Theile (1972) has proposed seven parameter model to capture age pattern of mortality for the whole age range. He has considered three different factors, exponential in nature,represnting young, adult and old age mortality pattern. Later Heligman and Pollard (1980), have developed eight parameter model and tested the same on Australian data at different points of time. They also gave three different factors like Theile. Even in other Countries, this misreporting is found. Caution is urged in interpreting data, since this data and any retrospective data is flawed by errors of recall and misreporting of date of birth and age of death. Child mortality rates ranges widely from 247/1000 in Mali to 32/1000 in Trinidad and Tobago. Infant mortality rates have a narrower range of 144/1000 in Liberia to 25/1000 in Sri Lanka. The widest variation in rates is in the age group 1-4 years (159/1000 in Mali to 4/1000 in Trinidad and Tobago). Mortality rates tend to be highest and the age pattern of mortality is the most distinct in Africa. Mortality declines over a 10 year period are evident for all countries; declines are the most rapid in the Near East/North Africa. Distinct age patterns in Africa show very high mortality for children aged 1-4 years of under 100/1000. The expected pattern of excess male mortality are evident (31percent neonatally, 6percent postneonatally, and 0percent among those aged 1-4 years). The effects of birth order and maternal age varied widely by region and age of the child (Sullivan, Rutstein and Bicego (1994). Regional Variations in Age pattern of Mortality among infants is found (NHFS, 1992-93, IIPS). Therefore, an investigation to understand the pattern of under-reporting of infant deaths in Indian context would be very important (Ramdev, 1998).

3. Need of the Proposed Study

It is well known that among the many available indicators of socio-economic development, infant mortality rate (IMR) is used as a sensitive and powerful index of development. Reduction in IMR is also known to reduce the fertility as probability of survivorship of children increases. It is also used to compare health situations among different communities. Infant deaths occur due to endogenous and exogenous cause (Pathak and Ram, 1992).Endogenous causes are biological and exogenous causes are environmental. As it has been observed in developing countries that infant deaths registration is a subject to error of misreporting which distorts the distribution of deaths by age during infancy. Hence a method for redistribution of infant deaths during the first year of life is important to get a clear picture on the number of deaths at various age points under one year. Keeping in view the above discussion, the following objectives have been framed:

4. Objectives

- To fit the proposed model for distribution of deaths among infants in two states, Uttar Pradesh and Kerala, India based on the NFHS (1992-93) data.
- To validate the model for the States of Kerala and Uttar Pradesh based on the NFHS (1992-93) data.

5. Data And Methodology

5.1. Source of Data

The data for the present study has been taken from the Vital Statistics (1960) of the United States of America and from the raw data of National Family Health Survey (1992-93) for the States of Kerala and Uttar Pradesh.

5.2. Finite Range Model

The model proposed by Mukherji and Islam (1983) with the use of reliability theory is as follows:

The probability density function (pdf) is

$f(x)=(p/x^*)(x/\beta)^p$

$0<x<\beta;0<p<1$ and $\beta>0$…………………………..(1)

f(x)=Prob of deaths distributed among total infant deaths by age

where β is the scale parameter and p is the shape parameter of the model.

The cumulative distribution function gives the cumulative proportion of deaths upto a desired age point x

$F(X<x)=(x/\beta)^p$

$0<x<\beta;0<p<1$ and $\beta>0$ ………………………(2)

6. Assumptions

- Effects of Seasonal variations on deaths at different ages, is independent of age during infancy.
- No misreporting of neonatal deaths.
- No over reporting or underreporting or misreporting of total deaths as far as the first year of life is concerned.

The application of the above model to the data on infant deaths by age has been developed by using the above mentioned assumptions.

The methodology involves the following steps:

Step 1

Let D and N be the total number of infant deaths and neo-natal deaths reported in a geographical area during the reference period.

$R=D/N$ …………………………………..(3)

Where R denotes the proportion of neo-natal deaths among the total infant deaths.

Step 2

By fixing β as 12 months, using data available as 'reported age of deaths in moths'.

By using equation (2),we can estimate the total proportion of deaths under the age point 'x'

$F(X<x)=(x/\beta)^p$

where
$0<x<12$……………………….(4)
In equation (4),if we put x=1 then we get the total proportion of deaths in the first month of age to total infant deaths.
$F(X<x)=(x/12)^p$
$0<x<12$ …………….(5)
By combining equation (4) and (5),we get;
$R=(1/12)^p$ ……………….(6)
Taking logarithm on both sides for equation (6),we get;
$Log=Log[(1/12)]^p$
$LogR= pLog(1/12)$
$P=logR/Log(1/12)$ ……………………(7)

Step 3
By taking equation (1) as
$f(x)=(p^{\wedge}/x)(x/\beta)^{p^{\wedge}}$
Where p=p^; $0<x<12$ …………………………………………..(8)
Then equation (2) becomes,
$F(X<x)=(x/12)^{p^{\wedge}}$
$0<x<12$ ……………………….(9)
By putting x=1,2,3……..12 in equation (9),we get cumulative proportion of deaths upto the first month of age, second months of age etc:

Step 4:
Proportion of death during a particular month of age
=proportion of deaths upto (x+1)-proportion of deaths upto age x
$=F(X<x+1)-F(X<x), x>1$………………………………………..(10)
Proportion of deaths during first month=$F(X<1)$………………………………(11)
By multiplying equation (11) with 'D' we get estimated number of deaths for the first months,
i.e.$d1=D*F(X<1)$ ………………………………………………………….(12)
Similarly
$dx=D*[F(X<x+1)-F(X<x)]$,
$x>1$

Chi-Square Test of Goodness of Fit
In this paper, Chi Square test for goodness of fit has been used to find out the homogeneity in the distribution of observed and expected deaths.
$\chi^2 = \Sigma(Oi-Ei)^2/Ei$, where O:Observed Deaths, E: Expected Deaths

7. Analysis and Findings
In this paper, the reliability model is proposed to estimate the infant deaths in the states of Kerala and Uttar Pradesh. The model has been verified before applying to the Indian States by fitting to the United States Vital Statistics data on infant deaths which is assumed to be reliable. Expected Deaths and Model parameters were first estimated for the purpose. The fit is fairly good as reported and expected deaths are fairly good during each months of infancy and are closer. The Chi-Square test for goodness of fit is fairly good (Table 1).

Age at Deaths (MONTHS)	Number of Deaths	F(X<x)	Probability of Death	Expected Death	Adjusted Deaths
0-1	1873	0.7193	0.7193	1873.00	1873
1-2	171	0.7885	0.0692	180.31	180
2-3	133	0.8321	0.0436	113.42	113
3-4	101	0.8644	0.0324	84.25	84
4-5	77	0.8904	0.026	67.6	68
5-6	575	0.9122	0.0218	56.74	57
6-7	47	0.931	0.0188	49.05	49
7-8	39	0.9477	0.0166	43.31	43
8-9	33	0.9626	0.0149	38.84	39
9-10	27	0.9761	0.0135	35.26	35
10-11	23	0.9885	0.0124	32.33	32
11-12	23	1.0000	0.0115	29.87	30
Total	2604			2604	2604

R	LogR	P	Chi Square (Cal)	(Tab)(10%)	
0.7173	-0.1431	0.1326	16.1000	16.8	

Table 1: Fitting For USA (Vital Statistics, 1960)

The model has also been tested for the states of Kerala and Uttar Pradesh in India as these two states are diametrically opposite in demographic parameters.

The fit for Kerala shows that the expected deaths is quite close to the reported deaths. The calculated Chi-square values are within the range of the tabulated value which shows that the fitted model is fairly good, though there are variations in the pattern of distribution of reported and expected deaths at each age (Table 2).

Age at Deaths (MONTHS)	**Number of Deaths**	**F(X<x)**	**Probability of Death**	**Expected Death**	**Adjusted Deaths**
0-1	93	0.7381	0.7381	93	93
1-2	7	0.8033	0.0652	8.22	8
2-3	5	0.8442	0.0408	5.14	5
3-4	5	8744	0.0302	3.81	4
4-5	1	0.8985	0.0242	3.05	3
5-6	1	0.9188	0.0202	2.55	3
6-7	4	0.9363	0.0175	2.2	2
7-8	1	0.9517	0.0154	1.94	2
8-9	2	9617	0.0138	1.74	2
9-10	0	0.978	0.0125	1.58	2
10-11	6	0.9894	0.0115	1.44	1
11-12	1	1.0000	0.0106	1.33	1
Total	126			126	126
R	LogR	P	Chi Square (Cal)	(Tab) (3.5%)	
0.7381	-0.1319	0.1222	1.9658	7.815	

Table 2: Fitting For KERALA (NFHS, 1992-93)

While applying the model to the state of Uttar Pradesh (Table 3), the distribution of observed and expected deaths shows a high variation except for the first month. For Uttar Pradesh, the calculated value of Chi-Square falls in the region of rejection. Hence, calculated value of chi-square is insignificant. Very lager variations at each month of the infancy in the reported and expected deaths are found.

Age at Deaths (MONTHS)	**Number of Deaths**	**F(X<x)**	**Probability of Death**	**Expected Death**	**Adjusted Deaths**
0-1	1395	0.623	0.6230	1395	1395
1-2	126	0.7109	0.879	196.81	197
2-3	111	0.768	0.0571	127.76	128
3-4	86	0.8112	0.0432	96.82	97
4-5	57	0.8465	0.0352	78.84	79
5-6	52	0.8764	0.0299	66.95	67
6-7	109	0.9025	0.0261	58.44	58
7-8	33	0.9257	0.0232	52.03	52
8-9	77	0.9467	0.021	47.01	47
9-10	57	0.9659	0.0192	42.95	43
10-11	76	0.9836	0.0177	39.6	40
11-12	60	1.0000	0.0164	36.79	37
Total	2239			2239	2239
R	Log R	P	Chi Square(Cal)	(Tab)(9.5%)	
0.623	-0.2055	0.1904	160.79	16.919	

Table 3: Fitting For UTTAR PRADESH (NFHS, 1992-93)

8. Conclusion and Policy Implications

Under-reporting and mis-reporting of infant deaths in both Uttar Pradesh and Kerala have been observed while level of misreporting in Kerala is lower than Uttar Pradesh. This high under-reporting in Uttar Pradesh may be due to the lower status of socio-economic conditions, lack of knowledge and particular way of functioning of government machinery responsible for survey or data collection regarding infant deaths. Also, there may be cultural factors affecting the reporting of infant deaths along with the response bias for such sensitive data sets. Socio-economic, general education and educational for attitudinal change and population education for reorienting the cultural value for development alongwith improvement in the physical conditions may be provided. Improved training and data collection facilities may be provided to the functionaries involved in the collection of vital statistics.

9. References

1. Heligman,L and Pollard, J. H.(1980): “Age Patterns of Mortality” Journal of Institute Actuaries, Vol 107, pp: 49-80
2. Mukherjee,S.P. and Islam,A. (1983): “ A Finite Range Dsitribution of Failure Time”; Naval Research Logistic Quarterly, Vol. 30, pp: 487-491
3. National Family Health Survey (1992-93): “National Family Health Survey”,IIPS and ORC Macro, Mumbai
4. Pathak,K.B. and Ram,F. (1992): “Techniques of Demographic Analysis “Himalaya Publishing House”, New Delhi
5. Ramdev,K. S. S. (1998): “Estimaiton of Infant Deaths:A Reliability Approach”IIPS,Mumbai
6. Sullivan JM; Rutstein SO; Bicego GT **(**1994**):** “Infant and child mortality” Calverton, Maryland, Macro International, 1994 Jun. VI, 57 p. Demographic and Health Surveys Comparative Studies No. 15
7. Theile,P.N (1972): On A Mathematical Formaula to Express the rate of Mortality throughout the Whole of Life, Journal of Institute of Actuaries, Vol 16, pp:313-329
8. Vital Statistics (1960): Vital Statistics of the United States accessed from http://www.cdc.gov/nchs/data/vsus/vsrates1940_60.pdf dated April 10, 2014