Analysis of Child Mortality for Muslim women of Uttar Pradesh, India: an application of Bayesian Frailty Model

G.P. Singh[1], A. Tripathi[1*], S.K.Singh[2], Umesh Singh[2] and

O.P. Singh[3]

[1] Department of Community Medicine and DST-CIMS ,BHU, Varanasi.
[2] Department of Statistics and DST-CIMS, BHU, Varanasi.
[3] Department of Statistics, U.P. College, Varanasi.

Abstract: In this paper an attempt has been made to analyze the child mortality by use of a hazard model in Bayesian environment, family effect through multiplicative random effect is also incorporated in the model. For fitting this model real data has taken from District Level Household and Facility Survey (DLHS)-3. The largest state (in population) of India i.e. Uttar Pradesh data is taken for analysis. Deviance information criteria are used for comparison of models. It found that the model with family frailty gives better fit. All the analysis is performed in winBUGS software, which is used Markov chain monte carlo simulation under gibbs sampling.
Key words: Bayesian Environment, Child Mortality, Deviance information Criteria, Gibbs sampling[1]

1. Introduction

Studies on child mortality are being performed since long time. Different types of analysis procedures have been used for drawing conclusion about child mortality estimation, its prediction and to find out various factors which are responsible to bring change in child mortality. As well as child mortality is concerned, it is a good indicator of social and health status of society. It also determines the life expectancy which is an important measure of well-being and development in poor countries (Sen, 1998). In all types of statistical analysis modeling is very dominant area in child mortality. In past decades binomial, negative binomial, exponential and weibull models etc. were used to calculate the child mortality and child survival, but They did not find out factors which are responsible for child mortality or survival, but in recent decades after development of better computational facilities various new techniques like logistic regression model, cox-proportional hazard model and various types of generalized linear models are used to find out child mortality or survival estimates and the factors which are responsible to bring changes in child mortality or survival. Mainly Socio economic conditions and bio-demographic conditions of family are responsible to change in child survival or mortality. Many studies have examined

---

* Corresponding Author

the empirical evidences concerning the influence of the demographic and socio-economic factors influencing child survival (Hobcraft et. al.,1984; Das Gupta, 1990, Uddin et. al 2008), including a number of studies in India(Jain and Visaria, 1988; Singh k k et al,2008 ; Bhattacharyya and Rao ,2009). Besides affecting by such factors, child mortality is also affected by family and community level characteristics (Shastry, 1997) which are immeasurable (genetic and environmental changes), these characteristics are not described properly among the factors which were already found in previous studies. Such unobserved heterogeneity is called as frailty. These frailties also distort the real effect of factors which are associated with child mortality (Vaupel et. al.1979). To find out such frailty effect a number of researches have been done. (Guo ,rodriguez 1992) and (Guo, 1993) used a proportional hazard model to find out family random effect in Guatemala, Curtis et al(1993) also found family random effect in Brazil by use of logistic model. Dasgupta M(1990) found clustering of child death in Punjab, India. Recently in the context of India (M.Kalaivane et al, 2012) used cox- proportional hazard model to family level heterogeneity. They found that mother's age at birth, place of delivery, sex of the baby, composite variable of birth order and birth interval, baby size at birth, and breastfeeding were significant determinants of under-five mortality, after adjusting for the familial frailty effect.

These all authors used the classical estimation procedure with parametric or non parametric approach. But after the discussion of Sinha and Dey (1997) about lacuna in Expectation maximization algorithm which uses for frailty analysis, various authors tend to perform Bayesian analysis for complex survival models. Recently (Gemperli et al 2004) fitted a Bayesian hierarchical geostatistical logistic model to infant mortality risk, by Markov chain Monte Carlo simulation. It confirmed that mother's education, birth order and interval, infant's sex, residence, and mother's age at infant's birth had a strong impact on infant mortality risk in Mali.(Kossi and Hognas2005) have also developed Bayesian frailty model to investigate family heterogeneity in child mortality data from Ivory Coast. They found that birth interval plays stronger role to determine child mortality as well as frailty.

Here we also tried to fit a piecewise exponential hazard model (M. Friedman, 1982) on child mortality with incorporation of frailty in model in India (On Muslim family of a state "Uttar Pradesh"). Bayesian environment was created for the analysis. Due to marriages between cross cousins, it is quite possible that genetic disorder might occur in future leading to child mortality (Park K, 2005), in Muslim family so we take Muslim family here.

2. Data

The analysis of child mortality performed here is used on data from District Level Household and Facility Survey (DLHS-3) which is preceded by DLHS-1 and DLHS-2, conducted in all over India in 2007-2008 on sample basis. Uttar Pradesh is the biggest state (census, 2011) and having poor level of health indicators, is taken here for the analysis. The state UP is mainly constructed by two religions Hindu and Muslim .India's Muslim population is the third largest in the world (Pew, 2009).So for investigating child mortality relationship with certain factors in Muslim population we plan this study. Here we have taken births with following condition: 1-cases of births taken place between 2004-2008 in Muslim female of Uttar Pradesh. 2- Female which have

three births during this period for averaging the birth interval, as well as for limit the no. of children each family at constant that the effect of birth interval and family size(no. of children per family) not grasp the effect of frailty. Then 198 female having 571 births were taken in study in which 75 deaths were observed in the period of study. The four independent variable or factors or covariates is taken in study, mother's education, mother's age, gender of child and family planning practise attitude of parents of child. The variables mother's education, mother's age and gender of child are taken in model guided by (Shastry, 1997; Gemperli et al 2004; Kossi and Hognas,2005; M.Kalaivane et al, 2012). Family planning practise attitude of parents of child is taken as new variable which is able to capture the role of birth interval on child mortality. It is well established fact that the birth interval effects child mortality and determine the no. of children for a family. Twin births and births for which at least any one of the above mentioned independent factors was missing, was excluded from the study.

3. Model

The Bayesian Model (Kossi and Hognas2005)

Let t_{ij} is random survival time of jth child from ith family $\theta=(\boldsymbol{\beta},w_i)$ the unknown parameters of the model corresponding to the data. The parameter w_i represent the family random effect and $\boldsymbol{\beta}= (\boldsymbol{\beta}1, \boldsymbol{\beta}2,....)$ is the vector of fixed effect coefficients.The family random effect is assumed to act on the conditional hazard h(t_{ij}| $\boldsymbol{\beta}$,w_i) in the following multiplicative way:

$$\text{h}(t_{ij}|\boldsymbol{\beta}, \text{w}_i)= w_i\lambda_0(t_{ij})\exp(\boldsymbol{\beta}\text{'X}_{ij})= w_i\lambda_0(t_{ij})\exp(\boldsymbol{\beta}1X1_{ij}+\boldsymbol{\beta}2X2_{ij}+.......) \tag{1}$$

our aim is to find out w_i .

The Bayesian approach updates the prior belief ($\pi(\boldsymbol{\beta},w_i)$) using the data, in order to obtain the posterior distribution (which represent new beliefs after having observed the data).The posterior distribution of ($\boldsymbol{\beta}$,w_i) conditioned on the data is proportional to the product of the likelihood function and the prior distribution.

$$\pi(\boldsymbol{\beta},w_i \mid t_{ij}) \propto \pi(\boldsymbol{\beta},\text{w}_i) \times \text{L}(t_{ij}|\boldsymbol{\beta},w_i) \tag{2}$$

The likelihood function

Let us introduce the censoring indicator δ_{ij} such that δ_{ij} equal 1 if child has died and 0 if not. If f_{ij} and S_{ij} are the density and the survival functions respectively, the following equality holds:

$$f_{ij}(t_{ij}|\boldsymbol{\beta},w_i) = S_{ij}(t_{ij}|\boldsymbol{\beta},w_i)\times\text{h}(t_{ij}|\boldsymbol{\beta}, w_i) \tag{3}$$

The contribution of child j from family i to the likelihood function is its density function if the child dies, and its survival function otherwise:

$$L(t_{ij}, \delta_{ij} | \boldsymbol{\beta}, w_i) = [\ f_{ij}(t_{ij} | \boldsymbol{\beta}, w_i)]^{\delta_{ij}} [S_{ij}(t_{ij} | \boldsymbol{\beta}, w_i)]^{1-\delta_{ij}} =$$

$$w_i^{\delta_{ij}} [\lambda_0(t_{ij}) exp(\boldsymbol{\beta}'X_{ij})]^{\delta_{ij}} exp(-w_i \Lambda(t_{ij})) \tag{4}$$

where $\Lambda(t_{ij}) = \int \lambda_0(t_{ij}) exp(\boldsymbol{\beta}'X_{ij}) dt$ is the integrated hazard function for the fixed effects.

The effect of covariates on child mortality do not have equal importance over the whole period of childhood. Considering it a piecewise exponential baseline hazard the study time period was split into five intervals with cut points at 1.8, 4.0, 12.0 and 24 months. Within each interval I_n the baseline hazard is assumed constant $\lambda(t_{ij}) = \lambda_n$ for $t_{ij} \in I_n$. The likelihood function for general λ(hazard) in case of exponential distribution is defined as:

$$L(\lambda) = \prod_j \lambda^{\delta_j} exp(-\lambda t_j) = \lambda^d \ exp(-\lambda T)$$

where $d = \sum_j \delta_j$ and $T = \sum t_j$

Here d is the total number of observed death and T is the total exposure, giving the total time the sample was known to be at risk of failure. It is easily seen that the maximum likelihood estimate of λ is d/T. If we assume instead that d is poisson, conditional on T with $E(d/T) = T\lambda$, then the likelihood under this sampling model

$$L_p(\lambda) \propto (T\lambda)^d exp(-T\lambda)/d! \ \propto L(\lambda) \tag{5}$$

Since the likelihoods under these two different sampling models are proportional, one can use the likelihoods interchangeably. So by using the property of equation (5). The likelihood function in our problem coincides with that of a Poisson distribution with mean $E_{ijn} \lambda_{ijn}$. In that expression, E_{ijn} denotes the time lived in the interval I_n by the jth child from the ith family and the parameter λ_{ijn} is the corresponding hazard function.

<div align="center">The Prior Specification β₁, β₂, …</div>

The Prior Specification $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, …

$$\boldsymbol{\beta}_k \sim N(0, 1000)$$

$$w_i \sim Gamma(\boldsymbol{\tau}, \boldsymbol{\tau}).$$

The hyperparameters $\boldsymbol{\tau}$ is also assumed Gamma distribution.

<u>The Full Model</u>

A proportional simplification of the posterior density is obtained from (2):

$$f(\boldsymbol{\beta}, w_i, \boldsymbol{\tau}, \{t_{ij}, \delta_{ij}\}) \propto f(\boldsymbol{\beta}) f(\boldsymbol{\tau}) \times \prod_i f(w_i | \boldsymbol{\tau}) \prod_j l_{ij}(\{t_{ij}, \delta_{ij}\} | \boldsymbol{\beta}, w_i). \tag{6}$$

Since that posterior distribution is analytically as well as numerically hard to obtain, a Markov Chain Monte Carlo (MCMC) simulation is performed.

4 Result

Table 1: Descriptive statistics of factors taken in study.

	Survive	Died	
Variables	Number(Percentage)	Number(Percentage)	Number(Percentage)
Sex of Child			
Male	269(87.06)	40(12.94)	309(100)
Female	227(86.64)	35(13.36)	262(100)
Mother education			
literate	94(91.26)	9(8.74)	103(100)
Illeterate	402(85.90)	66(14.10)	468(100)
Family Planning attitude			
Practiced	236(89.73)	27(10.27)	263(100)
Never Practiced	260(84.42)	48(15.58)	308(100)
Total	496(86.86)	75(13.14)	571(100)
Mother age	Mean(24.69) Sd(5.08)	Mean(25.88) Sd(5.03)	Mean(24.85) Sd(5.085)

In total 571 children 75 experienced death while 496 children survived. When comparison made between survive children to died children, It was found that in male child 87.06% had survived while 12.94% had died these percentage corresponding to female child was 86.64 and 13.36. For literate mother 91.26% children survived while in illiterate mother had 85.90% survived child. The couple who ever practised family planning method experienced 10.27 % child death while the couple who never practised family planning method experienced 15.58% child death. For the mother age variable the mean and standard deviation were found (24.69±5.08) and (25.88±5.03) for survive group and died group respectively.
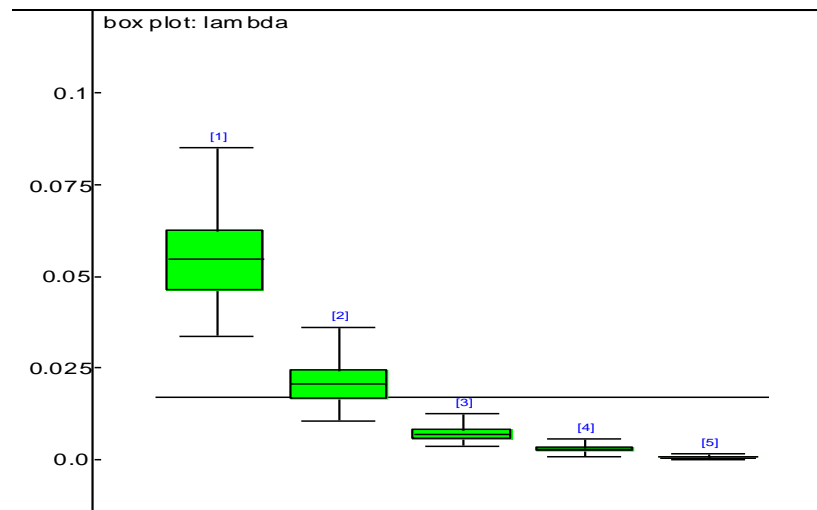
Figure-1 Box plot for the baseline hazard λ_j

Figure.1 presents the values of the baseline hazard λj by time interval. The observation period was split into the following five intervals (in months): [0,1.8), [1.8,4), [4,12), [12,24) and [24,) taken in box plot as (1),(2),(3),(4) and(5) in figure-1. Similarly to previous studies (Guo and Rodriguez 1992, Sastry 1997, Bolstad and Manda 2001,Kossi and Hognas 2005) the mortality risk is higher in the first two months of life and then, it continuously decreases.

We estimated two models: Model-1 used the covariates previously defined, without any family frailty term, and second Model included the same covariates as in Model -1 and allows for clustering by family.

Table 2 Effect pattern (regression coefficient) of covaraiates.

Variables	Model with frailty		Model without frailty	
Sex of Child	Mean(HR)	Sd	Mean(HR)	Sd
Male	0.054(1.05)	0.2565	0.078(1.08)	0.2339
Female	$		$	
Mother education				
literate	-0.475(0.62)	0.4593	-0.400(0.67)	0.3726
illiterate	$		$	
Family Planning attitude				
Practiced	-0.493(0.61)	0.3241	-0.368(0.69)	0.2482
Never Practiced	$		$	
Mother age	0.3156(1.37)	0.1569	0.216(1.24)	0.1074
Variance of frailty	1.71	0.74		
DIC	802		853	

Table 2 depicts the fixed effects coefficient obtained from both models for each covariate. The mean value and standard deviation of each fixed effect coefficient were derived from it's posterior distribution as shown in the table, for calculating these one lakh iteration were performed and ten thousand from them taken as burn in by Markov chain monte carlo (MCMC) method. The effect of gender and mother's education on child survival was not found significant by both models although it shows risk for male child death is higher in comparison to female and literacy shows a protective role from child mortality. Those who ever practiced family planning have better child survival in comparison to those who never practiced family planning. Mother's age directly effect the child mortality, as mother's age increases risk of child death also increases. In this analysis the exponetiated value of regression coefficient show the hazard rate(HR) which show the how many times risk higher for child death in particular category of a independent variable as compare to reference category of that independent variable. On that basis of HR it was found the risk of child mortality low for those children which parents ever practice family planning method by both model the value of HR are 0.61 and 0.69 by frailty and without frailty model respectively. Similarly for mother age HR are 1.37 and 1.24, which show increase in mother age risk of child mortality increases.

The deviance information criterion (DIC), as defined by Spiegelhalter et al. (2002), is used for models comparison: DIC = D + pD , where D is the posterior mean of the model deviance and pD is the effective number of parameters. The model with the smallest DIC offers the best fit. On the basis of DIC it was found frailty model gives better fit as compare to only covariate model.

The model with frailty shows smaller DIC as compared to model without frailty. A posterior mean of 1.71 is obtained for the variance of the frailty, after controlling for the variables defined in Table 1. The variance of frailty is zero means there is no relation of death of two children by same mother or family. This result means that, the death of one child in a family (without affected by family size or no. of children in family) increases the risk of death of the index child by about 1.71 times (Guo, 1993). Guo and Rodriguez (1992) found a variance of 0.22 for family random effect in Guatemala, Sastry (1997) obtained 0.516 for Northeast Brazil and Bolstad and Manda (2001), reported a variance of 0.843 for Malawi. kossi and hognas(2005) found it 0.34 in ivory cost. (M.Kalaivane et al, 2012) found it 2.16 for the India.

5. Conclusion

The primary goal of the study was to assess the determinants of under-five mortality by applying an appropriate model with use of frailty at family level and thus provide valid estimates of parameters .The well establish relationship between mother age and child death was found here which shows that the pregnancy at higher age is risky for child survival. Family planning practices were found to be protective for child death. The variance of frailty found here is 1.71 and statistically significant which shows higher familial effect on child death. This shows genetic and environmental component also play role in child mortality, in that case we also think for prenatal genetic screening for mothers of higher age and those who already have an afflicted child. The win BUGS software used here provides a simple way to perform suchtype of complex analysis.

6. Recommendation

Due to cultural and religious practices of Muslim couples they often show reluctance in practicing family planning methods and continue delivering babies even at older age. Marriages' between cross cousins are also practiced in this community, which is also a risk factor for infant mortality. So, on the basis of this study we suggest the government to spread public awareness regarding such problems.

Acknowledgment

References

[1] Bhattacharyya, R. and Rao, A. S. R. S. (2009) Estimation and analysis of child mortality for Indian states through a Bayesian approach. *Webpage, http:www.isical.ac.in/~arni.*

[2] Bolstd, W.M. and Manda, S.O. (2001). Investigating Child Mortality in Malawi. Using Family and Community Random Effects: A Bayesian Analysis. Journal of the American Statistical Association **96,** 453.

[3] Curtis, S. Diamond, I. and McDonald, J. (1993). Birth Interval and Family Effects on Post Neonatal Mortality in Brazil. *Demography* **301**, 33-43.

[4] Das Gupta, M. (1990). Death clustering mother's education and the determinants of child mortality in rural Punjab, India. Population studies **44**(**3**), 489-505.

[5] DLHS-1 District Level Health Survey (1998-1999).IIPS, Mumbai.

[6] DLHS-2 District Level Health Survey (2002-2004).IIPS, Mumbai.

[7] DLHS-3 District level health Survey (2007-2008). Conducted by International institute for population Sciences on Sample basis in all over India during the period December 2007 to December 2008.

[8] Friedman, M. (1982). Piecewise exponential models for survival data with covariates. *Ann Stat* **10**,101-113.

[9] Guo, G. (1993). Use of Sibling Data to Estimate Family Mortality Effects in Guatemala. *Demography* **30** (**1**), 15-32.

[10] Guo, G. and Rodriguez G. (1992). Estimating a Multivariate Proportional Hazards Model for Clustered Data Using the EM Algorithm, with an Application to Child Survival in Guatemala. *Journal of American Statistical Association* **87**, 969-976.

[11] Hobcraft, J. McDonald, J. and Rutstein, O. (1985). Demographic Determinants of Infant and Early Child Mortality: A Comparative Analysis. Population Studies **39** (**3**), 363-385.

[12] Indian Census, (2011). Office of Registrar general & census commissioner, India.

[13] Jain, A.K. and Visaria, P. (1988). Infant mortality in India: differential and determinants. Sage Publication, New Delhi.

[14] Koissi, M. C. and Hogans, G. (2005). Using winbugs to study family fraility in child mortality. African population studies **20**, 1-27.

[15] Mani, K. et al. (2012). Determinants of Under-Five mortality in rural empowered action group: An application of cox-frailty model. IJMA **1**(**1**), 60-72.
[16] Park, K. (2005). A book of Preventive and Social medicine. Bhanot publication, Jabalpur.

[17] Pew (2009).The Pew forum of religious and public life.

[18] Sastry, N. (1997). Family-Level Clustering of Childhood Mortality Risk in Northeast Brazil". Population Studies **51**, 245-261.

[19] Sen, A .(1998). Mortality as an Indicator of Economic Success and Failure. *Economic Journal* **108**, 1-25.

[20] Singh, K. K. , Pandey, N. and Gautam, A. (2008). A Probability Model for Describing the Distribution of Child Death and its Application. Population Poverty & Health: Analytical Approaches (ed. K.K. Singh, R.C. Yadava and Arvind Pandey), *Hindustan Publishing Corporation (India)*, *New Delhi*, 146-155.

[21] Sinha, D. and Dey, D. K. (1997). Semiparametric Bayesian Analysis of Survival Data. *Journal of the American Statistical Association* **92**, 1195-1212.

[22] Spiegelhalter, D.,Thomas, A., Best, N. and Lunn, D. (2003). WinBUGS version 1.4.*(http://www.mrc-bsu.cam.ac.uk/bugs)*

[23] Uddin, J. et al. (2008).Child mortality in a developing country: a statistical analysis. *JAQM* **4**, 271-283.

[24] Vaupel, J. Manton, K. and Stallard, E. (1979). The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality". *Demography* 16, 439-454.

G. P. Singh
Department of Community Medicine and DST-CIMS
Banaras Hindu University
Varanasi-221005, Uttar Pradesh, India
singhgpbhu@gmail.com

A. Tripathi
Department of Community Medicine and DST-CIMS
Banaras Hindu University
Varanasi-221005, Uttar Pradesh, India
ashoka07bhu@yahoo.in

S. K. Singh
Department of Statistics and DST-CIMS
Banaras Hindu University
Varanasi-221005, Uttar Pradesh, India
singhsk64@gmail.com

Umesh Singh
Department of Statistics and DST-CIMS
Banaras Hindu University
Varanasi-221005, Uttar Pradesh, India
umeshsingh52@gmail.com

O.P. Singh
Department of Statistics
Uday Pratap College
Varanasi-221005, Uttar Pradesh, India
abhinavt39@gmail.com