CHAPTER II

A COMPREHENSIVE REVIEW OF LITERATURE
2.1 INTRODUCTION

After the publication of Bayes' essay in 1763, many statisticians and scientists including Laplace, Edgeworth, Jeffreys, De finetti, Box, Savage, Lindley, Raiffa, Jaynes, Schlaifer, Pratt, Good, Geisser, De Groot, Tiao, and others have made important contributions to the developments of Bayesian Statistics. However, the modern Bayesian econometrics commenced in the early 1960's with papers by Jacques Dreze, Tom Rothenberg, Walter Fisher, Albert Ando, Gordon Kaufman, Zellner and a few others.

The importance of Bayesian approach lies in the fact that in econometrics and in other areas of science a good deal of prior information is employed, usually informally. The Bayesian approach provides formal methods for handling background or prior information. Prior information is the basis of Bayesian approach. Therefore, good procedures for incorporating prior information in analysis are needed. Fortunately, the Bayesian approach is one in which prior information, as little or as much as an investigator desires, can be flexibly and formally incorporated in estimation, prediction, testing, model selection, and policy analysis procedures.

The brief review in the following paragraphs reveals that many, if not all, non-Bayesian results can be produced by Bayesian methods under special assumptions. For example, with a large sample assumption, a posterior mean that is optimal relative quadratic loss is approximately equal to the maximum likelihood estimate. In large samples, according to Zellner, such approximate results may be satisfactory; however, in small to moderate sized samples these conditional posterior pdfs are poor approximations to the relevant marginal posterior pdfs and thus such approximations lead to poor estimates and other inferences.
Bayesian methods have been used successfully in a number of applied studies. Litterman has used Bayesian vector autoregressions in forecasting seven important U.S. quarterly macro economic variables including real GNP, the implicit deflator for GNP, the unemployment rate etc. In his work, Litterman formulated informative prior pdf for many parameters of his seven variable VAR which effectively reduces the number of free parameters. His forecasts compare very favorably with those obtained from unrestricted VARs, univariate AR models, Box-Jenkins’ models, and two large structural economic models. Akaike and Ishiguro and their colleagues have developed a Bayesian approach and computer program, BAYSEA, for seasonal adjustment procedures which indicate that the former results compare favorably with other results. Morris has discussed and referenced many applications of empirical Bayes’ procedures. Prescott, Bowman and Loparte, Harkema, and Zellner have compared Bayesian and non-Bayesian solutions to the undersized sample problem with the finding that the former are as good or better than non-Bayesian solutions. Miller reviews Bayesian applications in actuarial statistics.

Bayesian Approach has been widely used in Econometrics and Demography. Some of the research papers published on this topic are reviewed briefly in the following paragraphs:

2.2 BAYESIAN APPROACH - ECONOMETRICS

Cox (1958) discussed a model on the regression in binomial populations in detail. This paper follows exactly Cox’s approach, a classical regression analysis for the case of Poisson population. Using the Cox model it is assumed that the means are changing exponentially, that is, \( \log \theta_i = \alpha + \beta x_i \), and the exact Bayesian distribution for \( \beta \) is derived, a Bayesian approximation suggested
which proved to be very useful. Finally the author Sayyed (1961) showed the comparison of three methods i.e., classical approach, Bayesian method (the exact distribution) and Bayesian approximation.

The problems of assessment of prior distributions in Bayesian analysis was discussed by Winkler (1967). The results proposed by Winkler are of fundamental importance in Bayesian estimation.

Leonard (1972) obtained a Bayesian procedure for the simultaneous estimation of the parameters of m binomial distributions. He developed a method which uses logistic transformations for the parameters and an exchangeable prior distribution. Leonard (1975) also proposed estimation methods for the row, column and iterations effects in two-way contingency tables, the one-way table being treated as a special case. The methods were shown to be appropriate when the parameters are related 'a priori' to each other.

Lindley and Smith (1976) in their paper: 'Bayes estimates for the linear model' used Bayesian methods and the concept of exchangeability. The general methods are illustrated by applications to two-factor experimental designs and multiple regression. In this paper attention is confined to the linear model with normal errors. The authors develop thoroughly Bayesian approach.

Guthrie (1976) developed the Bayesian estimation of the parameters of Solow’s distributed lag model with implicit autocorrelation of disturbances in its autoregressive form. The estimation technique extends Chetry’s method for independent disturbances. The results of some Monte Carlo experiments were given comparing point estimates from the posterior distributions with the maximum likelihood estimators.
Leonard (1978) suggested an empirical Bayes method for smoothing two-way tables, based on the use of the log linear model and a normal prior distribution. Estimation of the variance component in the prior was also discussed by considering two approximations, one of which utilizes the Expectation Mathematical (EM) algorithm.

A Bayesian approach was given for various kinds of empirical Bayes problems by Deeley and Lindley (1981). Several examples of these concepts were also given besides a general theory showing the difference between an empirical Bayes model and a Bayes empirical Bayes model.

Estimation techniques for linear covariance components models were developed and illustrated with special emphasis on explaining computational processes by Rubin and Tsutakawa (1981). The estimation of fixed and random effects when the variances and covariances are known was presented in Bayesian form. Point estimates of the unknown variances and covariances were computed using the Expectation Mathematical (EM) algorithm for maximum likelihood estimation from incomplete data.

Leonard (1982) showed the interaction effects in two way contingency tables, when the parameters are thought a priori to be related to each other (row and column). The posterior estimates have the practical effects of smoothing the contingency table. The main case treated is where particular assumptions of exchangeability are reasonable a priori for the unknown parameters. Fienberg (1970) and Holland (1973) make the assumption of exchangeability of all the cell probabilities corresponding to cell in the same row and column will often be thought a priori to be more closely related than probabilities corresponding to cells
not in the same row and column, thus destroying the symmetry implicit in the exchangeability assumption. Sutherland et al. (1975) essentially assume exchangeability of the interaction effects with fixed marginal effects. Their prior distributions are however, data dependent and as such cannot really be recommended under a Bayesian approach.

A state-space model was developed by Dejong and Boyle (1983) which provides estimates of decrements in a dynamic environment. The model integrates the actual unfolding experience and a priori or Bayesian views of the mortality rates. The model was described and applied in the context of mortality estimation, which proved to be useful in actuarial applications.

Smith (1984) reviewed basic ideas and methods of Bayesian statistics both in foundational terms and in relation to the world of statistical practice. The limits and slope of the Bayesian approach were assumed, and tentative conclusions were drawn regarding the current situation and future prospects.

Tsutakawa (1985) presented a Bayesian method for estimating mortality rates of specific diseases when the frequency of deaths over a specified time period is assumed to have a Poisson distribution with mean proportional to the population size. The study of Tsutakawa was motivated by an epidemiological study on the geographic variation of cancer mortality in the state of Missouri. Most cancer types affect different age sex groups quite differently. Many epidemiologists like Mason and McKay (1974), Hill (1971), Breslow and Day (1975) find adjusted rates to be of limited value and use a multiplicative model on the raw rates. Finally, the author showed the Bayes estimate using the lung cancer data. He also used the Gauss-Hermite quadrature method for practical work.
Tsutakawa (1985) illustrates computational methods required for estimation in covariance components models. He presented Bayesian theory for the joint estimation of fixed and random effects when the variances and covariances are known. The posterior means and posterior variances and covariances thus defined are used for finding maximum likelihood estimates of variances and covariance components by means of Expectation Mathematical (EM) algorithm.

Ishak (1992) introduced an improved statistical technique that can sufficiently utilize all the available data i.e., both the national published data and the data obtained from sample surveys, censuses or any other sources to estimate infant mortality experience in the country. A Bayesian approach for estimating infant mortality in developing countries is applied. The applicability and reliability of the model is examined by applying it to the mortality data of Egypt.

Death rates of particular categories in epidemiological studies are often based on a small number of occurrences which can be well described by a Poisson distribution. This model was applied by Jordan et. al. (1997) for the analysis of multi centre study in five Japanese countries where the death rates of stomach cancer in four age groups were known. The model was estimated in a Bayesian framework by means of resampling techniques.

Bayesian learning model to assess the respective influence of different risk measurements on mortality risk perceptions was studied by Hakes and Viscusi (1997). According to the authors, risk-related variables were much better predictors of larger risks than of small risks, which rejected the role of information costs and benefits of learning about large risks.
Gandotra and Das (2001) analysed infant mortality rate as summary index of the socio-economic development of a region. The study also obtained data on census of infant deaths by using prospective data.

The effect of important biological and social factors on under five survival in Malawai was analysed by Bolstad and Manda (2001). The infant and early childhood survival model was suggested using family and community random effect multipliers on the fixed effect proportional hazards model, which allows the dependence between observations in the same family and community into the model.

Geweke et al. (2002) developed new econometric methods to infer hospital quality in a model with discrete random variables and non-random selection. Bayesian inference in these methods was made feasible using a Markov Chain Monte Carlo (MCMC) posterior simulator, and attached posterior probabilities to quality comparisons between individual hospitals and groups of hospitals.

Leyland and Davies (2002) reviewed empirical Bayes methods for disease mapping. A distinction was made between spatial models and non-spatial models. Several estimations were suggested and empirical Bayes methods were compared with full Bayes methods.

Empirical Bayes estimation of small area adult uncertainty risk was studied by Fergnson et al. (2004). They used Bayes theorem to generate smoothed age and sex specific probabilities of death by district between the ages of 15 and 59.

Inference about portfolio mortality trends was focused by Olivieri and Pitsacco (2006). A Bayesian inferential model was proposed, aiming at mortality adjustments based on prior information and statistical evidence. The authors cited some numerical examples and illustrated the inferential mechanism.
2.3 DEMOGRAPHY AND FORECASTING

Demographic analysis and forecasting has been attracting importance in view of the rapid growth of population and its impact on socio-economic conditions in the contemporary society. This research work has been carried out to analyse and forecast the demographic components such as mortality rates, birth rates, migration and urbanization. An attempt is made in the following paragraphs to review briefly some of the important research contributions in demographic analysis.


A simplest method of forecasting in to extrapolate life expectancy (a zero factor model), or some other life table measure, and to use empirically- based model life tables to obtain the age pattern. Coale & Guo (1989) have studied the case of older age and lower mortality. The independent extrapolation of age-specific rates commonly involves mortality reduction factors or some fraction of the reduction factor according to researches like Pollard (1987).

Parameterization (one-factor) functions of mortality were discussed by several authors including Keyfitz (1982) and Poppel (2002). Among the best-fitting are the three-term functions capturing the age pattern of mortality in childhood, young adulthood and senescence. Heligman and Pollard (1980) have considered eight parameters; each term takes positive values only at relevant ages, the whole function being estimated in one step. McNown and Rogers (1989) modeled the eight parameters as univariate ARIMA processes.
Carriere (1992) suggested a flexible model with potential advantages in forecasting. His three-term model with eight interpretable parameters gave a better fit to US data than the Heligman-Pollak model—a four-term version with eleven parameters gave significant improvements.

Rogers and Little (1995) found interdependencies among parameters; five were fixed, and four were modeled by three univariate ARIMA processes and a bivariate autoregression.

Congdon (1993) adopted the Zaba model and forecast the relatively stable parameters by univariate ARIMA models. Hannerz (2001) combined the features of relational models with parameterization functions and model life tables in a regression model.

The most well-known method of mortality forecasting is the Lee-Carter method (1992) for long term forecasting. The underlying two-factor model describes a one-parameter system with fixed age effects and has a homoscedastic additive Gaussian error structure. A unique least squares solution is found by singular value decomposition (SVD). Lee and Carter incorporated an adjustment of the level parameter so that fitted deaths match observed total deaths in any year; this avoids discrepancies arising from modeling on the logarithmic scale (Lee, 1992). The adjusted level parameter is modeled by time series methods; in almost all applications, a random walk with drift has been found to be applicable. Girosi and King (2006) reviewed the method in detail.

Recent developments extend the applicability of the Lee-Carter method. Li, Lee and Tuljapurkar (2004) demonstrate how, by assuming a linear trend in the level parameter, the method can be applied to populations with limited data at unequal time intervals.

It may be noted that the Lee-Carter method has close similarities to the principal components approach used by Bell and Monsell (1991). Hyndman and Ullah (2004) extend the principal components approach by adopting a functional data paradigm combined with nonparametric smoothing and robust statistics, and fit univariate time series models to each component coefficient (or level parameters); the Lee-Carter method is shown to be a special case of this generalized approach.


Giroai and King (2006) address the problem of life expectancy in the context of cause of death forecasting by incorporating prior information as covariates in a linear age-period regression model. The method suggested by them improves on existing Bayesian approach by borrowing strength based on expected values rather than on coefficients, and maximizes use of prior information hence maximizing automation.
The research contributions listed in the foregoing paragraphs indicate that in demographic forecasting, various methods of forecasting have been discussed along with their limitations with reference to the availability of data on different components.

Lee-Carter introduced a method for forecasting the age specific log mortality rates for the entire U.S population. Their method consist of fitting the following model to a matrix of log mortality rates indexed by age group a, and year ‘t’. 

\[ \ln(m_{a,t}) = \alpha_a + \beta_a K_t + \epsilon_{a,t}. \]

where \( \alpha_a \)'s are the averages over time of the \( \ln(m_{a,t}) \); \( \beta_a \)'s represent the rates of change of the age-specific mortality rate relative to changes in \( K_t \); and \( K_t \) represents an index of the level of mortality for year ‘t’.

Lee-Carter’s approach consists of three main steps:

(i) Estimating \( \alpha_a, \beta_a \) and \( K_t \).

(ii) Re-estimating the mortality index \( K_t \), assuring that \( \alpha_a \) and \( \beta_a \) are known; and

(iii) Modeling and forecasting \( \{ K_t : t=1,2, \ldots \} \).

From these forecasted values the mortality rates \( m_{a,t} \) for future years are calculated.
King and Pedroza (2002) used the Lee-Carter method and its reformulation as a state-space model. They compared it to a simple random walk model and showed that the forecasts from the two models are very similar. They also present a Bayesian model using Gibbs sampler method, besides Bayesian forecasts.

The stability of the Lee-Carter method to structural change and initial conditions was examined by Carter (2000) and Carter and Prskawet (2001). Tuljapurkar (2004, 2005) demonstrated the robustness of the method. Li and Chan (2005) proposed an outlier adjusted method. Lee – Miller (2001) studied the influence of the adjustment procedure on forecast bias and the variant. Their suggestions have been widely adopted as the standard Lee-Carter method. Girosi and King (2006) have reviewed the Lee-Carter method in detail.