Journal of Engineering and Applied Sciences 3 (2): 203-206, 2008
ISSN: 1816-949X

Modeling and Predicting Male Urban Forward Cumulative Population of Bangladesh

Rafiqul Islam
Department of Population Science and Human Resource Development,
University of Rajshahi, Bangladesh

Abstract: The purpose of the present study is to build up models to male urban forward cumulative population of Bangladesh. For this, the secondary data for male population of urban area of Bangladesh is taken from various issues of censuses. To check up the validity of the model, the model validation technique, Cross-Validity Predictive Power (CVPP) is applied. It is seen that forward cumulative distribution of population follows polynomial model.

Key words: Male urban forward cumulative population, polynomial model, Cross-Validity Prediction Power (CVPP), F-test

INTRODUCTION

In the up to date high technological and nuclear era, statistical models are very sophisticated apparatus to represent data. Statistical model is very much important in differentiating necessary and unnecessary characteristics among various socio-economic and demographic phenomena. Modeling is, in fact, essentially an attempt to find out the functional relations and their dynamic behaviors among a variety of components not only in demographic study but also all disciplines of knowledge. Conventionally, one can draw out some figures for the demographic parameters as well as socio-economic indicators due to cress sectional data or time trend data. But, in Bangladesh, very few of us comprehend which types of functional or mathematical form are more apposite for the parameters and social indicators. Model is finally important for prediction purposes. The age structure for male, female and both sexes population of Bangladesh was followed either negative exponential model or modified negative exponential model (Islam, 2003, 2005). Islam *et al*. (2003) found that age structure for male population of Bangladesh in 1991 follows modified negative exponential model. It was observed that age structure for population of both sexes of Bangladesh follows negative exponential model (Islam *et al*., 2005).

At this stage, an attempt would be made to mull over to build statistical model for male population of urban region of Bangladesh for the years 1981, 1991 and 2001 censuses using Quasi-Newton method by handling the software STATISTICA. Therefore, the basic aims and objectives of this study are to build up statistical models to male urban forward cumulative population of Bangladesh and thereafter the prediction is done.

MATERIALS AND METHODS

Data collection: To accomplish the objectives aforementioned, the secondary data of quinquennial age group on male population of urban area in Bangladesh is taken from various volumes of censuses 1981 (BBS, 1984), 1991 (BBS, 1994) and 2001 (BBS, 2003) counted as ten years apart. These have been utilized as raw materials in this study and shown in Table 1. The population is in thousands.

Data smoothing: If the male population of urban area by age group is put in graph paper then it is observed that there is some sort of unexpected distortions in the data. So, before going to fit the model to this data series, an adjustment is important and needed to alleviate these distortions. Therefore, an adjustment is made here using the Package Minitab Release 12.1 by the latest smoothing method "4253H, twice" (Velleman, 1980). Thereafter, the smoothed data are used to fit mathematical model to male urban population of Bangladesh and these smoothed data are launched in Table 1.

Model fitting: It seems from the scattered plot of forward cumulative distribution of male urban population of Bangladesh by age group that it can be distributed by polynomials for different ages. Therefore, an nth degree polynomial model is chosen and the structural formula of the model is given by:

$$y = a_0 + \sum_{i=1}^{n} a_i x^i + u \text{ (Montgomery and Peck, 1982)}$$

where, x is the mid value of the age group; y is forward cumulative distribution of male urban population; a_0 is the

Table 1: Observed, smoothed and predicted male urban forward cumulative population by age group of Bangladesh for the years 1981, 1991 and 2001 censuses

Age group	Forward cumulative distribution (in thousands)								
	1981			1991			2001		
	Observed	Smoothed	Predicted	Observed	Smoothed	Predicted	Observed	Smoothed	Predicted
0-4	955	955	998	1438	1474	1491	1588	1588	1407
5-9	1890	1890	1934	2962	2921	2926	3250	3330	3408
10-14	2837	2780	2793	4322	4239	4245	5132	5100	5254
15-19	3581	3596	3577	5360	5414	5449	6816	6818	6943
20-24	4322	4338	4286	6475	6510	6539	8410	8421	8476
25-29	5031	4998	4918	7579	7539	7513	9918	9887	9854
30-34	5559	5556	5475	8438	8455	8373	11169	11185	11075
35-39	6009	6008	5956	9244	9221	9117	12296	12288	12140
40-44	6369	6366	6361	9841	9823	9747	13205	13184	13049
45-49	6628	6641	6690	10259	10267	10261	13860	13867	13802
50-54	6861	6850	6943	10588	10581	10661	14374	14353	14399
55-59	6991	7006	7121	10783	10806	10945	14658	14688	14840
60-64	7144	7127	7223	10995	10976	11115	14959	14937	15125
65-69	7209	7242	7249	11097	11132	11169	15114	15171	15253
70+	7370	7370	7199	11301	11301	11108	15433	15433	15226

constant; a_1 is the coefficient of x^i (i = 1, 2, 3, ..., n) and u is the stochastic error term of the model. Here, an appropriate n is chosen such that the error sum of square is bare smallest amount.

Model validation: To test out out the adequacy or soundness of the models, the Cross Validity Prediction Power (CVPP), ρ^2_{cv}, is applied. The mathematical formula for CVPP is:

$$\rho^2_{cv} = 1 - \frac{(n-1)(n-2)(n+1)}{n(n-k-1)(n-k-2)}(1-R^2)$$

where, n is the number of classes, k is the number of explanatory variables in the model and the cross-validated R is the correlation between observed and predicted values of the dependent variables (Stevens, 1996). The shrinkage coefficient of the model is the positive value of $\lambda = (\rho^2_{cv} - R^2)$; where, ρ^2_{cv} is CVPP and R^2 is the coefficient of determination in the model. λ tends to zero indicating better prediction of the model. Additionally, 1-shrinkage is the stability of R^2 of the model. The estimated CVPP analogous to their R^2 and information on model fittings are presented in Table 2. It was informed that CVPP was also employed by Islam (2003, 2005) and Islam *et al.* (2003, 2005) as model validation technique.

F-test: To find out the measure of overall significance level of the fitted models as well as the significance of R^2, the F-test is employed here. The F-test is given by

$$F = \frac{R^2/(l-1)}{(1-R^2)/(n-l)}$$

with (l-1, n-l) degrees of freedom (d.f.); where, l = the number of parameters is to be estimated, n is the number of cases and R^2 is the coefficient of determination of the model (Gujarati, 1998).

RESULTS AND DISCUSSION

The polynomial model is assumed to fit the model to male urban forward cumulative population of Bangladesh in various censuses and the fitted models are described in the following way:

The fitted model for distribution of male population for urban area of Bangladesh in 1981 is:

$$y = 501.7943 + 202.2827x - 1.5159x^2 \quad (1)$$
t-stat (7.6114) (49.7029) (-28.8359)

The fitted model for forward cumulative distribution of population for male of urban area in 1991 is:

$$y = 730.8329 + 309.8861x - 2.3x^2 \quad (2)$$
t-stat (9.7832) (67.1968) (-38.6101)

The fitted model for forward cumulative distribution of population for male of urban in 2001 is:

$$y = 347.6746 + 431.5025x - 3.1212x^2 \quad (3)$$
t-stat (3.091) (62.1439) (-34.799)

The information on model fittings and estimated CVPP, ρ^2_{cv}, analogous to their R^2 of these models is shown in Table 2. From this table it appears that all the fitted models 1-3 are highly cross- validated and their shrinkage coefficients are very small quantity.

Table 2: Information on model fittings and CVPP

Models	n	K	R^2	ρ_{cv}^2	Shrinkage	Parameters	p-value	Cal. F	Tab. F (at 1% level)
1	15	2	0.99860	0.998258	0.000342			4279.71	
						a_0	0.000		6.93
						a_1	0.000		with
						a_2	0.0000		(2, 12) d.f
2	15	2	0.99925	0.999067	0.000183			7994	
						a_0	0.000		6.93
						a_1	0.000		with
						a_2	0.000		(2, 12) d.f
3	15	2	0.99917	0.998967	0.000203			7222.92	
						a_0	0.0093		6.93
						a_1	0.000		with
						a_2	0.0000		(2, 12) d.f

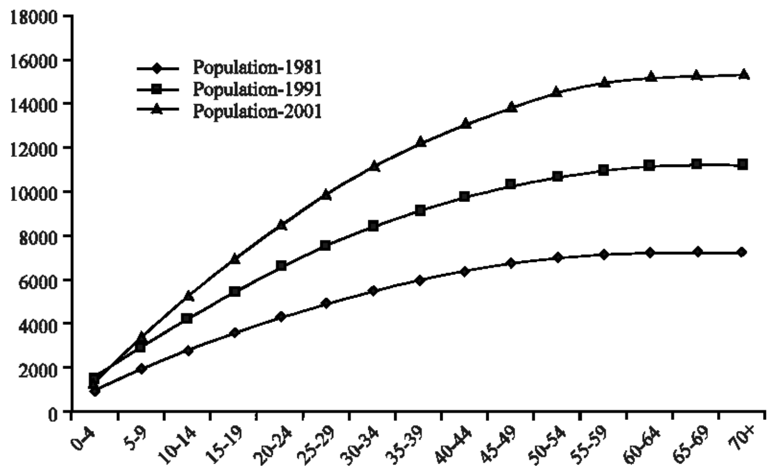

Fig. 1: Predicted male urban forward cumulative population of Bangladesh for the census years 1981, 1991 and 2001. X axis indicates age group in years and y axis represents predicted cumulative population

Furthermore, these fitted models are highly significance at 1% level that is seen from the Table 2. Therefore, from these statistics it is seen that these models and their analogous R^2 are highly statistically significant. In addition, it is also found from the same table that these models are explaining more than 99% variation of the total and more than 99% stable. Nevertheless, the stability of R^2 of these models is exceeding 99%. Hence, the fits of these models are well. Therefore, they provide better prediction.

These predicted cumulative populations are plotted in graph paper that is shown in Fig. 1. It is found that these are remarkably upward trend over time.

It is noted that other as usual models, that is, Logistic, Gompertz, Makeham, log linear model, semi-log linear model, double log linear model, inverse linear model were also tried to utilize here but seems to be worst fitted due to their coefficient of determination and shrinkage coefficients. For that reason, the outcomes of those models were not exposed here.

CONCLUSION

In this study, it is observed that forward cumulative distribution of population follows three parameters 2nd degree polynomial, i.e., quadratic polynomial model. Hope one might be used this predicted population of Bangladesh during 1981-2001 censuses for further higher study as more smoothened data than observed data.

REFERENCES

BBS, 1984. Bangladesh Population Census 1981. National Series, Government of the People's Republic of Bangladesh, Dhaka.

BBS, 1994. Bangladesh Population Census 1991. National Series, Government of the People's Republic of Bangladesh, Dhaka, Vol. 1.

BBS, 2003. Bangladesh Population Census 2001. National Report, Government of the People's Republic of Bangladesh, Dhaka.

Gujarati, D.N., 1998. Basic Econometrics. 3rd Edn. McGraw Hill, Inc., New York.

Islam Rafiqul, 2003. Modeling of Demographic Parameters of Bangladesh-An Empirical Forecasting, Unpublished Ph.D. Thesis, Rajshahi University.

Islam Rafiqul, Islam, Nurul, Ali, Ayub and Mostofa Golam, 2003. Construction of Male Life Table from Female Widowed Information of Bangladesh, Department of Statistics, University of Rajshahi, Bangladesh. Int. J. Stat. Sci., 2: 69-82.

Islam Rafiqul, 2005. Construction of Female Life Table from Male Widowed Information of Bangladesh. Pak. J. Stat., 21 (3): 275-284.

Islam Rafiqul, Islam, Nurul, Ali, M. Korban and Mondal Nazrul Islam, 2005. Indirect Estimation and Mathematical Modeling of Some Demographic Parameters of Bangladesh. The Oriental Anthropol., 5 (2): 163-171.

Stevens, J., 1996. Applied Multivariate Statistics for the Social Sciences. 3rd Edn. Lawrence Erlbaum Associates, Inc., Publishers, New Jersey.

Velleman, P.F., 1980. Definition and Comparison of Robust Nonlinear Data Smoothing Algorithms. J. Am. Stat. Assoc., 75 (371): 609-615.