



# Estimating Futuristic trend in Brass’s $\alpha$ and $\beta$ for Male population for Uttar Pradesh, India by Bayesian method

Pandey Abhishek

Department of Statistics, Udai Pratap Autonomous College, Varanasi, Uttar Pradesh, INDIA

Available online at: [www.isca.in](http://www.isca.in), [www.isca.me](http://www.isca.me)

Received 21<sup>st</sup> July 2015, revised 27<sup>th</sup> September 2015, accepted 4<sup>th</sup> October 2015

## Abstract

*Aim of this paper is to estimate the futuristic trend of Brass’s  $\alpha$  and  $\beta$  by Bayesian procedure using WinBUGS. We have used MCMC (Monte Carlo Markov Chain) method to estimate these values using Non-informative priors Bayesian Procedure. In WinBUGS we check the suitability of linear regression model for  $\alpha$  and  $\beta$ .*

**Keywords:** Monte Carlo Markov chain simulation, Bayesian approach, Linear regression.

## Introduction

Life table system was initiated by John Graunt by publishing a rudimentary life –table in 1662 based on the analysis of “Bills of Mortality”. He considered mortality data for this purpose; therefore it was claimed defective by other scientists. A systematic abridged life was proposed by William Farr. Many attempts were made after his method. A new method was proposed by Chiang’s for person years lived. Life table suggests the life history of a cohort. Life–table can be computed for different age groups using past and present data provided by Sample Registration System in India. For projecting the population we use the age specific data available for past and present time. Reliability of projected estimates is depend upon valid assumption. Projection are the probabilistic forecast so there is no requirement to check its accuracy. In our paper we have used Bayesian approach for projecting the population of Uttar Pradesh. Bayesian procedure gives reliable estimates for the unknown values (parameters) so this method is gaining popularity day by day. BUGS is capable of computing difficult integrals.

Sex ratio suggests that there is difference in the male and female population in Uttar Pradesh, it requires to project the  $\alpha$  and  $\beta$  estimate for male population separately.

Brass’s method is different among all the available methods; it relates two  $l_x$  values of two life tables.

Brass formula is given as-

$$\log it(l - l_x^*) = \alpha + \beta \log it(1 - l_x)$$

$$\log it(l - l_x) = 0.5 \log \left( \frac{1 - l_x}{l_x} \right)$$

Sample Registration System has provided the data of Age specific death rate for years ranging from 1971 to 2010. From these data  $\alpha$  and  $\beta$  values were calculated. The estimated values of  $\alpha$  and  $\beta$  for male population are given in table-1.

Congdon P<sup>1</sup>., Dyson T. et al<sup>2</sup>, Gelman A et al<sup>3</sup>, Gilks et al<sup>4</sup>, Gill J<sup>5</sup>, altogether gives the ideology about the Bayesian Method and about Monte Carlo Markov Simulation that how to simulate and analyze the samples from our observed data, and forecast the result in a given confidence interval. Rahul et al<sup>6</sup>, Pandey G.S. et al<sup>7</sup>, Singh G.P. et al<sup>8</sup> suggested the method Projecting Population applied to states of India and India as a whole using Time series data using a suitable model and running a program in Win BUGS. Registrar General of India, 2006<sup>9</sup>, provides report on growth and nature of futuristic population of India and its states. Spiegel halter eta al<sup>10</sup>, teaches the methodology of running of Win BUGS software through various worked out examples.

**Objective:** Our aim is to estimate the Brass’s  $\alpha$  and  $\beta$  for male population in Uttar Pradesh using Bayesian approach by linear regression model, so that project male population.

## Methodology

**Bayesian approach:** Bayesian method provides new technique of analyzing the data. This method of analyzing data got enormous popularity in the various discipline. At first our attempt is to make a probabilistic model that is considered to explain properly the underlying mechanism of the system based on our past study and procedure of collecting samples. After that our aim is to formulate appropriate prior distributions unknown quantities of the model. Baye’s rule is applied after observing the past data to get the posterior distributions for these desirable parameters, which depends on the conditional probability distributions given the observed data. The rule may be expressed symbolically as follows

$$P(\theta / x) = \frac{P(\theta) \times P(x / \theta)}{P(x)} = \frac{\text{prior} \times \text{likelihood}}{\text{Marginal}} \quad (1)$$

$$= \frac{P(\theta) \times P(x / \theta)}{\int P(\theta) \times P(x / \theta) d\theta}$$

Here,  $\theta$  is the set of unobserved quantities of interest/parameters,  $P(\theta)$  is the prior distribution of  $\theta$ ,  $P(x/\theta)$  is the probability distribution of data  $x$  given prior distribution and information of  $\theta$  which is popularly called likelihood function of data  $x$ , and  $P(\theta/x)$  is called posterior distribution of parameters/unobserved quantities of interest  $\theta$ . As soon as we obtain posterior estimates of the parameter  $\theta$ , we can use this distribution to provide estimates of parameter  $\theta$ .

**Model:** Let  $A_{4[i,j]}$  is matrix represents value of the  $\alpha$  and  $\beta$  for the male population for years 1971 to 2010 in Uttar Pradesh in the year  $i$  ( $i = 1, 2, \dots, 49$ ) where  $i$  starts from 1971 and lasts up to 2051. Where it takes values 1971, 2000, 2001,.....2009, 2011, 2016, 2021, 2026, 2031, 2036, 2041, 2046, 2051. The data of  $\alpha$  and  $\beta$  is given in the Table 2. We use Bayesian methods to check the suitability of linear regression model in Win BUGS.

All the parameters are projected using this model in the Bayesian framework.

For estimating parameters a program was written in BUGS language and checked in BUGS. Program was implemented successfully. We gave Non-informative prior for " $\alpha$ " and " $\beta$ " and precision estimator " $\tau$ " is given prior Gamma (0.001, 0.001).

All these priors are non-informative providing limited information and we do not have information of specific nature of their probability distribution. A more rigorous discussion on the choice of non informative priors is available in Win BUGS manual by Spiegelhalter, Thomas, Best, and Gilks (1996)<sup>10</sup>.

**Tools:** Posterior distributions of Bayesian method involves complicated mathematical terms. Most of them can be handled by Monte Carlo Markov chain simulation method. The Markov Chain Monte Carlo (MCMC) method is a repetition procedure of generating samples from our distribution. We have used this method for handling the difficulties which arises due to typical mathematical terms that involves expected value of the function of a random variable. The calculation can be made much easier by generating large number of independent samples by simulation procedure from the (complex) distribution of the random variable. After that we take the mean of obtained values of the function from these sample points. Win BUGS (Bayesian inference Using Gibbs Sampling for Windows) is a freely available software that helps us to find out the estimates of unobserved quantities of ultimate interest by using MCMC process.

This procedure requires running a number of chains starting with one chain initially which can be increased up to three (default) or more for each parameters. It requires large number of iterations to reach to the stationary distribution. If we further update the model then it is supposed that the samples are drawn randomly from the posterior distribution of the parameters. In

WinBUGS there are number of inbuilt functional tools that checks the convergence of the chains. Generally one can use multiple diagnostics on a single chain. In WinBUGS we can run multiple chains simultaneously for each parameter. We have used some of the diagnostics available with the WinBUGS that is briefly described below. For convergence of MCMC simulations we run a number of chains in it. WinBUGS provides dynamic trace plot of the chains while updating the model. When we cannot see sufficient mixing of chains even after lots of updates, it indicates lack of convergence of the chains. The bgr-diagnostics calculates the modified form of Gelman-Rubin convergence statistic<sup>3</sup>.

Green running plots are of the statistic in which the width of the central 80% interval of the pooled, Blue running plots of the average width of the 80% intervals within the and the red plot shows their ratio  $R$  ( $=$  pooled / within) are provided by WinBUGS. Brooks and Gelman<sup>3</sup> told that we should be concerned with convergence of  $R$  to 1, and pooled and within interval widths should converge to got stability. In WinBUGS we can get smooth density plots of the chains. The density curve takes bell (normal) shape when the chains approach to stationary. The absence of convergence of the chains indicates lack of normality.

There is another diagnostic tool available inside BUGS namely Auto-correlation. When the chains converge to the stationary distribution then autocorrelation decreases with the increase in the lags. The basis to reach the convergence of the chain is also provided by it. A detailed discussion on the diagnostics can be found in Gill<sup>5</sup>.

When it seems that chains have converged, then this simulation procedure can be continued for a further number of iterations to obtain the samples that can be used for posterior inference. The accuracy of our posterior estimates will increase when we generate and include more samples in the iteration process. After running the adequate number of updates and got satisfied by of history of chains, we can exclude the previous samples. Summary statistics can only be obtained from the further generated samples.

**Analysis:** Table 1 given below is estimate of  $\alpha$  and  $\beta$  for male population in Uttar Pradesh calculated by Brass's method taking 2010 as standard year. It includes the data continuous 40 years. SRS reported ASDR data starting from 1971 to 2010. Table-2 given below is the represents the estimated value of Brass's  $\alpha$  and  $\beta$  for males. Table-3 represents other parameters.

Figure-1 shows that Kernel density for all estimates is bell shaped. It shows that  $\alpha$  and  $\beta$  will be distributed normally in the future for male population. Figure-2 and 3 shows the mixing of chains. Figure-4 shows that autocorrelation will decrease with lag for all parameters.

**Table-1**

Year	$\alpha$	$\beta$
1971	0.370957	0.690075
1972	0.515134	0.736766
1973	0.330493	0.681715
1974	0.368228	0.723634
1975	0.455025	0.704739
1976	0.396507	0.762424
1977	0.355666	0.745026
1978	0.346002	0.714534
1979	0.275657	0.701429
1980	0.347206	0.806337
1981	0.290771	0.762001
1982	0.215059	0.739787
1983	0.271446	0.798313
1984	0.348837	0.773117
1985	0.307693	0.821377
1986	0.274234	0.776078
1987	0.260482	0.807326
1988	0.184326	0.782755
1989	0.179857	0.851728
1990	0.142863	0.875311
1991	0.126801	0.839048
1992	0.159368	0.864472
1993	0.123743	0.897395
1994	0.112021	0.862953
1995	0.051927	0.864675
1996	0.055365	0.870823
1997	0.082885	0.899909
1998	0.079237	0.931553
1999	0.088522	0.90038
2000	0.048743	0.914568
2001	0.044064	0.894527
2002	0.019015	0.903003
2003	0.057795	0.990583
2004	-0.01718	0.889945
2005	0.003523	0.910083
2006	0.005944	0.912052
2007	0.061871	1.028071
2008	0.037189	1.009766
2009	0.00847	1.009173
2010	0	1
2011	NA	NA
2016	NA	NA

Year	$\alpha$	$\beta$
2021	NA	NA
2026	NA	NA
2031	NA	NA
2036	NA	NA
2041	NA	NA
2046	NA	NA
2051	NA	NA

**Table-2**  
 Projected values of  $\alpha$ ,  $\beta$  and other parameters

Year	Parameters	2.50%	Mean	97.50%
2011	$\alpha[1]$	-0.09249	-0.06071	-0.02935
2016	$\alpha[2]$	-0.1583	-0.1205	-0.08317
2021	$\alpha[3]$	-0.1715	-0.1325	-0.0939
2026	$\alpha[4]$	-0.1846	-0.1445	-0.1047
2031	$\alpha[5]$	-0.1979	-0.1564	-0.1154
2036	$\alpha[6]$	-0.2112	-0.1684	-0.1262
2041	$\alpha[7]$	-0.2244	-0.1804	-0.1369
2046	$\alpha[8]$	-0.2376	-0.1924	-0.1475
2051	$\alpha[9]$	-0.2509	-0.2043	-0.1582
2011	$\beta[1]$	0.9809	1.001	1.022
2016	$\beta[2]$	1.016	1.04	1.065
2021	$\beta[3]$	1.023	1.048	1.073
2026	$\beta[4]$	1.03	1.056	1.082
2031	$\beta[5]$	1.037	1.064	1.091
2036	$\beta[6]$	1.044	1.071	1.099
2041	$\beta[7]$	1.051	1.079	1.108
2046	$\beta[8]$	1.058	1.087	1.116
2051	$\beta[9]$	1.065	1.095	1.125
	R <sup>2</sup> for $\alpha$	0.8507	0.9049	0.9398
	R <sup>2</sup> for $\beta$	0.8483	0.9037	0.9392
	phi[1,1]	0.8027	0.882	0.9619
	phi[1,2]	-4.865	-4.371	-3.882
	phi[2,1]	0.3341	0.3863	0.438
	phi[2,2]	2.535	2.851	3.168
	Sigma.C[1]	0.03938	0.04915	0.06184
	Sigma.C[2]	0.02577	0.03223	0.0406

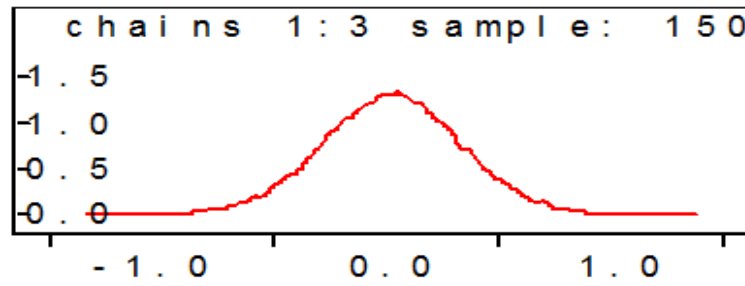


Figure-1  
Karnel density for all estimates

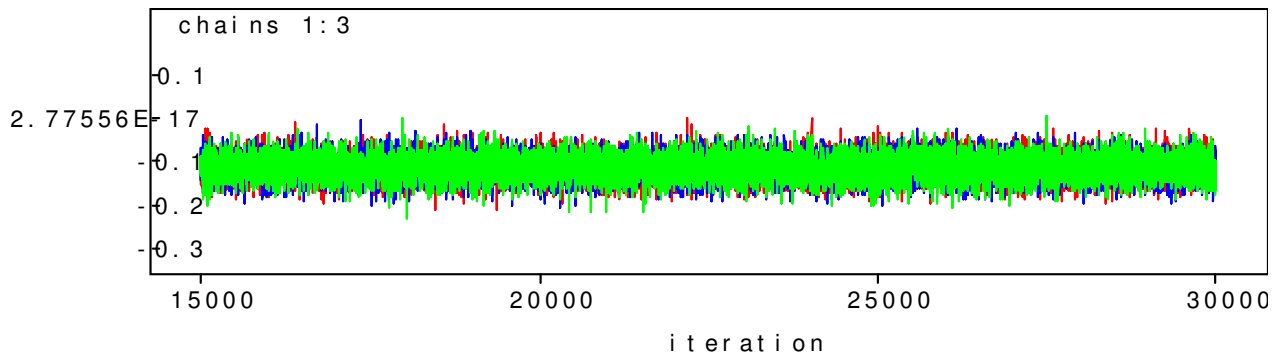


Figure-2  
Time series plot for checking convergence

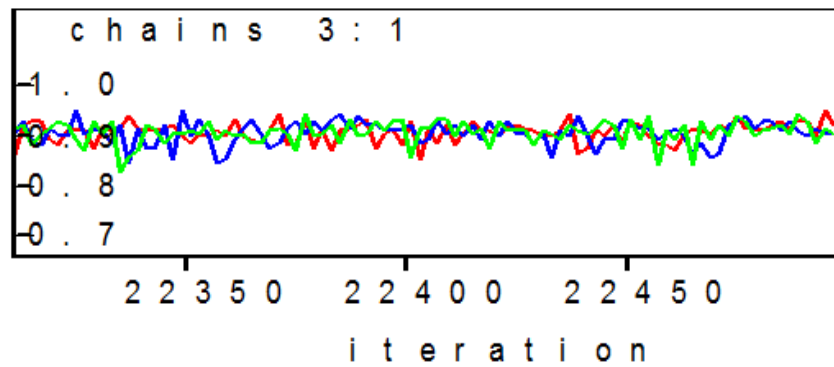


Figure-3  
Dynamic Trace for parameters

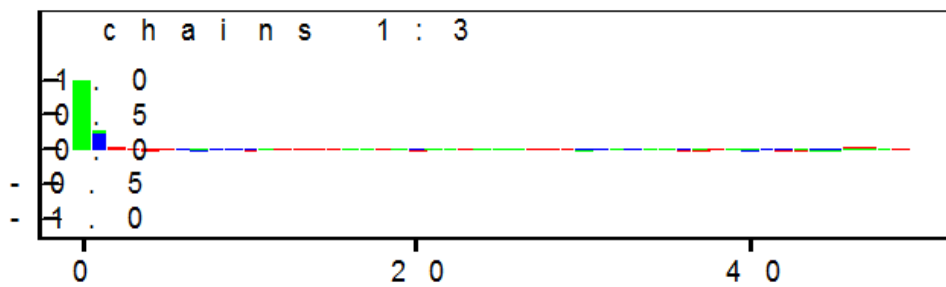


Figure-4  
Auto correlation Plot for all estimates

## Conclusion

Bayesian approach was implemented to check the suitability of the linear regression model. The estimated values of the parameters of proposed model are shown in table-2. The table shows interval estimates (95% Highest Posterior Density) for all the parameters of the model for different years. Values of Bayesian  $R^2$  suggest that our model is fitted good for linear regression model. These estimated values will be used to project future male population.

## References:

- 1 Congdon P., Bayesian Statistical Modeling, John Wiley and Sons, Chichester, U.K., (2001)
- 2 Dyson T., India's Population-The Future in T. Dyson R. Cassen and L. Visaria, Eds., Twenty-First Century India: Population, Economy, Human Development, and the Environment. Press, Oxford, 74-10., (2004)
- 3 Gelman A., Carlin J.B., Stern H.S. and Rubin D.B., Bayesian Data Analysis, Second Edition, Chapman and Hall/CRC, (2003)
- 4 Gilks W.R. Richardson S. and Spiegelhalter D.J., Eds., Markov chain Monte Carlo in Practice, Chapman and Hall, London, U.K., (1996)
- 5 Gill J., Bayesian Methods: A Social and Behavioral Sciences Approach, Chapman and Hall/CRC., (2002)
- 6 Rahul and Singh O.P., Population Projection Of Indian Using MCMC Technique in Bayesian Theory (2011-2051), published by LAP LAMBERT Academic Publishing, Germany, (2012)
- 7 Rahul Pandey G.S. and Singh O.P., Population Projection Of India Using MCMC Technique: A WinBUGS Implementation, Demography India, New Delhi, India, 36(1), 145-154, (2007)
- 8 Rahul Singh G.P. and Singh O.P., Population Projection of Kerala Using Bayesian Methodology, *Asian Journal of Applied Sciences*, 2(4), 402-413, (2009)
- 9 Registrar General Of India, Population Projection for India and States 2001-2026, Office of the Registrar General and Census Commissioner, India, (2006)
- 10 Spiegelhalter, D.J. Thomas, A. Best, N.G. and Gilks W.R., BUGS: Bayesian inference Using Gibbs Sampling, Version 0.5. MRC Biostatistics Unit, Cambridge, (1996)