Journal of Science and Technology, 3(2): 439-449, July 2013

FITTING OF POLYNOMIAL MODELS TO AGE PATTERN OF FERTILITY IN BANGLADESH

MASHFIQUL HUQ CHOWDHURY* and MD. GOLAM RABBANI[1]

Department of Statistics, Mawlana Bhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh, [1]Department of Statistics, Biostatistics and Informatics, University of Dhaka, Dhaka-1000.

Abstract

The modeling of fertility patterns is an essential method for researchers to understand population patterns. The purpose of this study is to build up statistical models to age specific fertility rates (ASFR's) of Bangladesh. For this, the secondary data of age specific fertility rates (ASFR's) have been taken from Bangladesh Demographic and Health Survey (BDHS). This article presents some polynomial model that can adequately capture the varying patterns of the fertility curves of Bangladesh. Testing various polynomial models it is observed that (ASFR's) follows 4th degree inverse polynomial model in 1999-2000, 3rd degree inverse polynomial model in 2004, 4th degree inverse polynomial model in 2007, 3rd degree inverse polynomial model in 2011 and average ASFR between 1999-2011 follows 4th degree inverse polynomial model. Usual statistical procedure has been applied to these models to check their validity and for selecting the model which best fit. The performance of this model was also compared with Adjusted R_a^2 . The inverse polynomial models that are used seem to be fit well to the ASFR of the indicated years.

Keywords: Age Specific Fertility Rates (ASFR's), Inverse Polynomial Model, Adjusted R_a^2

Introduction

Over the last four decades, Bangladesh has experienced a large decline in fertility from a Total Fertility Rate (TFR) of more than 6 children per woman in the 1970's to 2.3 children per woman in the 2011 Bangladesh Demographic and Health Survey (BDHS), indeed a historic record in demographic transition. The government of Bangladesh is targeting lower total fertility rates to meet the Millennium Development Goals (MDGs) (United Nations, 2000). To achieve such a target a better understanding of the current pattern of age specific fertility rates is required. Statistical models, when well-constructed, can aid in this understanding as they provide better insight into some characteristics of the distributional pattern of fertility in Bangladesh. Modeling fertility in Bangladesh has also become necessary to enable a meaningful comparison of fertility across the countries in the region of the current fertility transition.

The study of reproductivity is primarily concerned to model fertility curves. Therefore for many years, modeling fertility curves has attracted the interest of demographers and remains the area of research (Gayawan *et al*., 2010). Islam and Ali (2004), it was found that age specific fertility rates (ASFR's) follows slightly modified bi-quadratic polynomial model whereas forward and backward cumulative ASFR's follow quadratic and cubic polynomial model, respectively in the rural area of Bangladesh.

* Author for correspondence: mashfiq3026@gmail.com

Traditionally, one can sketch some graphs of the demographic parameters. But very few of us know what types of functional form are appropriate for the parameters in the context of Bangladesh. In this study, an attempt has been made here to find out what types of statistical models are more appropriate to age specific fertility rates of Bangladesh. We will prefer those types of model whose number of parameters are as less as possible and interpretable in physical terms and are good enough to approximate all the relevant variations that are observable in the data. These statistical models will also help us in having a good platform for comparing the fertility performances of different cohorts across regions and over time horizon.

In this article, we propose statistical models using the inverse polynomial function applied to the generally available data about the age specific fertility rates on different age groups over different BDHS (1999-2000, 2004, 2007 and 2011) data. The proposed model is a flexible one that can capture various shapes of ASFR. It also provides a mathematical description of some fertility indices through its interpretable parameters. The efficacy of the model was determined by comparing its performance with model validation criteria and Adjusted $R_a{}^2$. These conceptual statistical models will derive information about fertility characteristics over time. Thus the main objectives of this study are as follows:

- ✓ To investigate the underlying patterns of observed ASFR of indicated periods.
- ✓ To fit an appropriate statistical model capturing the pattern of ASFR.
- ✓ To establish a general statistical model of ASFR.

The paper is mainly organized into four sections. 1st section is introduction. The pattern of age specific fertility in Bangladesh is described in the 2nd section. Section three contains materials and methods which includes data sources and methodology. We propose our model in section 4. The graphs of observed and fitted model to the fertility data are presented in section 5. This paper concludes with a discussion of the implications of our findings.

The Pattern of Age Specific Fertility Rate's in Bangladesh

Fertility is one of the principal components of population dynamics that determine the size and structure of the population of a country. To assess the fertility pattern in Bangladesh, we have used estimates of total fertility rate (TFR), which is the most widely used index of fertility. Fig. 1 clearly show that the TFR declines during the period 1975 to 2011 in Bangladesh.

The total fertility rate in Bangladesh has declined rapidly in the late 1980's and early 1990's and plateaued at around 3.3 for most of the 1990's. Despite the increase in contraceptive prevalence rate (CPR) over the period, TFR has remained at the same level during 1993-1999. The 2004 Bangladesh Demographic and Health Survey (BDHS) data indicate that after a long stagnation, the TFR declined slightly from 3.3 to 3.0 between 1999-2000 and 2004. Between the 2007 and 2011 there has been almost a 15 percent decline in the total fertility rate from 2.7 to 2.3 births per woman. The rural urban difference in fertility has narrowed over the past decade, from 1.1 births in 1999-2000 to 0.5 births in 2011.

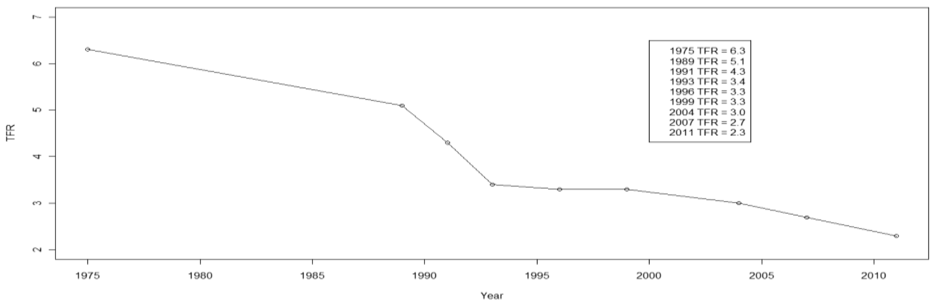

Fig. 1. Trends in Total Fertility Rates, 1975-2011 BDHS.

A plot of the ASFR reveals some differences in the age pattern of fertility among different BDHS survey. Figure II shows the five-year age pattern of fertility for Bangladesh. Some characteristics of fertility pattern can be observed from Fig. 2.

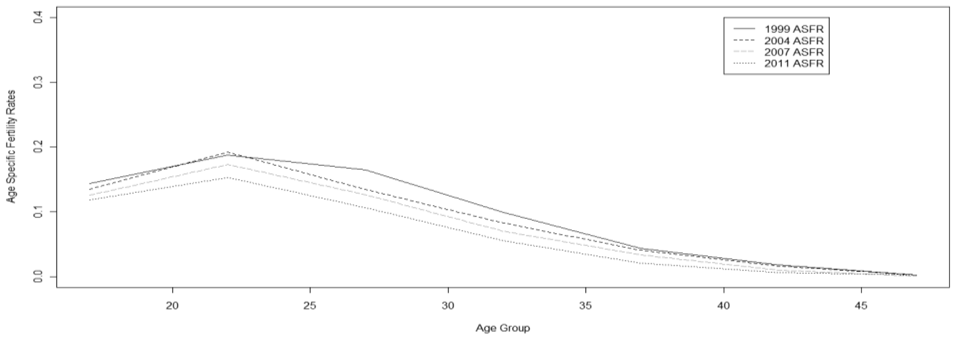

Fig. 2. Trends in Age Specific Fertility Rates from 1999 to 2011.

The pattern for 2011 fertility curves (Fig. 2) starts with a lower rate than those of the other fertility curves. It rises to reach the peak in the early twenties (age group: 20-24) before it starts dropping rapidly. Unlike BDHS-2011, rest of the fertility curves pattern starts with a relatively higher rate and peaks in the early twenties (age group 20-24). It does not demonstrate a rapid drop after the peak. In contrast to the other fertility curves, 1999 fertility curves depicts a broad peak shape that extends from the early twenties to the early thirties before a rapid drop. Additionally, whereas the tail of the curve almost approximates zero in the early forties for 2011, 2007 and 2004 BDHS, that of 1999-2000 is still much higher at the same point. Fertility curves also demonstrates that after a decade-long plateau in fertility (1993-1994 to 2000) at around 3.3 children per women there has been a steady and formidable decline in each subsequent years

Materials and Methods

Data Sources

This study uses secondary data from the Bangladesh Demographic and Health Survey (BDHS), 2011, a nationally representative sample survey. The 2011 BDHS survey was conducted under the authority of the National Institute for Population Research and

Training (NIPORT) of the Ministry of Health and Family Welfare (Mitra *et al.*, 2013). The data used in this study is publicly available. Other published reports from NIPORT are also used in order to support the analysis.

Methodology

It is widely known that the distribution of ASFR has a typical shape. The standard fertility shape depicting the age pattern of fertility with zero values before age 10, low but increasing positive values between 15-19, rising to a maximum between 20-29 and then decreasing slowly to reach zero by age 50. This non-linear pattern was the reciprocal of the V-shape. Some well-known families of models are polynomial models, exponential family and yield density models. To best capture the fertility pattern of Bangladesh, we have used polynomial models because other models cannot statistically describe the fertility pattern of Bangladesh.

Polynomial Model

In statistics, nonlinear regression is a form of regression analysis in which observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. The data are fitted by a method of successive approximations (Gujrati, 1998). In statistics, polynomial regression is a form of linear regression in which the relationship between the independent variable *x* and the dependent variable *y* is modeled as an n^{th} order polynomial. Polynomial regression fits a nonlinear relationship between the value of *x* and the corresponding conditional mean of *y*, denoted $E(y/x)$.

Using the scattered plot of age specific fertility rates by age group in years (Fig. 2), it is observed that age specific fertility rates can be fitted by polynomial model with respect to different ages in year. Therefore a n^{th} degree polynomial model is considered and the form of the model is

$$y = a_0 + a_1 x^1 + a_2 x^2 + \dots\dots + a_n x^n + u$$ (Gupta and Kapoor, 1997)

$$y = a_0 + \sum_{i=1}^{n} a_i x^i + u$$

Where, x is mid-value of an age group in years; y is ASFR's; a_0 is the constant term; a_i is the coefficient of x^i. $(i = 1,2,3,4,5,\dots\dots, n)$ and u is the stochastic error term of the model. Here, a suitable n is chosen for which the error sum of square is minimum.

Choice of Polynomial Model

Choosing of an appropriate model largely depends on the judicious selection of the functional form that enables to capture the inherent peculiarities of ASFR. A variety of polynomial models are tried to fit (Appendix 1) on the observed ASFR of indicated years and tested for statistical validation. Firstly, separate polynomial models are tested for different years and the attempt is taken to provide a general common model for ASFR. The common model is tried on the average ASFR's of indicated years. To fit the polynomial model we assume ASFR's the dependent variable (y) and mid-value of an age group is the independent variable (x). The functional forms of the selected polynomial models are as follows:

For the ASFR of 1999-2000: $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$

For the ASFR of 2004: $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$

For the ASFR of 2007: $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$

For the ASFR of 2011: $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$

The common model: For fitting a common model that represent a common pattern of ASFR for the years 1999-2000, 2004, 2007, 2011 we have averaged the ASFR of the study years. The average ASFR's are given in appendix. Recoursing various types of polynomial models on average ASFR the following form of the models seems to fit well to average ASFR.

$$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$$

The results of the fitted models are provided in table 1.

Model Validation

F-test: The F-test is used to the model to verify the overall measure of the significance of the model as well as the significance of R^2. The formula for F-test is given as with (*m-1*, *n-m*) degrees of freedom (degrees of freedom); where m is the number of parameters of the fitted model, n is the number of cases and R^2 is the coefficient of determination of the model (Gujarati, 1998).

$$F = \frac{R^2 / (m-1)}{(1-R^2) / (n-m)} \quad \text{with } (m\text{-}1,\ n\text{-}m) \text{ degrees of freedom}$$

Where m is the number of parameters of the fitted model n is the number of cases and R^2 is the coefficient of determination of the model (Gujarati, 1998).

Adjusted $R_a{}^2 = 1 - (1 - R^2)\left(\frac{n-1}{n-m}\right)$

Results

We present the following model for the ASFR pattern of BDHS-1999-2000, 2004, 2007 and 2011, respectively in Table 1.

Table 1: Fitted model of ASFR for BDHS 1999, 2004, 2007, 2011 and general model.

Proposed Model	Estimated Model (P-values of the coefficients are given in braces respectively)	Evaluation of the Model
1999 BDHS Model $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$	$\hat{y} = 1.67 - \frac{211.7}{x} + \frac{9297.44}{x^2} - \frac{163700.7}{x^3} + \frac{1008588.9}{x^4}$ (0.03) (0.02) (0.02) (0.02) (0.02)	This fitted model explains 99.82% of total variation and highly significant at 0.1% level of significance with F (4, 2) degrees of freedom. F statistic is 828.4.
2004 BDHS Model $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$	$\hat{y} = 0.1872 - \frac{27.325}{x} + \frac{1134.1}{x^2} - \frac{11612.01}{x^3}$ (0.008) (0.002) (0.0004) (0.0002)	This fitted model explains 99.94% of total variation and highly significant at 0.009% level of significance with F (3, 3) degrees of freedom. F statistic is 3232.
2007 BDHS Model $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$	$\hat{y} = 0.7195 - \frac{90.06}{x} + \frac{3764.13}{x^2} + \frac{59184.97}{x^3} + \frac{311194.06}{x^4}$ (0.03) (0.03) (0.03) (0.04) (0.04)	This fitted model explains 99.93% of total variation and highly significant at 0.04% level of significance with F (4, 2) degrees of freedom. F statistic is 2333.33.
2011 BDHS Model $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$	$\hat{y} = 0.2700 - \frac{32.68}{x} + \frac{1182.08}{x^2} - \frac{11399.23}{x^3}$ (0.04) (0.02) (0.007) (0.004)	This fitted model explains 99.37% of total variation and highly significant at 0.03% level of significance with F (3, 3) degrees of freedom. F statistic is 319.1.
General BDHS Model $y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$	$\hat{y} = 0.9312 - \frac{116.07}{x} + \frac{4913.06}{x^2} - \frac{80441.89}{x^3} + \frac{451050.5}{x^4}$ (0.006) (0.005) (0.005) (0.006) (0.008)	This fitted model explains 99.98% of total variation and highly significant at 0.01% level of significance with F (4, 2) degrees of freedom. F statistic is 8568.43.

Graphs

We present the following graphs for the empirical and estimated ASFR for BDHS 1999, 2004, 2007, 2011 and general model in fig. 3-7.

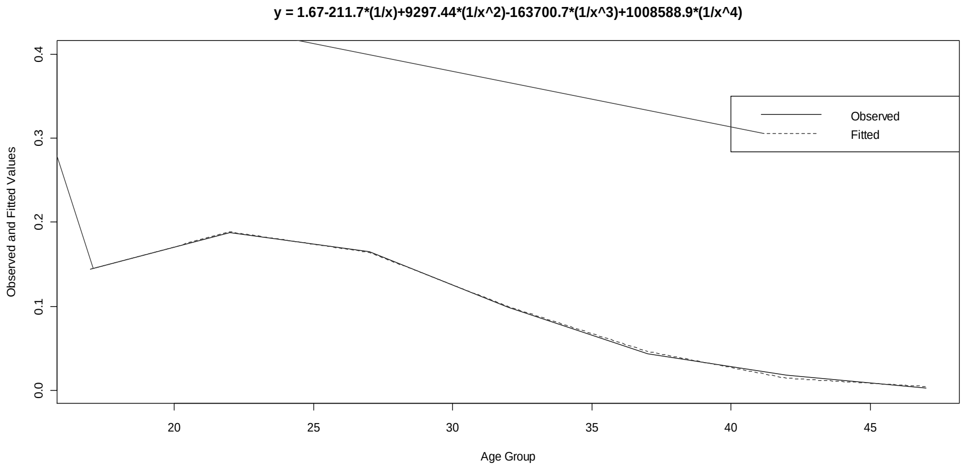

Fig. 3. Fitting of 4th degree inverse polynomial model for 1999-2000 BDHS.

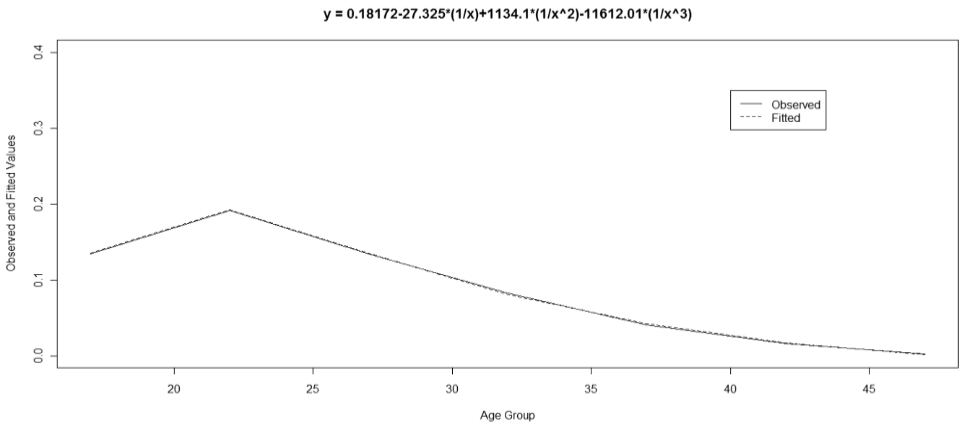

Fig. 4. Fitting of 3rd degree inverse polynomial model for 2004 BDHS.

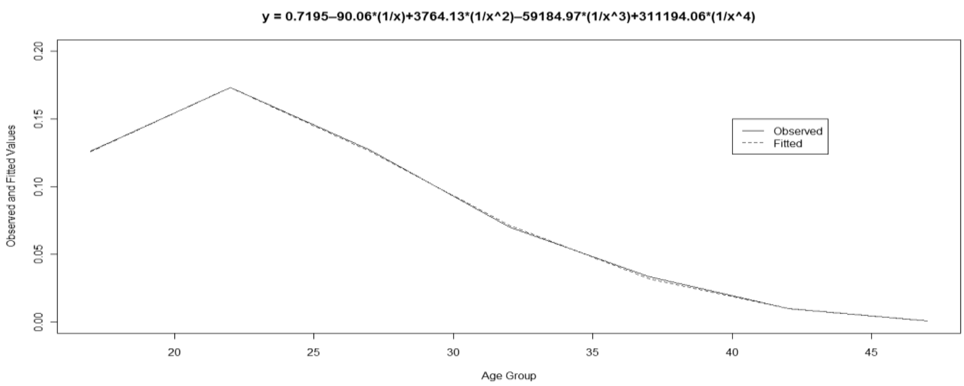

Fig. 5. Fitting of 4th degree inverse polynomial model for 2007 BDHS.

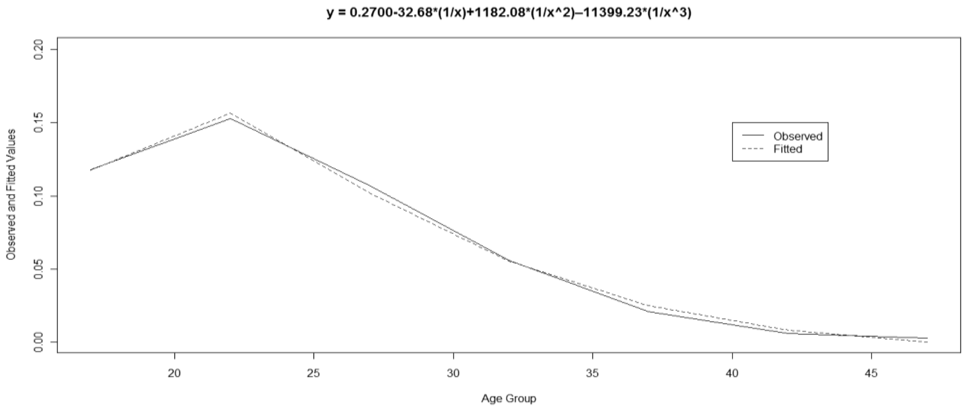

Fig. 6. Fitting of 3rd degree inverse polynomial model for 2011 BDHS.

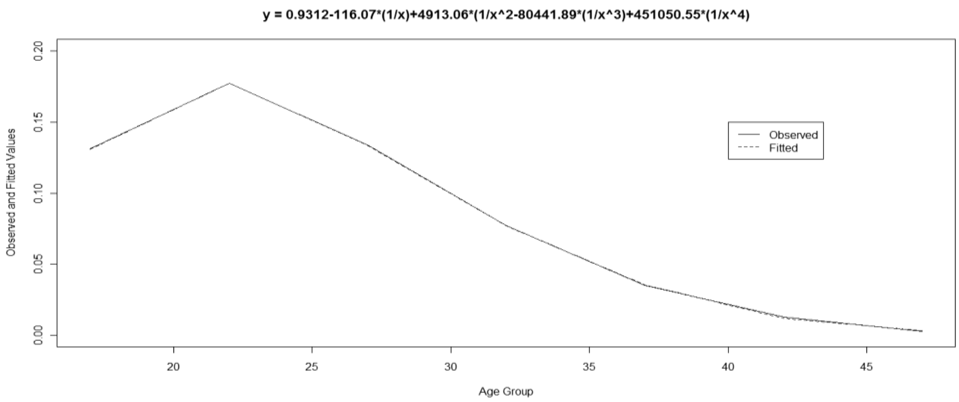

Fig. 7. Fitting of 4th degree inverse polynomial model for general ASFR (1999-2011).

Evaluation of the Model

Table1 present the results of the estimates. Fig. 3 -7 present how well the model fits to the observed fertility rates. The Tables provide the proposed model, significance of the model, variation explained by the proposed model and significance of each parameter. It is shown that all the parameters of the fitted models are statistically significant. Adjusted R^2 $(R_a^{\ 2})$ are more than 0.99 for all of the models. This means all of the proposed models explain more than 99% variation of the dependent variable, age specific fertility rate. The calculated value of F-test and probability value for all of the models parameter are presented in each table. These results are indicated that these models and their coefficients are highly statistically significant and hence we can conclude that all of these models are well fitted into the data set.

Implication

Knowledge about fertility trends and fitting of models may help us learn more about the consequences of low fertility and see clearer whether interventions may be justified and what specific steps one might take. This study of ASFR's in Bangladesh show the declining trend and the traditional reciprocal of V-shape pattern. It is also concluded that the woman of age interval 20-24 years is the most fertile and the age group 45-49 are the least fertile period in the reproductive life of Bangladeshi women.

The statistical models of ASFR of Bangladesh for the year1999 to 2011 have been fitted and these models have followed inverse polynomial models. Since all of these models are statistically significant and provides good fit for the data so these models could reveal important parameters which need to be taken into account when comparing fertility between countries and across time. This undoubtedly would increase our understanding of fertility pattern in the region.

References

Gayawan E., B. Samson Adebayo, Reuben A. Ipinyomi and Benjamin A. Oyejola. (2010). Modeling fertility curves in Africa, Volumes/Vol22/10/ DOI: 10.4054 / Dem Res. 2010.22.10

Gujarati, N. Damodar (1998). Basic econometrics, third edition, McGraw Hill, Inc., New York.

Gupta, S. C. and V.K. Kapoor. (1997). Fundamentals of mathematical statistics, Ninth Edition, New Delhi.

Islam, M. R. and Ali, M. Korban. (2004). Mathematical modeling of age specific fertility rates and study the reproductivity in the rural area of Bangladesh during 1980-1998, Pak. J. Stat., **20**(3): 379-392.

Mitra, S.N. and Associates. (2013). Bangladesh demographic and health survey, 2011, National institute of population research and training (NIPORT), Dhaka, Bangladesh.

United Nations (2000), Millennium development goals report, retrieved from http://www.un.org/millennium/goals

Appendix

Table 2: Trends in Age specific fertility rates and Total fertility rates among women age 15-49, selected sources, Bangladesh, 1999-2011.

Age Group	1999-2000 BDHS	2004 BDHS	2007 BDHS	2011 BDHS	Average ASFR
15-19	0.144	0.135	0.126	0.118	0.131
20-24	0.188	0.192	0.173	0.153	0.177
25-29	0.165	0.135	0.127	0.107	0.134
30-34	0.099	0.083	0.070	0.056	0.077
35-39	0.044	0.041	0.034	0.021	0.035
40-44	0.018	0.016	0.010	0.006	0.013
45-49	0.003	0.003	0.001	0.003	0.003
TFR	3.3	3.0	2.7	2.3	

Table 3: Fitted polynomial models of 1999-2000 BDHS.

Assume Model	Estimated Model (P-values of the coefficients are given in braces respectively)	Comment on the Model
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$	$\hat{y} = -0.1051 - \frac{3.6124}{x} + \frac{540.08}{x^2} - \frac{6920.12}{x^3}$ (0.65) (0.85) (0.34) (0.18)	This model explains 97% of total variation but is insignificant at 5% level of significance.
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$	$\hat{y} = 1.67 - \frac{211.7}{x} + \frac{9297.44}{x^2} - \frac{163700.7}{x^3} + \frac{1008588.9}{x^4}$ (0.03) (0.02) (0.02) (0.02) (0.02)	This is the valid model. Interpreted at table 1.

Note: Inverse polynomial model of first and second degree are also tested and found to be not significant at 5% level of significance.

Table 4: Fitted polynomial models of 2004 BDHS.

Assume Model	Estimated Model (P-values of the coefficients are given in braces respectively)	Comment on the Model
$y = a + \frac{b}{x} + \frac{c}{x^2}$	$\hat{y} = -0.4198 + \frac{24.41}{x} - \frac{252.32}{x^2}$ (0.004) (0.004) (0.007)	This model explains only 93% of total variation.
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$	$\hat{y} = 0.1872 - \frac{27.325}{x} + \frac{1134.1}{x^2} - \frac{11612.01}{x^3}$ (0.008) (0.002) (0.0004) (0.0002)	This Model is best than the above model because it explains 99.93% of the total variation.

Note: First degree inverse polynomial model also tested and found to be not significant at 5% level of significance.

Table 5: Fitted polynomial models of BDHS-2007.

Assume Model	Estimated Model (P-values of the coefficients are given in braces respectively)	Comment on the Model
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$	$\hat{y} = 0.1718 - \frac{25.8541}{x} + \frac{1062.10}{x^2} - \frac{10811.25}{x^3}$ (0.08) (0.02) (0.006) (0.003)	This model explains 99.62% of total variation but is insignificant at 5% level of significance.
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$	$\hat{y} = 0.7195 - \frac{90.06}{x} + \frac{3764.13}{x^2} + \frac{59184.97}{x^3} + \frac{311194.06}{x^4}$ (0.03) (0.03) (0.03) (0.04) (0.04)	This is the valid model. Interpreted at table 1.

Note: Inverse polynomial model of first and second degree are also tested and found to be not significant at 5% level of significance.

Table 6: Fitted polynomial models of BDHS– 2011

Assume Model	Estimated Model (P-values of the coefficients are given in braces respectively)	Comment on the Model
$y = a + \frac{b}{x} + \frac{c}{x^2}$	$\hat{y} = -0.3205 + \frac{18.1045}{x} - \frac{178.8959}{x^2}$ (0.01) (0.01) (0.02)	This model explains only 930% of total variation.
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$	$\hat{y} = 0.2700 - \frac{32.68}{x} + \frac{1182.08}{x^2} - \frac{11399.23}{x^3}$ (0.04) (0.02) (0.007) (0.004)	This Model is best than the above model because it explains 99.93% of the total variation.

Note: First degree inverse polynomial model also tested and found to be not significant at 5% level of significance.

Table 7:Fitted polynomial models for average ASFR.

Assume Model	Estimated Model (P-values of the coefficients are given in braces respectively)	Comment on the Model
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3}$	$\hat{y} = 0.1374 - \frac{23.01}{x} + \frac{996.68}{x^2} - \frac{10328.11}{x^3}$ (0.24) (0.06) (0.01) (0.01)	This model explains 99.31% of total variation but is insignificant at 5% level of significance.
$y = a + \frac{b}{x} + \frac{c}{x^2} + \frac{d}{x^3} + \frac{e}{x^4}$	$\hat{y} = 0.9312 - \frac{116.07}{x} + \frac{4913.06}{x^2} - \frac{80441.89}{x^3} + \frac{451050.5}{x^4}$ (0.006) (0.005) (0.005) (0.006) (0.008)	This is the valid model. Interpreted at table 1.

Note: Inverse polynomial model of first and second degree are also tested and found to be not significant at 5% level of significance.