Journal of Research in International Business and Management (ISSN: 2251-0028) Vol. 1(2) pp. 164-170 August 2011
Available online @http://www.interesjournals.org/JRIBM
Copyright ©2011 International Research Journals

Review

Developing a model for estimating infant mortality rate of Nigeria

***Mojekwu J.N. and Ajijola L.A**

Department of Actuarial Science and Insurance, University of Lagos, Akoka, Yaba, Nigeria.

Accepted 09 August, 2011

One of the Millennium Development Goals is the reduction of infant mortality by two thirds between 1990 and 2015. Although there has been a substantial reduction in infant and child mortality rates in most developing countries in the recent past, infant mortality remains a major public health issue in developing countries where it is estimated that over 10 million preventable child deaths occur yearly. Progress in infant mortality reduction remains unacceptable in Sub-Saharan Africa. With special reference to Nigeria, the giant of Africa, available statistics suggest that infant mortality levels continue to be high and exhibit wide geographic disparities. This study therefore attempts to estimate infant mortality rate in Nigeria using linear regression model. Crude death rate (CDR) has been selected as the minimum relevant parameter (independent variable) needed for estimating Infant Mortality Rate (IMR) which is the dependent variable because it represents the 'end result' of development. The IMR derived model is checked for adequacy by comparing the estimates of the present study with the estimates from other sources. The diagnostic test show that the regression model derived is quite adequate and reflects the true picture of Nigerian Infant Mortality Rate patterns.

Keywords: Infant mortality rate, crude death rate, simple regression model and diagnostic checking.

INTRODUCTION

Young children are in many ways the most vulnerable group to adverse effects of environmental health (Nuria, 2003). They are sensitive not only to conditions in their immediate environment after birth, but also to the pre- and post-natal health of their mother, and the quality of the health support services. Information on infant mortality thus provides both a specific indication of the health status of young children, and a more general indicator of the overall quality of the health conditions and effectiveness of health facilities. Infant mortality rate (IMR) is an important indicator of social development of a nation. It is widely used for assessing socio-economic and health situation in developing countries (Chandra, 1972: 1977; Jain and Visaria, 1988). Measurement is a fundamental aspect of research in the area of infant mortality. If vital registration is complete, IMR for each year can be calculated in the conventional manner directly from the system's data (Hill 1991:369). Unfortunately, complete vital registration system is practically non-existent in Nigeria (Shagodoyin et al., 2009). Most of the demographic studies on mortality of Nigerians are based on indirect estimates using indirect techniques proposed by Brass (1964), Trussell (1975), Feeney (1980), and Palloni and Heligman (1986) for estimating IMR for Nigeria using the census or survey data and from retrospective survivorship data. The estimation of IMR by these techniques needs accurate birth history data to be collected from census or survey; reliable and adequate age patterns of child mortality for selecting an appropriate method and model life table; and lastly, many assumptions to be satisfied by the population under study. But the irony is that the birth and death data collected from the censuses or surveys of Nigeria are highly inaccurate (Adewuyi and Feyisetan, 1984). Further, many assumptions underlying the models are unjustifiable in the population under study (Hill and Yazbek, 1993:1997). Moreover, there is no reliable age distribution of mortality for Nigeria to use in deciding the

*Corresponding Author E-mail: jnmoje@yahoo.com

right family that fits. Apart from this, Nigeria is a country which has fails to use her enormous resources to uplift the living standard of her citizen where 64 percent of people still live below the poverty line, and about 28 percent the total population are unable to read or write (SOWC, 2009).
Keeping in view the socio-economic and demographic realities of our own country, this paper presents a simple regression model for estimating IMR from the minimum relevant parameter. Crude death rate (CDR) has been selected as the minimum relevant parameter needed for estimating IMR because it represents the 'end result' of development; can be easily obtained from either suitable models or various publications; and more importantly, it is strongly correlated with the level of IMR (Arriaga, 1994). The model is applied to obtain the estimates of IMR for Nigeria, and its validity is ascertained by comparing the estimated IMRs for Nigeria with the other estimates available in the country; and by computing relevant tests for diagnostic checking for model adequacy.

MATERIALS AND METHODS

The proposed methodology of estimation is based on simple regression approach described elsewhere (Kumar 1981; Aryal and Gautam, 2001; Singh, 2003). The methodology of estimation developed here follows the usual path of establishing the relationships between the dependent variable, which in this case is the IMR and the independent variable, herein identified as CDR. Several empirical studies show almost a linear relationship between IMR and CDR. Therefore, it was decided to fit a regression model of type using:

$$\textbf{IMR} = \alpha + \beta X + \varepsilon \text{ ---------(1)}$$

Where IMR = the Infant Mortality Rate (per 1000 live births); X = Crude Death Rate; CDR (per 1000 population); ε is a random error term; and α and β are parameters to be estimated. All calculations were performed using statistical packages SPSS.
The next step is to estimate the value of the parameters. For this purpose, the regression model is fitted in by using the following set of data extracted from the United Nations Population Division (UNDP 2008).
Using the values in the table above, simple linear regression approach gives the regression model below:

$$\textbf{IMR} = - 5.251 + 6.887X \text{ ------------ (2)}$$

The R-Square =99.7% and Std. Error of the Estimate= 1.98713.

Diagnostic Checking for Model Adequacy And Discussions

A model that fails in diagnostic checking for model adequacy will always remain suspect and little faith can be put in the results (Kerlinger, 1998). Therefore, it is essential that the model fitted for estimation purposes should satisfy the important tests of model adequacy. In this study, diagnostic checking for model adequacy is done by applying the model in Nigeria's context; comparing the estimated IMRs for Nigeria with the other estimates available for the country from different sources (FOS 1966; NFS 1981; Bamgboye, 1986; NPC [PES] 1991; Shangoyin, 2001; SOWC, 2009); and by computing relevant tests for model adequacy described elsewhere. The comparison of the estimated IMRs for Nigeria with the other estimates available for the country is presented by the values in the following table and figure

The above table clearly shows declining trends of IMRs over the periods as are usually expected. More importantly, it shows close agreement between the estimated IMRs and the other estimates available for the country from different sources over a wide range of periods. However, it also shows a weak agreement between the estimated IMRs and other estimates especially, in 1991. This may be attributed to the limitations of data sources. In summary, the model seems to provide better estimates for distant past than more recent period. Diagnostic checking for model adequacy can also be done by computing major relevant tests described elsewhere.

Diagnostic procedures are intended to check how well the assumptions of multiple linear regressions are satisfied. Infringement of these assumptions cast doubt on the validity of the conclusions drawn on the basis of the results. A number of checks and tests help us to ensure that analysis has proceeded within the bounds of the basic assumptions. The checks fall into two groups viz. (i). Those for checking the pattern of the residuals by means of residual plots, and (ii). Those for individual data points.
Residual plots are the best single check for violation of assumptions, such as:
(i). Variance not being constant across the explanatory variables.
(ii). Fitted relationships being non-linear.
(iii). Random variation not having a normal distribution

The following tables and figures present the results of major relevant tests computed for diagnostic checking for model adequacy:

The coefficient of determination (R^2) is computed for testing the goodness of fit. For the given set of data, the computed value of coefficient of determination R^2 (=

Table 1: Estimates of CDR and IMR for Nigeria used for Fitting the Equation (1)

YEAR	X (=CDR)	IMR
1950 – 1955	27.7	186
1955 – 1960	26.6	178
1960 – 1965	25.3	169
1965 – 1970	24.2	162
1970 – 1975	22.9	153
1975 – 1980	20.7	133
1980 – 1985	20.1	131
1985 – 1990	20	132
1990 – 1995	19.7	133
1995 – 2000	18.8	127
2000 – 2005	17.4	114
2005 - 2010	16.5	109

Source: UNDP (2008)

Table 2: Comparison of the Estimated IMRs for Nigeria with other Estimates

Estimates of other Sources			Present Study
Years	**CDR**	**IMR**	**IMR***
1965[i]	22.2	141.4	147.6
1979[ii]	17	84.8	111.8
1985[iii]	16	72.4	104.9
1991[iv]	7.1	93	43.6
2001[v]	18	98.8	118.7
2007[vi]	17	97	111.8
2009[vii]	16.88	94.35	111.0

Notes

I. Refers to the estimates of FOS Demographic Survey (1966)
II. Refers to the estimates of Nigeria Fertility Survey (1981)
III. Refers to the estimates of Bamgboye (1986)
IV. Refers to the estimates of NPC Post Enumeration Survey (1991)
V. Refers to the estimates of Shangodoyin (2001)
VI. Refers to the estimates of UNICEF State Of The World Children 2009
VII. Refers to the estimates of Central Intelligence Agency (CIA)

* Refers to the estimates of Present Study

Table 3: Test for Goodness of Fit

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics			Durbin-Watson
					R Square Change	F Change	Sig. F Change	
1	.997(a)	.994	.994	1.98713	.994	1742.207	.000	1.357

a Predictors: (Constant), CDR
b Dependent Variable: IMR

Table 4: Test of Significance (ANOVA)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6879.430	1	6879.430	1742.207	.000(a)
	Residual	39.487	10	3.949		
	Total	6918.917	11			

a Predictors: (Constant), CDR
b Dependent Variable: IMR

Table 5. Coefficients (a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-5.251	3.619		-1.451	.178
	CDR	6.887	.165	.997	41.740	.000

a Dependent Variable: IMR

Table 6. Residuals Statistics (a)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	108.3897	185.5274	143.9167	25.00805	12
Std. Predicted Value	-1.421	1.664	.000	1.000	12
Standard Error of Predicted Value	.595	1.150	.791	.188	12
Adjusted Predicted Value	108.1676	185.2893	143.8746	25.00194	12
Residual	-4.31634	2.76951	.00000	1.89465	12
Std. Residual	-2.172	1.394	.000	.953	12
Stud. Residual	-2.277	1.503	.010	1.011	12
Deleted Residual	-4.74149	3.21912	.04202	2.12978	12
Stud. Deleted Residual	-3.112	1.620	-.049	1.216	12
Mahal. Distance	.070	2.769	.917	.894	12
Cook's Distance	.000	.255	.059	.084	12
Centered Leverage Value	.006	.252	.083	.081	12

a Dependent Variable: IMR

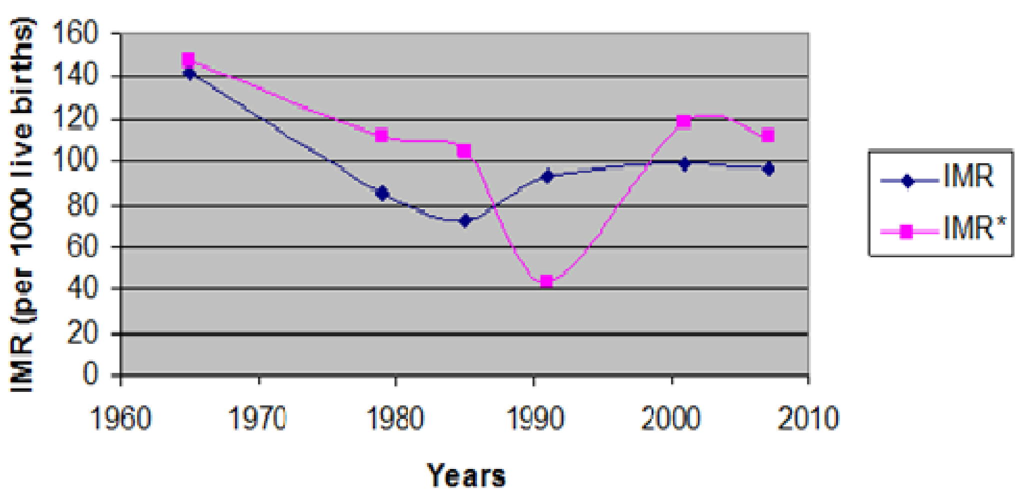

Figure 1. Comparison of the estimated IMRs for Nigeria with other Estimates

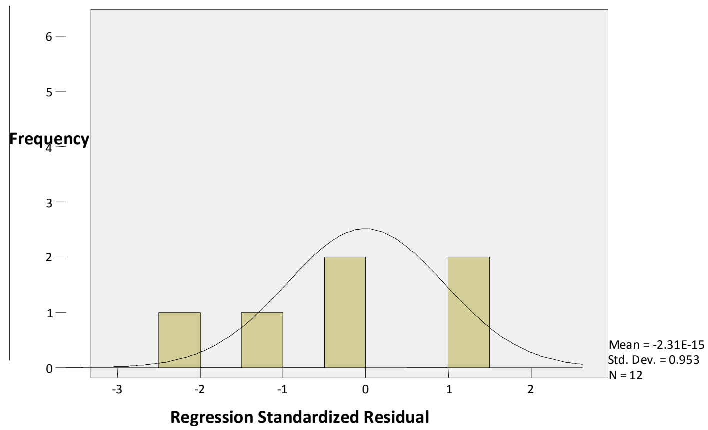

Figure 2. Histogram of residuals. Histogram Dependent Variable: IMR

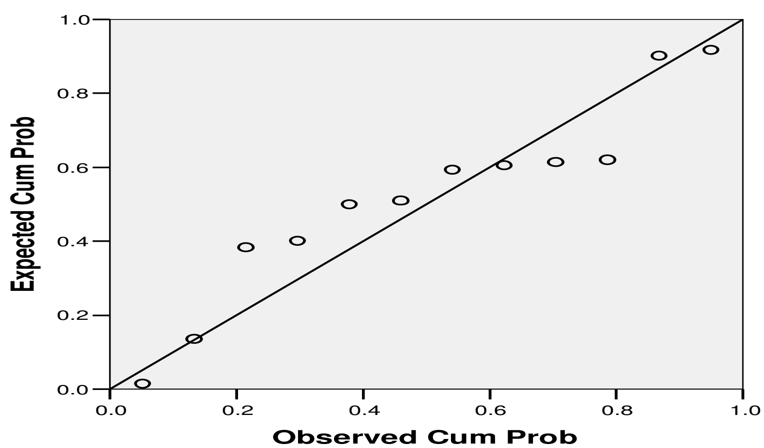

Figure 3. Normal Probability Plot of the Residuals. Normal P-P Plot of Regression Standardized Residual Dependent Variable :IMR

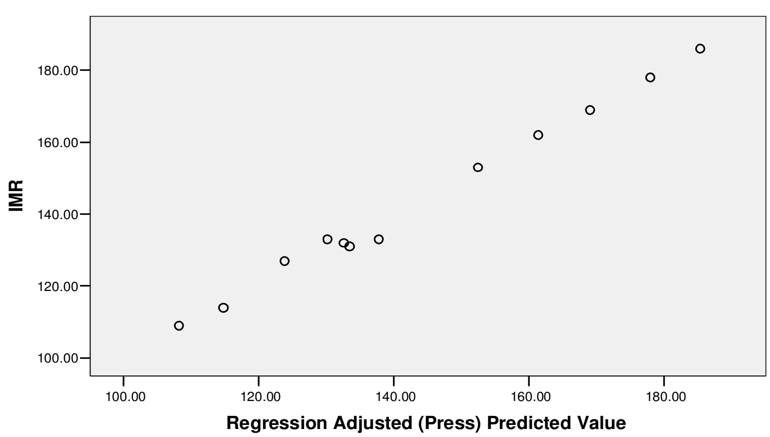

Figure 4. Residual Versus the Fitted Values. Scatterplot Dependent Variable: IMR

99.4%) is very high which indicates goodness of it as 99.4% of the variation in IMR among the periods, appears to be explained by the variation in the CDRs. Similarly, the smaller value of computed S.E. (Y) [= 1.98713] indicates the higher reliability of the model. The goodness of fit of a regression model is mostly affected by the estimated values of parameters. Similarly, the estimated parameters may be considered significant as they satisfied the't' test. For instance, the parameter [CDR] may be considered significant at t= 41.740 and small significant level. Another way to test the significance of the regression model is to compute the 'F' test. The F statistics represents a test of null hypothesis that the expected values of the regression coefficients are equal to each other and that they equal zero. If the null hypothesis were true, then that would indicate that there is not a regression relationship between the dependent variable and predictor variable(s). Since there is only predictor (CDR), then it could be use to predict the dependent variable, (IMR), as is indicated by a large F value (=1742.207) and a small Mean Square Error (=3.949).

Another check for co- linearity is the Durbin–Watson statistic. Normally its value should lie between 0 and 4. A value close to 2 suggests no correlation; one close to 0 negative correlations, and a value close to 4 positive correlations.

The presence of autocorrelation is a serious problem and therefore, Durbin – Watson (D-W) test is computed for detection of autocorrelation. The test leads us to accept the null hypothesis of no serious correlation beyond what is induced by first order autoregressive process. In this case it value is 1.357 so it indicates that doesn't exists autocorrelation of first order. Thus the model fits the data relatively well.

Residual is the difference between the calculated mean values of Y (IMR) (this is also the fitted value as determined by the regression line) and the actual observed value of Y for a given value of the explanatory variable. Thus the residuals tell us how well or otherwise the model fits the data. One problem with using residuals is that their values depend on the scale and units used. Since the residuals are in units of the dependent variable IMR there are no cut-off points for defining what is a “large” residual. The problem is overcome by using standardized residuals. They are calculated by residual ÷ standard error of the residual. The standard error of each residual is different, and using standardized residuals helps one to get round the problem. For example, a standardized residual of 1 means that the residual is 1 standard deviation away from the regression line, a standardized residual of 2 means that the residual is 2 standard deviations away from the regression line and a standardized residual of 0 means that the point falls on the regression line. In our own case the standard residual has mean 0 and standard deviation 0.953 which imply that the point falls on the regression line.Observations with standardized residuals exceeding 3 require close consideration as potential outliers. Plotting residuals on the Y axis against fitted value(s) on the X axis is a useful diagnostic procedure. If the model is appropriate for the data the plot should show an even scatter. Any discernible pattern in the plot means that the regression equation does not describe the data correctly, since pattern forms when the residuals are unevenly distributed about the regression line. Outliers may also get identified in such a plot. In addition one also needs scatter plots with standardized residuals on the vertical axis and predictor variable on the horizontal axis. These should show the same amount of variation in the standardized residuals for all the predictor (CDR). Looking at the Histogram of the Residuals (Figure 2) to check the assumption about normal distribution it could be seen that though negatively skewed the assumption about normal distribution could still be made. The normal Probability Plot of the residuals (figure 3) serves the purpose by showing a straight line for normal distribution and a plot of residual against fitted values (figure 4) shows that the fit is uniformly good for value of IMR.

CONCLUSIONS

The advantages of the indirect techniques in mortality estimation cannot be overemphasized in developing countries like Nigeria. The proposed model is very simple and easy to apply; does not need census or survey data and model life tables for estimation of IMR; and gives approximately reliable estimates for Nigeria. The results indicate that the model is effective in providing approximately reliable estimates of IMR for Nigeria during the last few decades. The model seems to provide comparatively better estimates for distant past than for more recent periods. However, the model seems to be affected by accuracy of data and age structure of the population under study.

Conclusively, the model may be considered suitable for estimating IMR for Nigeria for few more decades.

REFERENCES

Adewuyi AA (1984). “Correlation of Infant Mortality; Empirical Evidence from Ile-Ife, Nigeria”. Department of Demography and Social Statistics, O.A.U. Ile-Ife, Nigeria. Pp 31-36

Arriaga EE (1994). Population Analysis with Microcomputers, Presentation of Techniques. Washington D.C.: United States Bureau of the Census. 1: 66-70

Aryal JP, Gautam A (2001). Quantitative Techniques, New Hira Book Enterprises. Pp 113-140

Brass W (1964). "Uses of Census and Survey Data for Estimation of Vital Rates," Paper Presented at the African Seminar on Vital Statistics, Addis Ababa. Pp 11-15

CIA World Fact Book, (2009). www.theodora.com.

Chandra SS (1972). Infant Mortality, Population Growth and Family Planning in India. London: George Allene and Unwine Ltd. Pp 142-155

Federal Office of Statistics of Nigeria (FOS) Demography Survey of 1966. FOS publications. Pp 212-222
Radhakrishna GC (1992/93). "Population Data Quality," Tathyank
Gatibidhi (Four Monthly Statistical Bulletin), Kathmandu: Central Bureau of Statistics, Kathmandu, 13 (2):6-9.
Hill K (1991). "Approaches to the Measurement of Childhood Mortality:
A Comparative Review," Population Index, 57(3):368-382.
Hill K, Yazbeck A (1993). Trends in Child Mortality, 1960-1990: Estimates for 83 Developing Countries. New York: United Nations Children's Fund.
Jain AK, Visaria P (eds. 1988). Infant Mortality in India: Differentials and Determinants. New Delhi: Sage Publications. Pp 113-119
Karki YB (1989). "Vital Registration System in Nepal." A paper presented at National Seminar on Social Statistics, Central Bureau of Statistics, pp 20-23 Kathmandu.
Kerlinger FN (1998). Foundations of Behavioural Research. Delhi: Surjeet Publications. Pp 21-33
Kumar R (1981). "Use of Regression Models in Determining Accurate Vital Statistics in Developing Countries: A Methodology," Demography India, 10:216-233.
Nigeria Fertility Survey (NFS 1981). National Population Commission. Pp 13-27
National Population Commission (NPC)(1991). Post Enumeration Survey (PES). Pp 48-51
Nuria CV (2003). "Socio-Economic and Geographical Determinants of Infant Mortality. Pp 15-18 in Spain: 1975-2000". M.Sc. Dissertation, Faculty of Actuarial Science and Statistics, Cass Business School, City of London.
Palloni A, Heligman L (1986). "Re-estimation of Structural Parameters to Obtain Estimates of Mortality in Developing Countries." Population Bulletin of the United Nations, 18:10-33.
Pant PD (1996). "The Quality of the Reported Age of Mother, Date of Birth and Death of Children in the 1976 NFS and the 1986 NFFS Data," Nepal Population J. 5(4): 1-8.
Shangodoyin DK (2008). "Nigeria Population Dynamics and its implication for Development". Medwell J. ISSN:1818-6125. Pp 50-57
Singh ML (1979). "Population Dynamics of Nepal: An Analytic Study." Unpublished Ph.D. Thesis, Tribhuvan University, Kathmandu. Pp 62-71
Singh ML (2003). Understanding Research Methodology (Third Edition). Kathmandu. Pp 35-45
Trussell TJ (1975). "A Re-estimation of the Multiply Factors for the Brass Technique for Determining Childhood Survivorship Rates," Population Studies, 29(1):97-108.
UNDP. 2002. "Poverty Reduction and Governance," Human Development Report 2001. Kathmandu. Pp 15-25
UNICEF. 2001. The State of World's Children. New York
UNICEF. 2001. The State of World's Children. New York: Oxford University Press.
UNICEF. 2009. The State of World's Children. New York: Oxford University Press.