ORIGINAL RESEARCH PAPER
Volume-7 | Issue-6 | June-2018 | PRINT ISSN No 2277 - 8179
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

USE OF SIMPLE EXTRAPOLATION METHODS FOR POPULATION PROJECTIONS OF INDIA

Statistics

Pandurang V. Thatkar* Research Scholar, Department of Statistics and Mathematics, NES Science College, Nanded *Corresponding Author

Kishor Y. Ingale Assistant Professor of Statistics, Dept. of Mathematics and Statistics, NSB College, Tarasingh Market, Nanded

D. D. Pawar Professor & HOD, Department of Statistics, NES Science College, Nanded

KEYWORDS

INTRODUCTION

Generally the information about population totals is not available except Census year. Hence we need to estimate population totals for the intermediate years. The information about future or past population totals is called an estimate.

The population estimates can be obtained in many ways. Some methods use the past or most recent census using component, regression or ratio technique. for these methods data from administrative records or sample surveys is used. Some other methods develop estimates using interpolation technique over the dates between censuses. few methods gives estimated population figures by age, sex, race and other variety of socioeconomic and demographic characteristics. whereas some methods provide the estimated directly for the total population.

Generally, the term projection or a forecast is often used interchangeably by the demographers. Information about the future is termed as either a projection or a forecast. These terms can be differentiated according to the likelihood of their expected outcome.

A projection, under certain set of assumptions, is defined the numerical outcome of the future population. The conditional calculations provides a projection which can answer the future population if a particular set of assumptions hold true. Although the projected population figures can be judged by the merits of its assumptions.

A forecast is a projection that is most likely accurate population prediction. Forecast represents a precise standpoint concerning the legality of the primary data and assumptions. A forecast replicate a verdict and it can be established as right or wrong by future events. The term projection is a more inclusive term than forecast: Forecasts are subset of projections that is all forecasts can be projections but not vice-versa. Both Projections and forecasts include information regarding age, sex, race, and other properties of population totals.

METHODOLOGY:

Definitions of terms used in study:

1. *the base year:* the initial year of the data used for predictions.
2. the launch year: the latest year of the data used for predictions.
3. the target year: the year for which the predictions were obtained.
4. the base period: the difference between the base year and launch year, while
5. the projection horizon: the difference between the launch year and target year.
6. the projection interval: the number of years between the two predictions made.

Statistical Methods of Projection of Population

1. Trend Extrapolation

i. Simple extrapolation methods
 a. Linear Change
 b. Geometric Change
 c. Exponential Change

ii. Complex extrapolation methods
 a. Linear Models
 b. Polynomial Models
 c. Logistic Models
 d. ARIMA Time Series Models

iii. Ratio extrapolation methods
 a. Constant-Share
 b. Shift-Share
 c. Share-of-Growth

2. Cohort-Component Method

a. Projecting Mortality
b. Projecting Fertility
c. Projecting Migration

For the purpose of this research work, only the simple extrapolation methods will be discussed.

Trend Extrapolation:

Historical or past data is fitted into a mathematical model and these models are used to project the future population figures. This method is called as trend extrapolation. Such models require relatively less data and less resources. Hence costs involved in obtaining population projections using trend analysis is generally low. Hence use of models is inevitable not only in population studies but also in other fields as well (Armstrong, 2001).

There are many different methods of trend extrapolation which are generally the trend extrapolation methods are organized into three categories as below.

i. Simple extrapolation methods
ii. Complex extrapolation methods
iii. Ratio extrapolation methods

Simple extrapolation methods requires data for only two data points (dates). In simple extrapolation methods, we use three approaches as:

a. Linear change,
b. Geometric change, and
c. Exponential change

Complex extrapolation methods requires data for a number of data points (dates). In complex extrapolation methods we use four approaches as:

a. Linear trend,
b. Polynomial curve,
c. Logistic curve, and
d. ARIMA time series;

Ratio extrapolation methods considers the population of a lesser area and is articulated as a proportion of the population of its bigger, area (parent area). In ratio extrapolation methods, we use three approaches as:

a. Constant-share,
b. Shift-share
c. Share-of-growth.

Trend extrapolation techniques are frequently used for the population progressions of total population of a larger area. These techniques are rarely used for projection of subgroups of the population.

Linear Change: Linear change is a method of simple extrapolation which considers that the future population will change by the rate of change similar to the given period as accrued during the base period.. Here the period may be a month, a year or a decade.

Geometric change: Geometric method is used under the assumption that the population will be changing by the same rate (percentage) over a given addition of time in the future as during the base period.

Exponential Change: The another simple extrapolation method is known as the exponential change method. This exponential change method is very much associated to the geometric change method, but it assumes change as occurring continuously instead of the change occurring at discrete intervals.

Linear Trend: Among the complex extrapolations methods, linear trend models are the simplest In linear trend the basic assumption is that the population change in future will follow the similar pattern as in the past. This means population will change at a same rate in the future as in the past. The assumptions underlying linear trend are similar to the assumptions underlying the simple linear method.

Polynomial Models: For population projections, polynomial models are useful when change is not guarded to be linear or the change is non-linear. The polynomial models includes more than one independent variable (here it is time) and the coefficients include both a measure of the linear trend (b_1) and measure of non-linear patterns (trends) (b_2,b_3, ..., b_n).

Logistic Models: The logistic model approach unequivocally allows us to put an upper limit on the final population size for a particular region. The logistic model is dependable through Malthusian and other theories of constrained population growth.

ARIMA Time Series Models:

ARIMA models (Autoregressive Integrated Moving Average) have been used in the analysis and projection of populations as a whole and of their demographic attributes (Alho and Spencer, 1997; Carter and Lee, 1986; De Beer, 1993; Land, 1986; Lee, 1993; McDonald, 1979; McKnown and Rogers, 1989; Pflaumer, 1992; and Saboia, 1974). The procedures used in ARIMA models are complex, making them hard to execute and put in plain words to data users.

ARIMA models endeavor to expose the stochastic mechanisms that produce historical data series and use this information as a basis for developing new projections. Following processes are involved in the ARIMA models that can describe the stochastic mechanism:

1. autoregressive;
2. differencing; and
3. a moving average.

The autoregressive process has a memory in the sense that it depends on the correlation of each value of a variable with all previous figures. The impact of earlier figures gets reduced over the period of time exponentially. The number of preceding figures/values that are included in the model decides the order of model. As a result, all previous figures/values manipulate current values, though with a declining impact. In a second-order ARIMS model, the current value is clearly a function of the two directly previous values; again, all previous values have an indirect impact.

Generally, the ARIMA model is written as ARIMA (p,d,q).
Where,p: is the order of the autoregression,
d: is the degree of differencing, and
q: is the order of the moving average

Ratio Extrapolation Methods:

Ratio extrapolation methods are useful when an area containing the population to be projected is part of a parent area for which projections are available. Ratio extrapolation methods are generally used where areas is a perfect hierarchical arrangement; which means, the geographical areas at each level are mutually exclusive and exhaustive, so that they can be summed to higher levels, producing in one all-inclusive unit.

For example, consider census blocks in the India, which can be aggregated successively into Districts, Regions, States and finally at the country level.

Mainly there are three commonly used ratio methods:

1. Constant Share;
2. Shift-Share: and
3. Share-of-Growth.

Here the point to be noted is that, above all three methods require projections of a parent area which covers the area of interest.

Constant Share method:

This technique assumes, the subarea area's contribution of the parent area's population is considered as constant at a stage recorded during the base period. Naturally it will be the contribution of the launch year.

The constant share method needs only historical data from one point in time; hence, it is greatly useful for areas in which it is not possible to form the reliable data series due to poor quality of records and when geographical boundaries changes.

One more important characteristics of this method is that the projections of all smaller areas can be added to form the population of projection for the parent area. The demerit of this method is that this method considers same growth rate (as parent area) for all smaller areas

Shift-Share:

The shift-share method is planned in such a way that it can deal with changes in shares of population. The main demerit of shift-share method is that in areas which grows very slowly, it can guide to substantial population during the base period, This problem generally arises while projecting population for a long range horizons (e.g., 20 or 30 years). Consequently, for the areas with very high rapid population growth, this method can lead to ridiculously high projections .

Hence, while using shift-share method for projections, one needs to be very cautious for long-range projections, especially for places whose population shares have growing slowly or declining or increasing rapidly.

Share-of-Growth:

This is the third ratio method which considers shares of population change relatively than size of the population. This method assumes same share of the smaller area's share of population change in the parent area over the projection horizon as it was during the base period.

In many cases, the share-of-growth method is found to generate more reasonable projections than both constant or shift-share methods. The demerit of this method is that, projections generated by this method will be quiet unreliable when a growth rate in a smaller area has the opposite sign than that for the parent area.

DATA COLLECTION:

For the purpose of obtaining population Projections of India using various statistical techniques, the population figures were collected for each of the census starting from 1901 to 2011.

SYSTEM IMPLEMENTATION:

The data collected was entered into a excel sheet and cleaned for any errors. The formulas explained in the above section were applied for each method of population projection separately and the population projections were obtained as shown in the table 1. These population projections were obtained for Males, Females and Total Population and also compared with the Population figures given by Indian Census 2011.

RESULTS:

Table1: Projection Results for India according to Gender-wise and Total population Using Different Extrapolation Methods, 2006, 2011 & 2016 (in '000)

	Male Population			Female Population			Total Population		
Method/Year	**2006**	**2011**	**2016**	**2006**	**2011**	**2016**	**2006**	**2011**	**2016**
Projected Population (According to Census 2001)	601,213	642,105	682,953	560,875	605,341	643,849	1,162,088	1,247,446	1,326,802
SIMPLE EXTRAPOLATION METHODS									
Linear	566,893	601,562	636,231	528,610	560,705	592,801	1,088,809	1,148,881	1,229,032
Geometric	575,675	622,675	673,512	537,051	580,897	628,323	1,112,726	1,203,572	1,301,835
Exponential	575,801	622,948	673,955	537,169	581,152	628,737	1,112,970	1,204,100	1,302,691
COMPLEX EXTRAPOLATION METHODS									
Linear Trend	563,478	600,628	637,777	527,155	561,908	596,661	1,090,633	1,162,536	1,234,439
Quadratic	575,806	628,025	682,983	540,076	590,622	644,038	1,115,883	1,218,646	1,327,021
Logistic	595,857	658,523	727,781	561,718	620,794	686,084	1,157,479	1,279,213	1,413,749
ARIMA	630,272	700,140	777,754	588,500	653,737	726,207	1,218,771	1,353,877	1,503,961
RATIO EXTRAPOLATION METHODS									
Constant	578,013	616,793	658,395	534,624	570,198	608,663	1,113,777	1,188,243	1,268,394
Shift	611,686	687,985	772,377	495,636	487,027	475,480	1,116,969	1,194,339	1,278,133
Share	592,318	648,026	708,201	518,208	533,710	550,461	1,115,262	1,190,918	1,272,645
RANGE									
Absolute	66,794	99,513	141,522	92,863	166,710	250,727	129,962	204,996	274,928
Percent	11.85%	16.57%	22.24%	18.74%	34.23%	52.73%	11.75%	16.49%	22.37%

CONCLUSIONS:

The population projections seem to be varying according to the technique/method used. The complex extrapolation methods and Ratio extrapolation methods were giving closer estimates as compared to the simple extrapolation methods. Among ratio extrapolation method the projections obtained from share method were closely related with the Projected population as per Census 2011. The percentage range of population projections varies between 11.75% to 52.73 % among all statistical techniques of population projections discussed above.

References

1. Fildes, R. 1992. The evaluation of extrapolative forecasting methods. International Journal of Forecasting, 8: 81-98.
2. Marchetti, C., P. Meyer, and J. Ausubel. 1996. Human population dynamics revisited with the logistic model: How much can be modeled and predicted. Technological Forecasting and Social Change, 52, 1-30.
3. McNown, R., A. Rogers, and J. Little. 1995. Simplicity and complexity in extrapolative population forecasting models. Mathematical Population Studies 5: 235-257.
4. Pearl, R. and L. Reed. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proceedings of the National Academy of Science 6: 275-287.
5. Schmitt, R. and A. Crosetti. 1953. Short-cut Methods of Forecasting City Population. Journal of Marketing 17: 417-424.
6. Swanson, D. A., D. Beck., and J. Tayman. 1995. "On the Utility of Lagged Ratio Correlation as a Short-cut Population Projection Method: A Case Study of Washington State."Journal of Economic and Social Measurement 21: 1-16.