J. Asiat. Soc. Bangladesh, Sci. 42(2): 243-249, December 2016

AN ANALYSIS OF BIRTH INTERVALS IN BANGLADESH USING FRAILTY MODELS

HASINA ISLAM[1]
Department of Mathematics and Statistics, Bangladesh University of Business and Technology, Dhaka-1216, Bangladesh

Abstract

This research dealt with the extension of the Cox's proportional hazards model that allows for heterogeneity among the responses due to random effects of covariates using frailty (random effect) approach. The results of this study showed that the Gaussian frailty model is better for the data than gamma frailty model and the unobserved cluster effect has a sizeable impact on the second birth spacing in Bangladesh. This research also showed the current pattern of the second birth spacing in Bangladesh and different demographic and socio-economic factors which affect the second birth spacing. It was found that mother’s education, survival status of 1st birth, region, place of residence and mother's age of marriage have great influence for the variation of the second birth spacing.

Key words: Frailty model, Birth space, Cox proportional hazards model

Introduction

Birth spacing (interval) refers to the time interval from one child's birth until the next child's birth i.e; length of the time between two successive births. Identification of the factors causing variation in the length of birth interval could have of great importance for it's direct relation to fertility. The covariates may play some key role to the second birth spacing. Determination and identification of the factors causing variation in the length of birth interval are of great importance for its direct relation to fertility.

Conventional Cox's model assumes that the investigated subjects under given experimental conditions are independent and identically distributed and hence homogeneous by nature. There may be situations where there may exist some factors other than the measured covariates which can significantly influence the parameters and hence modify the distribution of the survival time. There may be various reasons for such unmeasured or neglected covariates. If there are too many covariates to consider, it is nearly impossible for the researchers in practice to include all the relevant covariates. Then they are tempted to overlook some of the relevant covariates. Another common reason may be that researchers are not aware of the influence of the potential covariate that might exist. For example, if there is a genetic risk factor responsible for potential occurrence of some diseases, which may be unknown to us, it is not possible for the researcher to include that

[1] Corresponding address : E-mail: hasinaislam13@yahoo.com

as a covariate. Such covariates are said to be the unobserved covariates. Therefore, for practical reasons such unobservable covariates are ignored by considering them as a part of the error component and not controlled in conventional survival analysis. This may greatly simplify the calculation, but this advantage comes at a great price.

Therefore, as individuals in any group are dissimilar in their own rights, the model has to be improved to account such hidden heterogeneity and modification of homogeneity assumption is necessary. Keyfitz (1978) and Vaupel *et al.* (1979) suggested for the mixture of individuals with different hazards in the same population. Experiences in their studies suggested that there exists a considerable variation in the risk of developing various diseases and thus, individuals in a biological population differ substantially in susceptibility for various mortality and morbidity events. Individuals have different frailties, and that those who are most frail will die earlier than the others. The reason can be different patterns of gene they carry, the distinct life style they follow etc (Zahan 2010). To deal with such problem in the survival analysis, frailty models have been suggested by various researchers. This model corrects the bias of the regression coefficients in the Cox's Proportional Hazard model (Chamberlain 1979). The frailty models also plays role in describing the non-proportionality of the conditional hazards, which in turn improves the fit. The frailty model was historically first introduced by Clayton (1978) as a bivariate model.

Materials and Methods

Secondary data extracted from the Bangladesh Demographic and Health Survey (NIPORT 2011) conducted under the authority of the National Institute for Population Research and Training (NIPORT) of the Ministry of Health and Family Welfare, Government of the People's Republic of Bangladesh have been used for this study. The sample for the 2011 BDHS is nationally representative and covers the entire population residing in no institutional dwelling units in the country (NIPORT 2011). The second birth interval was considered because if second birth interval is higher then the chance of getting more children is lower. At first the gamma frailty model and Gaussian frailty model were fitted with the help of Bangladesh Demographic and Health Survey data to estimate the corresponding standard errors of the regression coefficients. Then they were compared by using Akaike Information Criterion (AIC) value.

Birth space was taken as dependent variable and defined as continuous variable. Birth spacing for 1st two births was considered. The 2nd birth after 2005 was considered for showing the current pattern of 2nd birth space, 1st birth was considered at any time.

The important factors of demographic and socio-economic had been identified as explanatory variables on the basis of the previous studies. They were mother's age of marriage, mother's education, mother's working status, wealth index, region, place of residence, religion, gender of 1st child and survival status of 1st child.

Estimation Procedures of Frailty Model: There are several procedures for the estimation of frailty model. In this study the penalized approach was used. The penalized approach is based on a modification of the Cox's partial likelihood so that both the regression coefficients and frailties are included and optimized over. Specifically the likelihood is described as a product where the first term is the partial likelihood including the frailty terms as parameters. The second term is a penalty introduced to avoid large differences between the frailties for the different groups. In practice it is fitted by first setting the frailty values to 1. Then an iterative procedure was used with a first step of optimizing the partial likelihood, treating the frailties as fixed and known parameters. In the second step the frailties were evaluated as the conditional means giving their observations, using the formulas, like the EM-algorithm. The experiment was repeated until convergence.

Assuming that the data for subject *i* , who is member of the j^{th} family, follows proportional hazards shared frailty model is given by,

$$h_{if} = h_o(t) \exp (x^{p}{}_{i}\beta + z_i\, w_j)$$

where x is the vector of covariate for subject *i* and β is a vector of regression coefficients, W_j is the frailty for family, *j* with independently and identically distributed from some positive scale family with density function $f(w;\ \theta)$, having mean 1 and variance θ. Z is matrix of indicator variables such that $z_{ij} = 1$ when subject *i* is a member of family j and 0 otherwise, and each individual belongs to only one family.
Estimation under this model is done by maximizing the penalized partial log-likelihood,
$PPL = PL(\beta, W\text{i}\ data) - g(w;\ \theta)$,
over both β and w. Here PL is the log of the usual Cox's partial likelihood function,

$$PL(\beta, w) = \sum_{i=1}^{n} \int_0^{\infty} \left[Y_i(t)(x_i\beta + z_i\ w) - \log\left(\sum_k Y_k(t) \exp(x_k\beta + z_k\ w) \right) \right] dN_i\ (t);$$

Where $Y_i(t)$ is an observed process taking the value 1 or 0 according to whether or not subject i is observed at time *t* and *g* is a penalty function chosen by the investigator to restrict the values of *w*. Typically, one would choose the penalty function to shrink w toward zero and use to control the amount of shrinkage. A penalized Cox model with penalty function $g(w;\theta) = \frac{1}{\theta}\sum[w_i - \exp(w_i)]$ is equivalent to the gamma frailty model discussed in Klein (1992) and Nielson *et al.* (1992). The w_is are distributed as the logs of independently and identically distributed gamma random variables and the tuning parameter θ is their variance. A penalized Cox model with penalty function $g(w;\theta) = \frac{1}{2\theta}\sum_i w_i^2$ is equivalent to the Gaussian random effects model.

Results and Discussion

Different choices of distributions for the unobserved covariates are possible. In this study gamma and Gaussian frailty model were chosen. The variance of frailty distribution determines the degree of heterogeneity in the study population. It is apparent from Tables 1 and 2 that for Gaussian frailty model unobserved cluster variance was more than for gamma frailty model (for gamma 0.004 and for Gaussian 0.041). Thus Gaussian frailty model was observed to be more heterogeneous than gamma frailty model.

Table 1. Parameter estimates, standard errors (SE), p-values and hazards ratio (HR) obtained from Gaussian frailty model.

Covariates	Estimate	SE	p-value	HR
Child's sex				
Female	---			
Male	-0.022	0.038	0.552	0.978
Region				
Dhaka	---			
Barisal	-0.016	0.075	0.827	0.984
Chittagong	0.266	0.065	0.000	1.302
Khulna	-0.040	0.071	0.571	0.961
Rajshahi	-0.012	0.070	0.866	0.988
Rangpur	0.034	0.072	0.640	1.034
Sylhet	0.405	0.072	0.000	1.497
Place				
Rural	---			
Urban	-0.140	0.044	0.002	0.870
Mother's Education				
No education	---			
Primary education	-0.101	0.063	0.108	0.904
Secondary education	-0.380	0.063	0.000	0.684
Higher education	-0.378	0.095	0.000	0.685
Wealth Index				
Poor	---			
Middle	0.003	0.054	0.960	1.002
Rich	0.040	0.055	0.485	1.041
Working women				
No	---			
Yes	0.112	0.056	0.057	1.121
Survival status of 1st child				
Death	---			
Alive	-1.026	0.071	0.000	0.358
Religion				
Other	---			
Islam	0.040	0.062	0.512	1.041
Mother's age of marriage				
Linear Effect	-0.113	0.044	0.010	0.893
Squared Effect	0.002	0.001	0.156	0.002
Cluster Variance	0.041		0.000	

$SE_{\text{Standard Error}}$, $HR_{\text{Hazard Ratio.}}$

Table 1 shows the hazard ratio of different variables. When the 2nd birth interval was compared with different divisions', it was found that the birth interval was more in Chittagong and Sylhet than in Dhaka. It was observed that in urban areas 2nd birth interval

Table 2. Parameter estimates, standard errors (SE), p-values and hazards ratio (HR) obtained from gamma frailty model.

Covariates	Estimate	SE	p-value	HR
Child's sex				
Female	---			
Male	-0.022	0.038	0.552	0.978
Region				
Dhaka	---			
Barisal	-0.016	0.075	0.827	0.984
Chittagong	0.266	0.065	0.000	1.302
Khulna	-0.040	0.071	0.571	0.961
Rajshahi	-0.012	0.070	0.866	0.988
Rangpur	0.034	0.072	0.640	1.034
Sylhet	0.405	0.072	0.000	1.497
Place				
Rural	---			
Urban	-0.140	0.044	0.002	0.870
Mother's Education				
No education	---			
Primary education	-0.101	0.063	0.108	0.904
Secondary education	-0.380	0.063	0.000	0.684
Higher education	-0.378	0.095	0.000	0.685
Wealth Index				
Poor	---			
Middle	0.003	0.054	0.960	1.002
Rich	0.040	0.055	0.485	1.041
Working women				
No	---			
Yes	0.112	0.056	0.057	1.121
Survival status of 1st child				
Death	---			
Alive	-1.026	0.071	0.000	0.358
Religion				
Other	---			
Islam	0.040	0.062	0.512	1.041
Mother's age of marriage				
Linear Effect	-0.113	0.044	0.010	0.893
Squared Effect	0.002	0.001	0.156	0.002
Cluster Variance	0.004		0.000	

SE Standard Error, HR Hazard Ratio.

was 0.870 times lower than in rural areas. It was also observed that if previous child is alive then birth space has 0.358 times lower compared to survival status of 1st child is dead. From the Table 1 it is apparent that 2nd birth space was 0.684 and 0.685 times lower in secondary and higher educated mother respectively compared to non educated mother. Mother's age of marriage was found statistically significant at 1% level of significance (Table 1). It was observed from this findings the linear effect and square effect for mother's age of marriage were $-0.113 < 0$ and $0.002 > 0$ respectively. These results indicate that with the increase of mother's age of marriage birth spacing decreases up to a particular level of age of marriage, after that birth spacing increases with the increase of mother's age of marriage. Gender of 1st child, religion and wealth index were found statistically insignificant. The results are in agreement with the work of Rabbi *et al.* 2012.

For the data for a given set of candidate model, the preferred model is that have minimum Akaike Information Criterion (AIC) value. Results presented in Table 3 show that the AIC value for Gaussian frailty model was minimum indicating that among the frailty models Gaussian fit well for the data. The AIC value for this study for gamma frailty model was higher than Gaussian frailty model but for gamma model the study gave the significant result also, so gamma frailty model was moderately well for the data.

Table 3. Comparison of frailty models.

Frailty Model	Log likelihood value	AIC	Cluster Variance
Gamma	-23370.01	46776.2	0.004
Gaussian	-23278.85	46593.7	0.041

AIC Akaike Information Criterion.

This study evaluated the current effect of some selected demographic and socio-economic variables on subsequent birth interval using 2011 BDHS data. Among the nine explanatory variables examined mother's education, survival status of 1st child, region and place of residence were found to have strong impact on 2nd birth interval.

Normally educated women always have longer birth interval than non-educated women but this result shows that secondary and higher educated women 2nd birth interval is shorter than non-educated women. Urban mothers' have smaller birth interval than that of their rural counterparts which indicates that the lack of development in fertility behavior among the rural families, who are not aware of high parity progression. Chittagong and Sylhet divisions have larger 2nd birth interval compared to Dhaka which indicates that they are aware about fertility behavior. Other divisions differ insignificantly with Dhaka. The result also shows that if 1st birth was dead then 2nd birth interval increases. It might be due to cause of mothers' physical complication of pregnancy for which mother's are bound to wait for recovery and having the next child. Thus 2nd birth interval increased.

The rate of fertility is started to decline in Bangladesh but in near future a rapid reduction of current fertility trends is needed for achieving replacement level of fertility. In this study it was found that mother’s education, survival status of 1st birth, region, place of residence and mother's age of marriage have much influence for the variation of the birth spacing. For getting replacement level of fertility these factors should be considered.

Acknowledgement

The author expresses her gratitude and thanks to Prof. Dr. Wasimul Bari, Department of Statistics, Bio-statistics and Informatics, University of Dhaka for his valuable directions and criticisms during the preparation of this paper.

References

Chamberlain, G. 1979. Heterogeneity; Omitted Variable Bias and Duration Dependence. *Harvard Insti tute of Economic Research.* **25:** 102-109

Clayton, D.G. 1978. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrics.* **55**: 1162-1170.

Keyfitz, N. 1978. Improving Life Expectancy: An Uphill Road Ahead. *American Journal of Public Health*. **10:** 954-956.

Klein, J.P. 1992. Semi parametric estimation of random effects using the Cox model EM algorithm. *Biometrics*. **48**: 795-806.

Nielsen, G.G., R.D. Gill, P.K. Andersen and T.I.A. Srensen. 1992. A counting process approach to maxi mum likelihood estimation in frailty models. *Scandinavian Journal of Statistics.***19**: 25-44.

NIPORT, Mitra and Association & ORC Macro. 2009. Bangladesh Demographic and Health Survey . 2011.

Rabbi, A.M.F., S. C. Karmaker, S.A. Mallick, S. Sharmin. 2013. Determinants of Birth Spacing on Fertility in Bangladesh. *Dhaka University Journal of Science*. **61(1):** 105-110

Vaupel, J.W., Manton, K.G. and E. Stallard. 1979. *Demography*. **16**: 439-454.

Zahan, R. 2010. Analysis of child mortality data using Bayesian and frequentist approaches. Unpublished MS Thesis. 95.

(Revised copy received on 12/12/2016)