Global Scientific JOURNALS

GSJ: Volume 8, Issue 4, April 2020, Online: ISSN 2320-9186
www.globalscientificjournal.com

CORRELATEDFACTORS OF EARLY MARRIAGE IN BANGLADESH: LOGISTIC REGRESSION ANALYSIS

Author: Dr. Md. Roshidul Islam

Associate Professor, Department of Statistics, Begum Rokeya University, Rangpur, Bangladesh.
E-mail: mroshid.brur@gmail.com

Abstract: This paper explores the issue of early marriage in Bangladesh. Early marriage is nothing new in Bangladesh; it is deeply embebed in the improvised and traditional cultural settings. The main objective of this study was to identify on factors contributing to early marriage in Bangladesh based on national surveys of Bangladesh Demographic and Health survey (BDHS) 2014 that are associated with early marriage. The study used secondary data from Bangladesh Demographic and Health Survey (BDHS) 2014. Logistic regression analysis was applied to find out the most significant factors of early marriage in Bangladesh. It was found thatrespondents education level, division, religion, husband/partner's education level, respondent work status level most significant factors on early marriage in Bangladesh.

Keywords: Early marriage, Odds ratio, Affecting factors, Logistic regression analysis, Bangladesh.

INTRODUCTION

Marriage is an important universal social institution for the person and society at large. For the person, it is a significant and memorable event in one's life cycle as well as the most important base in the family formation process. In addition, marriage marks the begging to an end of the changeover to adulthood as the person separates from the parental home, even if generations continue to the socially and economically interdependent through the extended family.

The definition of marriage varies according to different cultures but in general, marriage is a socially or ritually accepted union or legal contract between an adult man and an adult woman that establishes rights and responsibility between them. In many different countries and parts of the world, young women's marriage before the age of 18 is a reality to be lived.

According to UNICEF, early marriage is a marriage contracted before the age of 18, and this is common among young girls.According to Child Marriage Restraint Act (1929), the legal marriage age for a girl is 18 years and for a boy is 21 years in Bangladesh. The punishments for early marriage of women according to this act are imprisonment for one month or a fine up to taka 1000 or both.

All over the world, marriage is considered as a moment of celebration and a milestone in adult life. Unfortunately, the practice of early marriage gives no such reason for celebration. All too often, the infliction of a marriage partner upon a child means that a girl childhood is cut short and their essential rights are compromised.

This paper attempts to clear up the issue by analyzing the affecting factors of early marriage for women in Bangladesh, an area with one of the highest rates of early marriage worldwide.

Objectives of the study

The overall objectives of this study are as bellow:

i) To investigate the age at first marriage in teenage women in Bangladesh.

ii) To find out the affecting factors that affects the early marriage in Bangladesh.

Coverage of the sample

The data are extracted from the 2014 Bangladesh and Health Survey (BDHS), according to the study 1172 respondents have been considered out of 43,772 respondents.

METHODOLOGY

Logistic regression analysis

Logistic regression is used to model the binary relationship between a binary response variable and one or more predictor variables, which may be either discrete or continuous. A variety of multivariate techniques can be used to predict a binary dependent variable from a set independent variables. We can make age at marriage a binary variable by considering women married before a certain age as early married and married after that age as late married. Two important multivariate techniques are multiple regressions and discriminate analysis. Sometimes it is difficult to apply these techniques when the dependent variables are categorized (dichotomous and polygamous). In this case the assumption necessary for hypothesis testing in regression analysis is violated. For instance, the distribution of errors cannot be assumed as normal. In addition, predicted values obtained from multiple regression analysis cannot be used as probability as they are not contained to fall in the interval between 0 and 1. Linear discriminate analysis is, on the other hand, based on the assumption that the independent variables are normally distributed with equal variances. In such situation the multivariate technique called Cox's (1970) linear logistic regression model is most appropriate, since it does not require any distributional assumption. Cox (1958) is the pioneer of logistic regression model. It can be used to identify risk factors as well as predict the probability of 'success' (P_i) to 'failure' ($1-P_i$) and relating it to the independent variables, the logistic parameter can easily be interpreted in terms of odds and odds ratio, relative odds can be estimated for the categories of each independent categorical variables or combination of such variables.

Linear logistic regression model

Let us suppose that there are 'n' individual to whom an event occurred. These are termed as the successes and the 'failures'.Let,

$Y_i=1$, if i^{th} individual is success And

$Y_i=0$, if i^{th} individual is a failure, for $i=1, 2, \dots, n$

Let us suppose again that for each of the 'n' individuals, 'p' independent variables x_{i1}, x_{i2}, x_{i3}, …,x_{ip}were measured. These variables can both be qualitative or quantitative. Our aim is to relate the set of independent variables to the dichotomous dependent variable.

In the linear logistic regression model the dependency of probability of success on independent variables is assumed to be

$$\pi(x_i) = p(y_i = 1\backslash x) = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip})}$$

and the probability of failure is

$$1 - \pi(x_i) = p(y_i = 0\backslash x)$$

$$= \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip})}$$

A transformation of $\pi(x_i)$ that will be the central part to the study of logistic regression model is the logit transformation. The transformation is defined in terms of $\pi(x_i)$, is as follows:

$$\lambda_i = g(x_i) = \text{logit}\ \pi(x_i) = \ln\left[\frac{\pi(x_i)}{1 - \pi(x_i)}\right]$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots\ldots\ldots + \beta_p X_{ip}$$

The importance of logit transformation is that $g(x_i)$ has many of the desirable properties of a linear regression model. The logit $g(x_i)$ is linear in its parameter, may be continuous and may range from $-\infty$ to $+\infty$, depending on the range of X.

Estimation Technique

The most common method used to estimate unknown parameters in linear regression is least squares. Under usual assumptions, least squares estimators have some desirable properties. But when least squares method is applied to estimate a model with dichotomous outcome the estimators no longer have these same properties. In such situations, the general method for estimating the parameters of logistic regression models is the method of maximum likelihood.

In regression, the likelihood equation is non-linear and explicit function of unknown parameters. Therefore, we use a very effective and well-known iterative method, Newton-Raphson method. Cox and Hinkley (1974), Kalbfleisch and Prentice (1980) and Lawless (1981) have given a detail of this method.

Now, let us consider a single regression model as

$Y_i = \beta_o + \beta_1 X_i + \varepsilon_i$; (i=1, 2,…………, n) ………………………………… (i)

Where, Y_i is a dichotomous variable, it takes two values 0 and 1.

So, Y_i is a Bernoulli random variable. So the p.d.f of Y_i is given by

$f_i(Y_i) = p_i^{Y_i}(1-p_i)^{1-Y_i}$; ($Y_i$ = 0 or 1) ………………………………… (ii)

i= 1, 2,………, n

Where, p_i is a probability that define as

$$p_i = \frac{e^{\beta_o + \beta_1 X_i}}{1+e^{\beta_o + \beta_1 X_i}}$$

$$\therefore \frac{p_i}{1-p_i} = e^{\beta_o + \beta_1 X_i} \text{ ………………………………… (iii)}$$

Implies that, $log_e\left(\frac{p_i}{1-p_i}\right) = log_e\left(e^{\beta_o + \beta_1 X_i}\right)$

Since those are assumed to be independent, the joint probability density function is

$$g(Y_1, Y_2, \ldots\ldots Y_n) = \prod_{i=1}^{n} f_i(Y_i)$$

$$= \prod_{i=1}^{n} p_i^{Y_i}(1-p_i)^{1-Y_i} \text{ ……………………………(iv)}$$

Since the algorithm is a monotonic function, so, taking logarithm on (iv),we get

$$\log_e[g(Y_1, Y_2, \ldots\ldots Y_n)] = \left(\log_e \prod_{i=1}^{n} p_i^{Y_i}(1-p_i)^{1-Y_i}\right)$$

$$= \sum_{i=1}^{n}\left[Y_i \log_e p_i + (1-Y_i)\log_e(1-p_i)\right]$$

$$\log_e[\,g(Y_1, Y_2, \ldots\ldots Y_n)] = \sum_{i=1}^{n}\left[Y_i\{\log_e p_i - \log_e(1-p_i)\} + \log_e(1-p_i)\right]$$

$$=\sum_{i=1}^{n}\left[Y_i \log\left(\frac{p_i}{1-p_i}\right) + \log(1-p_i)\right]$$

$$=\sum_{i=1}^{n}\left[Y_i(\beta_o + \beta_1 X_i) + \log\left(\frac{1}{1+e^{(\beta_o+\beta_1 X_i)}}\right)\right]$$

$$I_i(\text{say}) = \sum_{i=1}^{n}\left[Y_i(\beta_o + \beta_1 X_i) - \log\left\{1 + e^{(\beta_o+\beta_1 X_i)}\right\}\right] =$$

$$=\sum_{i=1}^{n} Y_i\,(\beta_o + \beta_1 X_i) - \sum_{i=1}^{n} \log\left\{1 + e^{(\beta_o+\beta_1 X_i)}\right\} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{(v)}$$

Now, differentiating eq.(v) with respect to β_o and β_1 respectively,

$$\frac{\delta I_i}{\delta \beta_o} = \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n}\left[\frac{e^{\beta_o+\beta_1 X_i}}{1+e^{\beta_o+\beta_1 X_i}}\right]$$

$$\text{Or,} \frac{\delta I_i}{\delta \beta_o} = \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} p_i$$

And $$\frac{\delta I_i}{\delta \beta_1} = \sum_{i=1}^{n} X_i\, Y_i - \sum_{i=1}^{n}\left[\frac{X_i e^{\beta_o+\beta_1 X_i}}{1+e^{\beta_o+\beta_1 X_i}}\right]$$

$$\text{Or,} \frac{\delta I_i}{\delta \beta_1} = \sum_{i=1}^{n} X_i\, Y_i - \sum_{i=1}^{n} p_i$$

And

$$\frac{\delta I_i}{\delta \beta} = \begin{pmatrix} \frac{\delta I_i}{\delta \beta_o} \\ \frac{\delta I_i}{\delta \beta_1} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} p_i \\ \sum_{i=1}^{n} X_i\, Y_i - \sum_{i=1}^{n} p_i \end{pmatrix}$$

$$= X^T\text{Y} - X^T\text{p}$$

$$= X^T(\text{Y-p}) \cdots\cdots\cdots\cdots\cdots\cdots \text{(vi)}$$

Where,

$$X = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}$$

Now, we put $\frac{\delta I_i}{\delta \beta} = 0$, then we get

X^T(Y-p) = 0 ……………………………………… (vii)

Or, $X^T Y = X^T$p

Or, $\hat{Y} = \hat{p}$

The solution to equation (vii) will satisfy

X^T(Y-$\hat{p}$) = 0 ………………………………… (viii)

Equation (viii) is generally solved by using Newton-Raphson method. This entails first determining,

$$\frac{\delta}{\delta\beta}(X^T Y - X^T \mathrm{p}) = \frac{\delta}{\delta\beta}\{X^T(\mathrm{Y} - p)\} = -\frac{\delta}{\delta\beta}X^T p = -\left[\frac{\delta}{\delta\beta}p\right]X$$

But we have,

$$p_i = \frac{e^{\beta_o + \beta_1 X_i}}{1 + e^{\beta_o + \beta_1 X_i}}$$

$$\frac{\delta p_i}{\delta \beta_o} = \frac{\left\{1 + e^{(\beta_o + \beta_1 X_i)}\right\} e^{(\beta_o + \beta_1 X_i)} - e^{(\beta_o + \beta_1 X_i)} e^{(\beta_o + \beta_1 X_i)}}{\left\{1 + e^{(\beta_o + \beta_1 X_i)}\right\}^2}$$

$$= \frac{e^{(\beta_o + \beta_1 X_i)}}{\left\{1 + e^{(\beta_o + \beta_1 X_i)}\right\}} \times \frac{1}{e^{\beta_o + \beta_1 X_i}}$$

$$= p_i(1 - p_i)$$

$$\therefore \frac{\delta p_i}{\delta \beta_1} = \frac{\left\{X_i e^{\beta_o + \beta_1 X_i}\right\}}{\left\{1 + e^{(\beta_o + \beta_1 X_i)}\right\}^2}$$

$=X_i p_i(1-p_i)$

$\therefore \frac{\delta p}{\delta \beta} = X^T \mathrm{W}$

Where, W= diag And$X^T = \begin{pmatrix} 1 & 1 & \vdots & 1 \\ X_1 & X_2 & \vdots & X_n \end{pmatrix}$

$$\frac{\delta p}{\delta \beta} = X^T \mathrm{WX}$$

Iteratively estimates of β are then obtained as

$\hat{\beta} = (X^T\mathrm{WX})^{-1}(X^T\mathrm{WX}) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$ (ix)

With Z playing the role of Y in this iteratively reweighted least squares approach.

Specifically,

$$Z_i = \hat{\eta} + \frac{Y_i + \widehat{p_i}}{\widehat{p_i}(1-\widehat{p_i})}$$

With $\eta_i = \log_e\left(\frac{p_i}{1-p_i}\right)$. Notice that $\widehat{\eta_i}$ plays an important role of Y_i and $\frac{Y_i+\widehat{p_i}}{\widehat{p_i}(1-\widehat{p_i})}$ is the residual corresponding to Y_i divided by the estimated variance of Y_i.

If we wish to write equation (ix)

In an equivalent form that shows the updating of $\hat{\beta}$, we may write Z as:

$$\mathrm{Z} = \hat{\beta}^{(i)} + W^{-1}\mathfrak{e}$$

With e = Y$-\widehat{p_i}$, we then obtain

$\hat{\beta}^{(i+1)} = (X^T\mathrm{WX})^{-1}X^T\mathrm{W}(X\hat{\beta}^{(i)} + W^{-1})$

$= \hat{\beta}^{(i)} + (X^T\mathrm{WX})^{-1}X^T\mathrm{e} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$ (x)

The updating formula given by equation (x) is used until the estimates converge.

The first step is to obtain the initial estimates;$\beta^{(0)}$.Various approaches are used to obtain these. As originally derived by Lachebrunch (1975) and displayed by Hosmer and Lemeshow (1989), the initial estimates obtained using the discriminant function are given by

$$\hat{\beta}^{(0)} = \begin{pmatrix} \hat{\beta}_0^{(0)} \\ \hat{\beta}_1^{(0)} \end{pmatrix} = \begin{pmatrix} \log_e \left(\frac{\widehat{\boldsymbol{\Theta}}_1}{\widehat{\boldsymbol{\Theta}}_0}\right) - 0.5(\hat{\mu}_1^2 - \hat{\mu}_0^2) + \hat{\sigma}^2 \\ \frac{\hat{\mu}_1^2 - \hat{\mu}_0^2}{\hat{\sigma}^2} \end{pmatrix}$$

Here, $\hat{\mu}_0 = \bar{X}_0$ and, $\hat{\mu}_1 = \bar{X}_1$, where $\bar{X}_0$and $\bar{X}_1$ are the average of the n-values when Y=0 and Y=1 respectively. And$\widehat{\boldsymbol{\Theta}}_1 = \bar{Y}$ and $\widehat{\boldsymbol{\Theta}}_0 = 1 - \widehat{\boldsymbol{\Theta}}_1$

$$\text{And}\hat{\sigma}^2 = \frac{(n_0-1)s_0^2 + (n_1-1)s_1^2}{n_0+n_1-2}$$

Where s_0^2 *and* s_1^2 are the usual sample variances computed using Y=0 and Y=1 respectively, and n_0 and n_1 are the corresponding sample sizes.

RESULTS

The result of logistic regression analysis is shown in Table 1.

Table 1:Results of Logistic Regression Analysis for Early Marriage patterns

Variables	Coefficient B	S.E.	Wald	Sig.	Odds Ratio	95.0% C.I. for EXP(B)	
						Lower	Upper
Educational level			31.903	.000			
no Education (1)	-1.734	.507	11.686	.001	.177	.065	.477
incomplete primary (2)	-1.598	.373	18.366	.000	.202	.097	.420
complete primary (3)	-1.446	.334	18.735	.000	.236	.122	.453
incomplete secondary (4)	-1.137	.218	27.149	.000	.321	.209	.492
higher							

Religion			6.284	.043			
Islam (1)	-.250	1.275	.038	.845	.779	.064	9.484
Hinduism (2)	.409	1.295	.100	.752	1.505	.119	19.033
Buddhism							
Husband/partner's Education level			10.291	.016			
no education (1)	-.716	.359	3.976	.046	.489	.242	.988
primary (2)	-.238	.306	.605	.437	.788	.433	1.435
secondary (3)	.080	.283	.080	.777	1.083	.622	1.888
higher and don't know							
Respondent Work status No **Yes**	-.464	.206	5.075	.024	.629	.420	.941
Division			29.746	.000			
Barisal (1)	-.765	.282	7.334	.007	.465	.268	.810
Chittagong (2)	-.188	.247	.577	.447	.829	.511	1.345
Dhaka (3)	-.470	.255	3.394	.065	.625	.379	1.030
Khulna (4)	-.825	.277	8.858	.003	.438	.254	.754
Rajshahi(5)	-1.066	.278	14.686	.000	.344	.200	.594
Rangpur (6)	-1.151	.280	16.941	.000	.316	.183	.547
Sylhet							

DISCUSSIONS

The result of logistic regression analysis in Table 1 we can say that such variables are Educational level and Division are significant at 1% level of significance, the variables Religion, Husband/partner's Education level, Respondent Work status are significant 5% level of significance. The six column of the Table 1 shows the odds ratio.

The odds ratio of Educational level, we see that odds ratio for no Education is .177, incomplete primary is .202, complete primary is .236, incomplete secondary is .321.These indicate that odds ratio are .177,.202,.236,.321 times higher for those respondents who have no Education, incomplete primary, complete primary, incomplete secondary than those higher.

The odds ratio of Religion, we see that odds ratio for Islam is .779, Hinduism is1.505. These indicate that the odds ratio .779 and1.505 times higher for those respondents who areIslam, Hinduism than those Buddhism.

The odds ratio ofHusband/partner’s Education level, we see that odds ratio for no education is.489, primary is.788, and secondary is1.083. These indicate that the odds ratio .489,.788, 1.083 times higher for those respondents who are no education, primary, secondary than those higher and don’t know.

The odds ratio of Respondent Work status, we see that odds ratio for not working is .629.These indicate that the odds ratio .629 times higher for those respondents who are not working than those working.

The odds ratio of Division, we see that odds ratio for Barisal 0.465, Chittagong 0.829, Dhaka .625, Khulna .438, Rajshahi .344, Rangpur .316. These indicate that the odds ratio .465, .829, .625, .438, .344, .319 times higher for those respondents who are Barisal, Chittagong, Dhaka, Khulna, Rajshahi, Rangpur than those Sylhet.

Therefore most important significant variables that influence early marriage are Educational level, Division, Religion, Husband/partner’s Education level, Respondent Work status.

CONCLUSION

Findings of the study revealed that early marriage strongly influenced by education, socio-economic status, determined by assets owned. Marriage is one of the demographic events of human life, especially for women in Bangladesh that is determined by socio-demographic factors and also by culture. Educational level, Division, Religion, Husband/partner’s Education level, Respondent Work status, are most significant factors that influence early marriage.

Education is also an important factor that makes change human behavior significantly. This study also proved that higher educated women have lower risk of early marriage. In context of Bangladesh, women education has not significantly increased still today. This study is an attempt to focus on various factors related to this issue. There is a scope of further improvement on this analysis considering many other demographic and economic variables.

REFERENCES

Bangladesh Demographic and Health Survey, 2007; National Institute of Population Research and Training (NIPORT) Dhaka, Bangladesh; Mitra and Associates, Dhaka, Bangladesh; ORC Macro, Calverton, Maryland USA.

Bangladesh, Unlocking the Potential, National Strategy for Accelerated Poverty Reduction; General economic division, Planning Commission, Government of people's Republic of Bangladesh (2005).

Bangladesh: Strategy for Sustained Growth, Bangladesh Development Series, Paper No.18;the World Bank Office, Dhaka, July 2007.

NIPORT 2013, Bangladesh Demographic and Health Survey 2011, Mitra& Associates, Dhaka, viewed 7 August 2015.

WMS (2009). Report on Welfare Monitoring Survey. Bangladesh Bureau of Statistics. Government of People's Republic of Bangladesh.

Cagatay N (1998) Gender and Poverty. Social Development and Poverty Elimination Division. United Nations Development Programme.

Hirschman, C. 1985,"Premarital Socioeconomic Roles and The Timing of Family Formation: A comparative Study of Five Asian Socities", Demography, 22(1):35-38.

Caldwell, J.C., Reddy, P.H. and Caldwell, P. 1983, "The Causes of Marriage Change in South India", Population Studies, 37: 343-361.

Ahmed, M., S.R.Nath and K.S. Ahmed, 2003, Education Watch 2002, literacy in Bangladesh need for a new vision overview of main report (Campaign for Population Education).

Begum, L., 2003. Meaning given to adolescent's reproductive health in Bangladesh. J.Popul.,9(1):82-103.

Central Intelligence Agency, 2007.The World Fact Book: Demographics of Bangladesh.

Islam, M.M. and M. Mahmud 1996. Marriage patterns and some issues related to adolescent marriage in Bangladesh. Asia-PacificPopul. J., 11(3):27-42.

Mitra, S.n., A.A.Sabir, T.Saha and S. Kumar, 2001. Bangladesh demographic and health survey 1999-2000. National Institute of Population Research and Training (NIPORT), Mitra and Associates, Dhaka; ORC Macro; Calverton, USA.

David W.Hosmer, S.Lameshow,(1989). Applied Logistic Regression (New York: John Wiley and Sons).

Davis Lucas (1979) Beginning Population studies Demography Teaching Notes No.**2**. Canberra: Development Studies Centre, Australian National University.

Kumchulesi, G, Palamuleni, M and Kalule-Sabiti, I 2011, 'Factors affecting age at first marriage in Malawi', Sixth African Population Conference, Ouagadougou- Burkina Faso, 5-9 December, viewed 17 September 2015.

Choe, MK, Thapa, S, and Mishra, V. 2005, 'Early Marriage and Early Motherhood in Nepal', Journal of Biosocial Science , vol. 7, no. 2, pp. 1-20, viewed 25 August 2015.