Applied Mathematics 2013, 3(3): 98-106
DOI: 10.5923/j.am.20130303.03

On the Estimation of Average HIV Population Using Various Bayesian Techniques

Gurprit Grover, Ravi Vajala*, Manoj Kumar Varshney

Department of Statistics, University of Delhi, Delhi, 110007, India

Abstract This paper models the HIV population through a Poisson distribution and obtains the expressions for the estimators of the average number of HIV individuals (Incidence Rate of HIV). Conventional methods for obtaining such estimates have used the Maximum likelihood Principle that does not take into account, any prior information about the parameter. Bayesian perspective accommodates this missing link and hence obtains the estimators where data is refined using the prior information. Three different types of prior distributions including Jeffreys non-informative priors have been considered and the corresponding estimates along with standard errors have been obtained assuming a squared error loss function. However, computational techniques like Markov Chain Monte Carlo (MCMC) have been avoided by using the Empirical Bayes Perspective. These procedures were applied on the state and year-wise data of HIV patients in India and relevant estimates are obtained and compared with actual figures. When year is considered as random variable, M.L.E proved to be better than the Bayes estimates but vice-versa is seen when states were considered as a random variable.

Keywords Poisson Model, Human Immuno Deficiency Virus (HIV), Informative and Non-Informative Priors, Bayes Estimators, Empirical Bayes Methods

1. Introduction

The Kolmogrov equations for the various Birth and Death processes yield the Poisson distribution as the distribution of number of infectives at time t. This may be thought of as an intuitive result considering the fact that when we are building a model for the HIV infectives in the population, the area of opportunity is very large and the opportunity of infection is very small, so that both of them multiply to a finite quantity. This finite quantity is the average number of HIV cases in the population at time t or the HIV incidence rate per time period and may be considered as a time dependent or independent constant. The scenario may be suitably modeled through a Poisson distribution as follows:

Let X denote the number of HIV infected individuals in the population. Therefore,

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, x = 0, 1, \ldots \text{and } \lambda > 0 \qquad (1)$$

where the parameter λ is (assumed as time independent) the average number of HIV cases in the population.

Poisson distribution is widely used by many researchers in modeling the HIV data ([5, 19, 20], etc.) Considerable amount of work has been done in the estimation of the parameter λ using various techniques and such an estimate is used for further modeling of the Viral Load, estimating the time since infection **[7], obtaining incidence rates [14],** etc.

Deuchert et. al.[6] used Poisson distribution while modeling HIV/AIDS epidemic by assuming that heterosexual transmission is the major or sole transmission mode. The empirical estimates for relevant model parameters were obtained by the Maximum Likelihood approach and were compared with parameters in mathematical models using the Chi-square goodness of fit test, Akaike Information criterion, etc.

However, the usual method to estimate λ is the Maximum likelihood approach where the estimate is obtained as $\hat{\lambda} = \bar{x}$ or a weighted mean of observations for a sample $x_1, \ldots, x_n$ of n observations. Moreover, despite satisfying the properties of a good estimator asymptotically, the Maximum likelihood estimator (M.L.E) fails to take into account any additional information available on the parameter λ prior to taking the sample. This additional information may be incorporated into the estimation process by the so-called prior distribution and hence Bayesian approach may be used to evolve a much more refined estimator. Bayesian methods do not require large samples or asymptotics for their validity. They allow for incorporation of expert knowledge through the specification of prior distribution.

Classical M.L.E approach to the problem of estimation relies on an estimator which is obtained theoretically and remains the same for whatever may be the data set. However, Bayesian approach obtains a separate set of estimators for

* Corresponding author:
vravi.vr@gmail.com (Ravi Vajala)
Published online at http://journal.sapub.org/am

every set of prior information and adjusts these estimators for changes in the data set. Such an estimator provides a logical alternative because it not only incorporates the additional information on the parameter, but also relies on the data to a great extent.

2. Bayes Approach for Estimating HIV Incidence Using Various Prior Distributions

The use of prior distributions is the best way to summarize the available information (or rather lack of information) about the parameter of interest i.e., average number of HIV persons in the population. It may be helpful in incorporating the experiences from previous studies or subjective beliefs of the experimenter into the analysis. These beliefs may be put into various kinds of functional forms depending on the amount of information available.

Let x_i denote the number of HIV infected individuals in the population for the i^{th} entity/time point, with probability $P(x_i \mid \lambda)$ where λ is the parameter denoting the average number of HIV infected individuals in the population. Let the *prior* probability (or "*unconditional*" or "*marginal*" probability) of λ be $P(\lambda)$ and the joint distribution of $x_1, x_2, \ldots, x_n$ be $P(\underset{\sim}{x} \mid \lambda)$. Then the posterior density of λ is given by

$$P(\lambda \mid \underset{\sim}{x}) = \frac{P\left(\underset{\sim}{x} \mid \lambda\right) P(\lambda)}{P\left(\underset{\sim}{x}\right)} = \frac{P\left(\underset{\sim}{x} \mid \lambda\right) P(\lambda)}{\int P\left(\underset{\sim}{x} \mid \lambda\right) P(\lambda) d\lambda} \quad (2)$$

provided that the probability of $\underset{\sim}{x}$ does not equal zero.

When *substantial information* about the average HIV cases is available, we may look at the (Natural) Conjugate Priors wherein the functional form of the prior and posterior remains same and when *no information* is available, we may consider the Non-informative Prior (Jeffreys)[10].

The subsequent sections develop theory for modeling the HIV incidence λ, using various prior distributions in the population. This prior information is refined to posterior distribution by means of additional information provided by the data and estimates of λ are obtained from the posterior distribution. The estimates are obtained so as to provide minimum risk (which is expected loss) with respect to the posterior distribution. Of course, there is no consensus opinion on defining the loss, although the Quadratic loss is popularly used and found to be sufficient in majority of the situations.

2.1. Conjugate Prior for Modeling HIV Incidence

Cole et. al[18] and Berry et al.[3] had used the Gamma distribution as a prior for the incidence of HIV infection and multiplied it with a pseudo-person-time to find the average number of recent HIV infections. The estimates were obtained using the Markov Chain Monte Carlo (MCMC) procedures. Kpozhouen et al.[2] attempted to test the Bayesian approach as a tool for optimizing management of a chemoprophylaxis trial in HIV infection by allowing interim analysis with a reduced number of patients or follow-up time. The Bayesian proportional hazards model was considered for this purpose. The unknown coefficients of the covariates and the baseline cumulative incidence were assigned three different kinds of prior distributions. Gamma prior was assumed for the baseline cumulative incidence since the variable under consideration followed Poisson distribution and also to facilitate the determination of conjugated distribution. Finally, posterior distribution and the estimates of the parameters (as well as hyperparameters) were obtained using Markov Chain Monte-Carlo (MCMC) methods. White et al.[16] developed an age-stratified model that accounts for transmission due to unsafe injections, unsafe transfusions, and mother-to child transmission. The confidence intervals for HIV incidence rates with respect to unsafe injections were based on the Poisson assumption with a Gamma prior. The estimates of relative contribution of HIV-contaminated injections, and other routes of HIV transmission in this age-structured transmission model were obtained using the MCMC techniques. Grover et al.[9] used the Bayesian approach for estimating the proportion of HIV infected population converting to AIDS. They had modeled the number of HIV infected persons using Binomial distribution and assumed the beta distribution as a conjugated prior for the proportion of HIV infected population converting to AIDS. Finally Maximum likelihood estimates were for the posterior distribution.

Even though the widespread use of Computers in analysis have highlighted the importance and popularized the use of MCMC techniques, the procedure itself seems to be complex. Moreover, at times MCMC procedures involve large number of iterations and still fail to converge to any particular value. We present an easy method of obtaining the estimates of the parameters by using the Empirical Bayes approach[17] as this will save the time and complexities of an MCMC technique. The aim of suggesting the Empirical Bayes approach is not to portray the MCMC procedures in poor light, but is solely intended to provide an easy and convenient alternative. Bartolucci et al.[1] had used the Empirical Bayes analysis for estimation of incidence and intervention parameters for the Intervened Poisson (IP) model.

Let us assume that the prior distribution for HIV incidence rate λ follows a Gamma distribution with parameters (α, β). On using the Bayes theorem, for a given set of data $x_1, \ldots, x_n$, the posterior distribution of $\lambda \mid x_1, \ldots, x_n$ becomes Gamma $(\alpha + \sum x_i, \beta + n)$. The posterior mean of the distribution is $\frac{\alpha + \sum x_i}{\beta + n}$ which provides an estimate of λ with variance $\frac{\alpha + \sum x_i}{(\beta + n)^2}$.

The problem in finding the estimator of intensity λ is that it is based on the data as well as the parameters of the prior

distribution α and β (known as *hyperparameters*). Authors[16, 18] of related studies have, on the basis of their past knowledge, judgment or intuition taken various predetermined values for these hyperparameters and obtained the estimate of the average HIV cases. These estimators were further studied for robustness with respect to the prior parameters. However, we believe that the information about the parameters of interest lies in the data itself and hence the hyperparameters have been estimated using the Empirical Bayesian Procedure[17].

Let $\mu_f(\lambda)$ and $\sigma_f^2(\lambda)$ denote the conditional mean and variance of the random variable X which denotes the HIV cases in the population. Let μ_m and σ_m^2 denote the marginal mean and variance of these HIV cases. Assuming that these quantities exist, we have

$$\mu_m = E^{\pi(\lambda)}\left[\mu_f(\lambda)\right] \tag{3}$$

and

$$\sigma_m^2 = E^{\pi(\lambda)}\left[\sigma_f^2(\lambda)\right] + E^{\pi(\lambda)}\left[\mu_f(\lambda) - \mu_m\right]^2 \tag{4}$$

Further, if $\mu_f(\lambda) = \lambda$ and $\sigma_f^2(\lambda) = \sigma_f^2$ then,

$$\mu_m = E^{\pi(\lambda)}\left[\lambda\right] \text{ and } \sigma_m^2 = \sigma_\pi^2 + \sigma_f^2$$

Therefore the estimates of the hyperparameters when the prior distribution of λ is Gamma(α, β) are obtained as $\hat{\alpha} = \frac{\bar{x}^2}{s^2 - \bar{x}}$ and $\hat{\beta} = \frac{\bar{x}}{s^2 - \bar{x}}$ where $\bar{x}$ and s^2 are the sample mean and variance respectively. These may in turn be used to find the estimate of HIV incidence rate along with its standard error.

2.2. Non-Informative Priors for Modelling HIV Incidence

Kpozhouen et al.[2] used three different kind of prior distributions, one being a non-informative prior to assess the efficacy of Cotrimoxazole prophylaxis in reducing severe morbidity in adults at early stages of human immunodeficiency virus infection. The authors modeled the intensity of serious events (mortality) using the Bayesian proportional hazards model proposed by Spiegelhalter et al. (1996) and used the non-informative prior to represent the weak prior information on the coefficients of the covariates of the model. Many other authors have used the non-informative priors in terms of assigning negligible values to the parameters of the conjugate/informative prior distributions. However, none of them have modeled weak prior information in the lines of Jeffreys perspective which is a formal and admissible approach to specifying negligible information on the parameter of interest.

When prior information about the HIV incidence parameter λ is not available and the intention is to use the available clinical data to determine the parameter, Harold Jeffreys[10] approach may be used to obtain the following non-subjective reference prior in terms of the Fisher's Information matrix:

$$\pi(\lambda) \propto \sqrt{\left|\mathbf{I}\left(\lambda \mid \underset{\sim}{x}\right)\right|} \tag{5}$$

where

$$\mathbf{I}\left(\lambda \mid \underset{\sim}{x}\right) = E\left[\frac{\partial \log l\left(\lambda \mid \underset{\sim}{x}\right)}{\partial\lambda}\right]^2 = -E\left[\frac{\partial^2 \log l\left(\lambda \mid \underset{\sim}{x}\right)}{\partial\lambda_i \partial\lambda_j}\right]$$ is the Fisher's Information matrix.

Therefore, the prior distribution for the HIV incidence rate (λ) according to Jeffrey's rule may be taken as $\pi(\lambda) \propto \sqrt{\frac{1}{\lambda}}$. Using the Bayes rule, the posterior distribution is obtained as ($\lambda|x_1, \ldots, x_n$) ~ Gamma($\sum x_i + \frac{1}{2}, n$). The posterior mean of the distribution is $\frac{\sum x_i + 1/2}{n}$ which provides an estimate of λ with variance $\frac{\sum x_i + 1/2}{n^2}$.

The situation of no prior information about the Incidence rate λ, may be also modeled through the improper prior, $\pi(\lambda) = 1$ where $0 \le \lambda < \infty$ for which the posterior distribution is given by ($\lambda|x_1, \ldots, x_n$) ~ Gamma ($\sum x_i + 1, n$). The estimate of the HIV incidence rate is then $\frac{\sum x_i + 1}{n}$ with variance $\frac{\sum x_i + 1}{n^2}$.

Also, in case of large sample sizes, Brenner et al.[4], Fraser and McDunnough[8] gives the asymptotic posterior distribution for the HIV incidence parameter λ on assuming non informative prior.

If $\hat{\lambda}$ is the M.L.E of λ and prior density of λ is non-informative (or likelihood dominates the prior density), then the posterior density of λ is given by $\lambda \mid x_1, \ldots, x_n \sim$ Normal $\left(\hat{\lambda}, \frac{-1}{L''(\hat{\lambda})}\right)$ where L(λ) is the logarithm of likelihood of ($\lambda \mid x$).

Using this result, we obtain the posterior distribution of the HIV incidence as

$$(\lambda \mid x_1, \ldots, x_n) \sim \text{Normal}\left(\frac{\sum x_i}{n}, \frac{\sum x_i}{n^2}\right) \tag{6}$$

The estimate of HIV incidence rate, λ is the posterior mean i.e., $\bar{x} = \frac{\sum x_i}{n}$ with variance $\frac{\sum x_i}{n^2}$ which is the Maximum Likelihood Estimator of λ.

The objective of using these non-informative priors is to

highlight the '*weak*' or '*negligible*' information over '*no*' prior information as assumed while obtaining M.L.E. Also, this paper assumes an informative prior modeled by Gamma distribution in comparison with *weak* (Jeffreys), *negligible* (Improper) and *no* (M.L.E) prior information. The consideration for these non-informative priors is to bridge the gap between *substantial* amount of information as compared to *no* information. Various other informative priors could also have been considered for this purpose but then, these could also be obtained by giving suitable values to the parameters of Gamma distribution being one from the exponential class of distributions. In such situations, the solution would simply be obtained by doing a grid search for the best values of hyperparameters that will yield minimum standard errors of the estimators. However, doing so would dilute the concept of Empirical Bayes procedure which recommends for the estimation of hyperparameters from the sample itself.

3. Application

The year-wise record (2002-2011) of the number of HIV patients across 19 States/Union Territories of India have been taken from the National AIDS Control Organisation (NACO)[13] and National Institute of Medical Statistics (NIMS)[11]. NACO is an agency that is committed to contain the spread of HIV in India by building an all-encompassing response reaching out to diverse populations. They strive to provide people with accurate, complete and consistent information about HIV, promote use of condoms for protection, and emphasise treatment of sexually transmitted diseases. National Institute of Medical Statistics (NIMS) aims to promote and undertake research in statistical techniques and methodology in the field of health research, exercise surveillance to ensure the statistical adequacy and validity in various programmes of the Government of India. One of their main thrust areas is modelling, estimation and projection of HIV/AIDS.

4. Results and Discussion

Table 1. Bayes estimates and Standard errors of the average HIV persons (Lakhs) in various states/UT's of India

States/UTs	Gamma Prior		Jeffrey's Non-Informative Prior		Improper Non-informative prior		M.L.E	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Andhra Pradesh	5.188	0.7595	5.238	0.7237	5.288	0.7272	5.188	0.7202
Karnataka	2.558	0.5339	2.608	0.5106	2.658	0.5155	2.558	0.5057
Maharashtra	4.899	0.7471	4.949	0.7035	4.999	0.7070	4.899	0.6999
Manipur	0.267	0.1721	0.317	0.1779	0.367	0.1915	0.267	0.1633
Nagaland	0.169	0.1370	0.219	0.1478	0.269	0.1639	0.169	0.1298
Tamil Nadu	2.314	0.5106	2.364	0.4862	2.414	0.4913	2.314	0.4810
Gujarat	1.372	0.3913	1.422	0.3771	1.472	0.3837	1.372	0.3705
Bihar	0.687	0.2787	0.737	0.2716	0.787	0.2806	0.687	0.2622
Chhattisgarh	0.359	0.2004	0.409	0.2022	0.459	0.2142	0.359	0.1894
Delhi	0.346	0.1961	0.396	0.1989	0.446	0.2111	0.346	0.1859
Haryana	0.317	0.1885	0.367	0.1916	0.417	0.2043	0.317	0.1781
Jharkhand	0.151	0.1295	0.201	0.1417	0.251	0.1583	0.151	0.1228
Kerala	0.527	0.2437	0.577	0.2402	0.627	0.2504	0.527	0.2296
Madhya Pradesh	0.528	0.2435	0.578	0.2404	0.628	0.2506	0.528	0.2298
Orissa	0.353	0.1996	0.403	0.2007	0.453	0.2128	0.353	0.1878
Punjab	0.290	0.1803	0.340	0.1843	0.390	0.1974	0.290	0.1702
Rajasthan	0.448	0.2244	0.498	0.2231	0.548	0.2340	0.448	0.2116
Uttar Pradesh	1.030	0.3396	1.080	0.3287	1.130	0.3362	1.030	0.3210
West Bengal	1.161	0.3699	1.211	0.3481	1.261	0.3552	1.161	0.3408
National Estimate	23.890	1.6925	23.940	1.5473	23.990	1.5489	23.890	1.5457

Table 2. Bayes estimates and standard errors of the average HIV persons (Lakhs) in India for the years 2002-2011

Prior Distributions		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Gamma Prior	per state estimate (λ)	1.32	1.31	1.30	1.29	1.28	1.22	1.26	1.23	0.97	0.91
	National estimate (19*λ)	25.00	24.92	24.75	24.51	24.24	23.16	23.88	23.38	18.51	17.28
	std. error	0.2588	0.2578	0.2561	0.2539	0.2516	0.2443	0.2459	0.2411	0.2201	0.2136
Jeffreys Prior	per state estimate (λ)	1.34	1.34	1.33	1.32	1.30	1.25	1.28	1.26	1.00	0.94
	National estimate (19*λ)	25.50	25.42	25.25	25.01	24.74	23.66	24.38	23.88	19.01	17.78
	std. error	0.2658	0.2654	0.2645	0.2632	0.2618	0.2560	0.2599	0.2572	0.2295	0.2219
Improper Prior	per state estimate (λ)	1.37	1.36	1.36	1.34	1.33	1.27	1.31	1.28	1.03	0.96
	National estimate (19*λ)	26.00	25.92	25.75	25.51	25.24	24.16	24.88	24.38	19.51	18.28
	std. error	0.2684	0.2680	0.2671	0.2658	0.2644	0.2587	0.2625	0.2599	0.2325	0.2250
M.L.E	per state estimate (λ)	1.32	1.31	1.30	1.29	1.28	1.22	1.26	1.23	0.97	0.91
	National estimate (19*λ)	25.00	24.92	24.75	24.51	24.24	23.16	23.88	23.38	18.51	17.28
	std. error	0.2632	0.2627	0.2618	0.2606	0.2591	0.2533	0.2572	0.2545	0.2264	0.2188

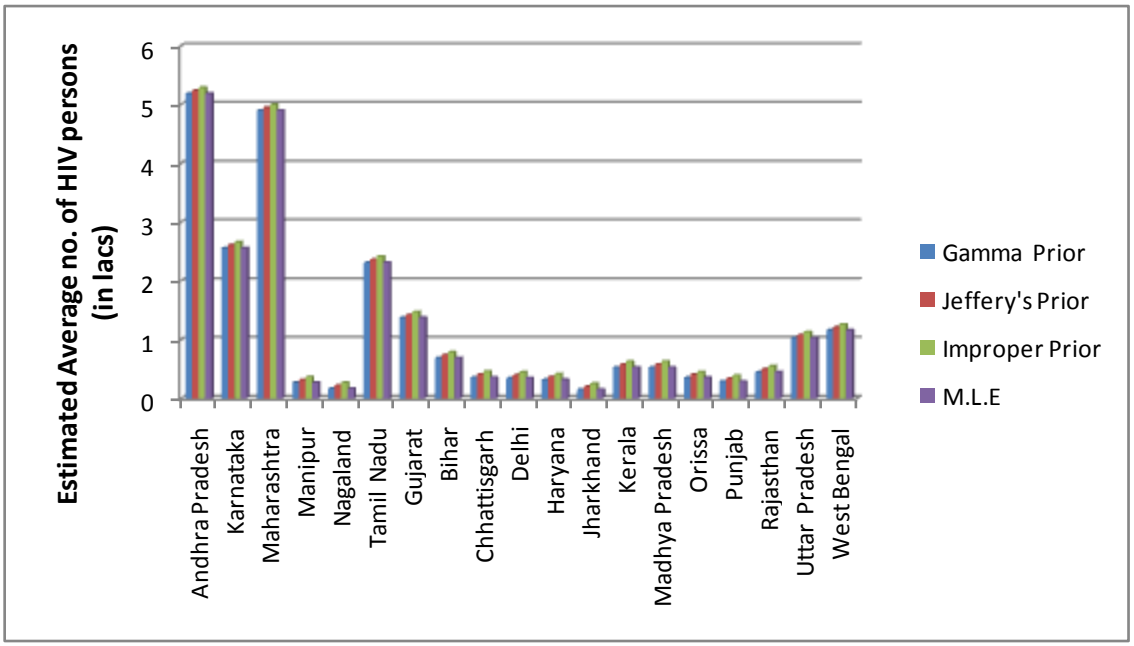

Figure 1. Bayes estimates of the average HIV persons in various states/UT's of India

The National and State-wise Bayes estimates for average HIV population along with their Standard Errors using various prior distributions discussed above are obtained in table 1. Also, the year-wise estimates of average number of HIV patients in India using various priors discussed above are obtained in table 2.

Estimates of the average HIV persons in various States/UT's spread across different years and their corresponding standard errors have been obtained in the tables 1 and 2. As can be seen from figure 1 and 3, the Maximum likelihood estimators are exactly similar to that of Empirical Bayes estimators for a Gamma prior. That is the reason why Empirical Bayes procedure is labeled as partially-Bayesian or sometimes non-Bayesian rather than fully-Bayesian. However, the Standard Errors of these estimates (figure 2 and 4) are not identical and thus provide

the criteria for preferring or rather not preferring Bayesian procedure of introducing prior distributions.

In case of the state-wise estimation of the average number of HIV people (table 1), time (years) was treated as a random variable and it is found that the Maximum Likelihood estimates have the minimum standard errors (figure 2). This implies that the additional information routed through the use of prior distributions is not so much as to produce estimators which are better than the classical M.L.E. In fact, it may be noted that as we gradually move from informative priors to the non-informative ones (M.L.E depicting the most non-informative one as there is no prior distribution involved), the estimates too showed reduction in their standard errors. One of the possible reasons for such a trend could be that the estimators of the incidence were obtained by assuming it to be a time independent constant and hence when time plays the role of a random variable, such estimators may not be viable. The solution to this situation may be explored by considering the hyperparameters as time dependent.

On the other hand, when time was kept fixed and States/UT's were considered as random, the Bayes estimates with Gamma prior had the minimum standard errors followed by M.L.E and the estimates obtained using Jeffrey's and improper prior (table 2). Thus in this case, Bayes estimates are not only admissible but prove to be better than the classical Maximum Likelihood estimates. As can be seen from figure 4, the standard errors of the estimates are overwhelmingly better than the M.L.E. The reason for such results may again be attributed to the last point made in the previous paragraph. Note that, even though the differences in the Standard errors appear to be meager, these are considered significant enough to warrant one procedure over the other because of all the numbers being in lakhs.

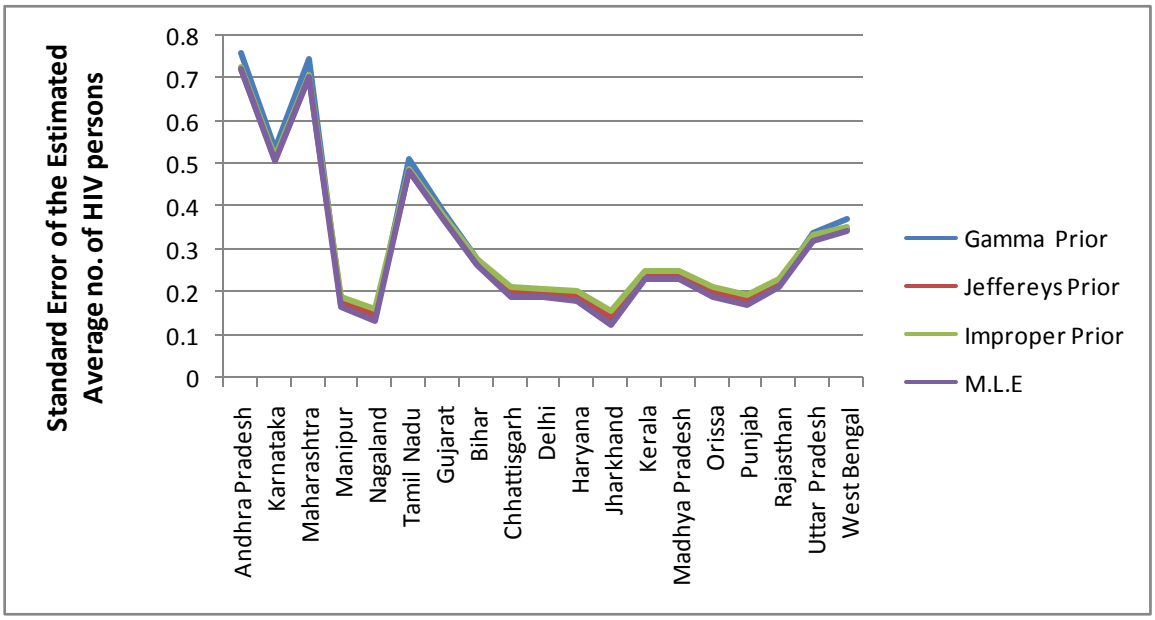

Figure 2. Estimated Standard Error for the Bayes estimates of the average HIV persons in various states/UT's of India

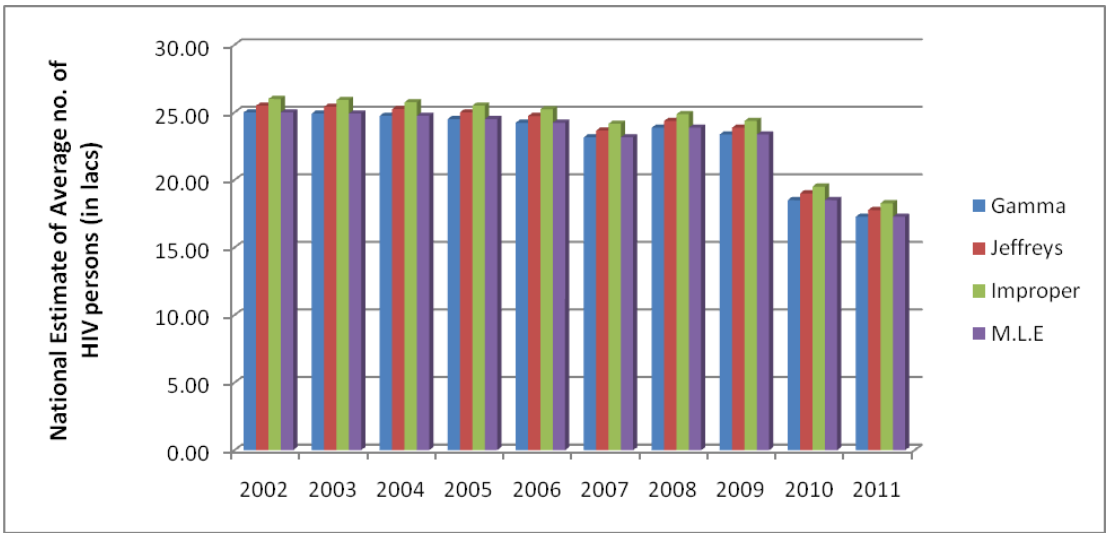

Figure 3. Bayes estimates of the average HIV persons in India for the years 2002-2011

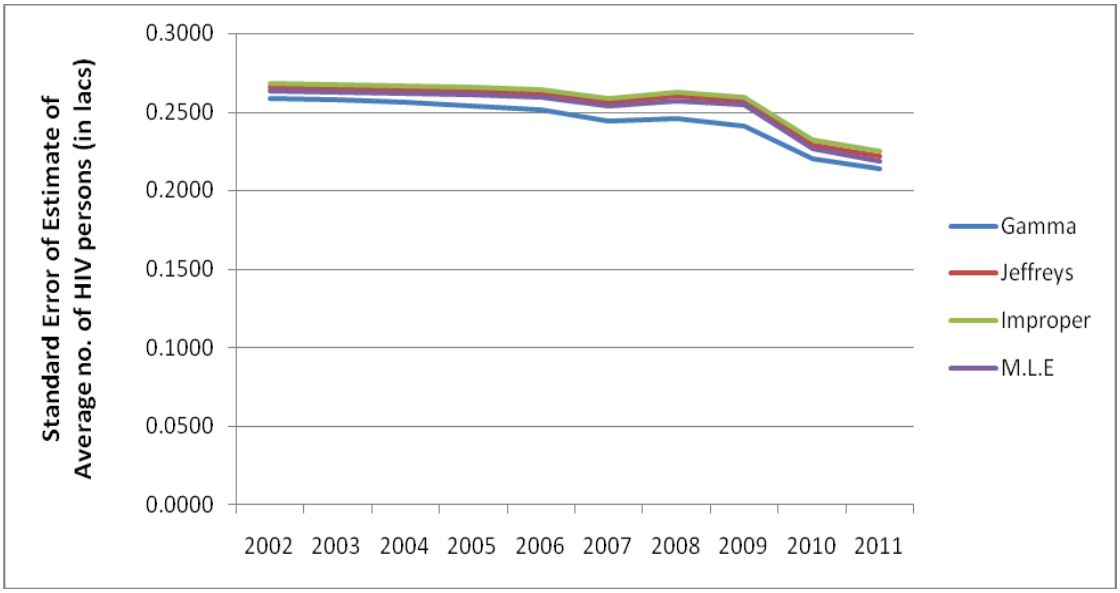

Figure 4. Estimated Standard Error for the Bayes estimates of the average HIV persons in India for the years 2002 -2011

At the end of year 2011, the number of HIV cases in India were estimated to be 2.39 million (Table 1). While Andhra Pradesh reported the highest number of HIV cases i.e., 5,18,800 which is 21.72% of the national estimate, Maharashtra was not far behind with 4,89,900 cases which is 20.51% of the national estimate. Together the high prevalence states of Andhra Pradesh, Maharashtra, Karnataka and Tamil Nadu accounted for 1.5 million HIV cases i.e., 62.62% of the national estimate. On the other hand, Gujarat, West Bengal and Uttar Pradesh recorded 1,37,200, 1,16,100 and 1,03,000 HIV cases which combined form 14.92% of the national estimate. The number of HIV cases in Bihar, Madhya Pradesh, Kerala, Rajasthan, Chhattisgarh, Orissa, Delhi and Haryana range from 30,000 to 1,00,000 while Punjab, Manipur, Nagaland and Jharkhand registered below 30,000 HIV cases.

Pandey et al.[15] estimated the total number of HIV cases in India for the year 2008 to be 2.44 million while in 2009 it was 2.4 million. Their estimates indicated a slowdown in the AIDS epidemic which is corroborated by our national estimate of 2.39 million at the end of year 2011. Hence we notice a decreasing trend in the number of HIV cases in India. However, comparison of our estimates with state-wise estimates of[15] showed a different trend. According to[15], high prevalence states showed a stable or decreasing trend while low-to-moderate prevalence states exhibit a varying trend. The high prevalence states account for 62.62% of all the HIV infections as compared to their 57%. This shows that the number of infections in high prevalence states has increased by 5.62% as compared to year 2009. West Bengal, Gujarat, Bihar and Uttar Pradesh together account for 17% of the national estimate while it was 22% in the year 2009. The number of HIV cases in Punjab, Orissa, Rajasthan, Madhya Pradesh together account for 6.78% which is a steep decrease from 12% of the national estimate in the year 2009. Among the high prevalence states, the number of HIV cases in Andhra Pradesh increased from 21% to 21.72%, Maharashtra increased from 18% to 20.51%, Tamil Nadu increased from 10% to 10.70% and Karnataka increased from 7% to 9.68% of the National estimate as compared to year 2009.

Table 3. Estimated Number and %HIV cases with respect to Area(sq. Km) and Population

STATE	No. of HIV (per sqkm)	% HIV (in population)
Andhra Pradesh	1.8862	0.6128
Maharashtra	1.5921	0.4360
Karnataka	1.3337	0.4184
Tamil Nadu	1.7792	0.3208
Gujarat	0.6999	0.2272
West Bengal	1.3081	0.1271
Uttar Pradesh	0.4234	0.0516
Bihar	0.6925	0.0662
Madhya Pradesh	0.1713	0.0727
Kerala	1.3560	0.1578
Rajasthan	0.1309	0.0653
Chhattisgarh	0.2655	0.1406
Orissa	0.2265	0.0842
Delhi	23.3311	0.2065
Haryana	0.7170	0.1250
Punjab	0.5758	0.1047
Manipur	1.1948	0.9810
Nagaland	1.0194	0.8533
Jharkhand	0.2022	0.0458
National Estimate	**0.8777**	**0.2105**

The above findings give an idea about the concentration of HIV cases across various states of India on the basis of which we could segregate high prevalence areas from the moderate and low-prevalence ones. On the other hand, if we look at the density of the HIV cases in terms of area of each state(in sq. Km.), Delhi records a whopping surge over all the remaining states with 23 persons per square Kilometer followed by

Andhra Pradesh, Tamil Nadu and Maharashtra and others as distant next few positions. Chhattisgarh, Orissa, Madhya Pradesh and Rajasthan record the minimum number of HIV persons per square Kilometer area. Also, if the population of States is considered, Manipur records the highest percentage of HIV population closely followed by Nagaland and Andhra Pradesh. The lowest percentage i.e., less than 0.1% of HIV cases in the population is noticed in the States of Orissa, Madhya Pradesh, Bihar, Rajasthan, Uttar Pradesh and Jharkhand. The results are given in Table 3 and illustrated in Figure 5. Also, the census data pertaining to the Area (in Sq. Km) and the population is obtained from[12].

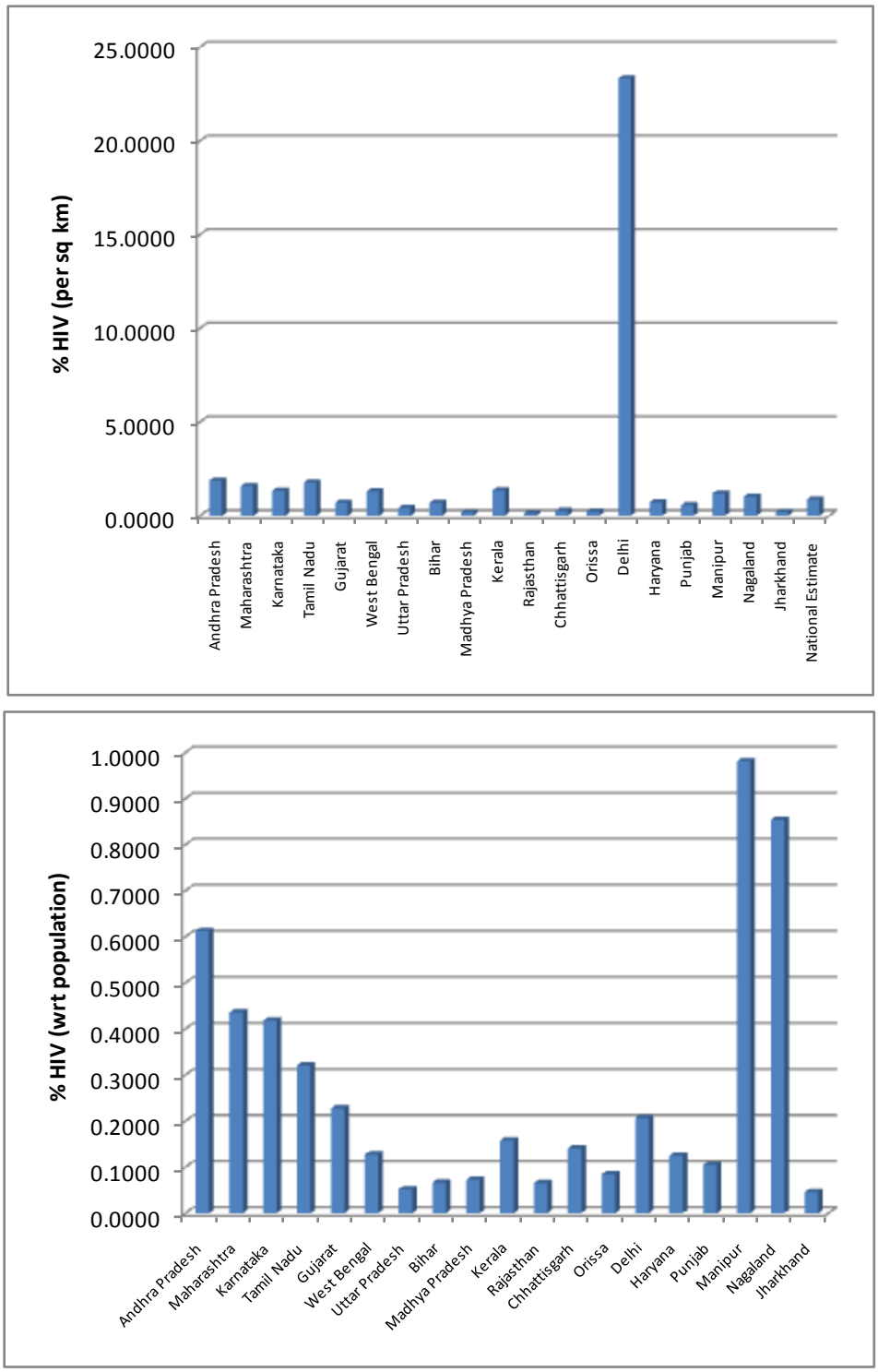

Figure 5. Estimated % HIV with respect to Area(sq. Km) and Population

5. Conclusions

From the results, we note that the average number of HIV infected persons in India is declining over the years. Before coming to any conclusion on this, it may be explored whether the decrease is indeed a good sign or is a situation where they are either dying or converting to AIDS. It may also be noted that the states of Andhra Pradesh and Maharashtra have recorded high incidence of HIV cases while the lowest is seen in the Jharkhand and Nagaland. The high prevalence states of Andhra Pradesh, Maharashtra, Karnataka and Tamil Nadu show an increase in the incidence in 2011 as compared to 2009. Delhi records the highest prevalence in terms of the number of HIV cases per square Kilometer area while the lowest is seen in Rajasthan. In terms of population of each state, the highest percentage of HIV cases are seen in Manipur and Nagaland while the lowest is seen in Jharkhand.

The results invariably strike a balance between Classical and Bayesian procedures by not discriminating one over the other. However, it may be noted that the Bayesian procedure is a fairly general procedure which may encompass the classical procedure by simply making assumptions on the hyperparameters. This paper obtains very good estimates of the HIV infection rate by relating it to the number of infected people and hence, subject to the assumptions of constant infection rate, prior distributions provides a good utility in the estimation procedure.

Even though, this procedure can be generalized for calculating the national averages of HIV infectives, further improvements may be done by developing a procedure that incorporates time dependent Incidence Rate in the Poisson model. Such a model would refine the Bayes estimators to perform well, even for a time-series data. The incidence rate may be verified for its dependence on certain covariates and suitable Bayesian approach may be applied to it. Also, the estimators may further be improved to accommodate the case of incomplete data sets.

REFERENCES

[1] Alfred A. Bartolucci, Karan P. Singh, Ramalingham Shanmugam, 2004, "Empirical Bayesian analysis of the Poisson intervention and incidence parameters," Mathematics and Computers in Simulation, Vol. 64, 393-399.

[2] Alphonse Kpoz ĉhouen, Ahmadou Alioum, et al., 2004, "Use of a Bayesian Approach to Decide When to Stop a Therapeutic Trial: The Case of a Chemoprophylaxis Trial in Human Immunodeficiency Virus Infection," American Journal of Epidemiology, Volume 161, Issue 6, 595-603.

[3] Berry, D.A., Wolff, M.C., Sack, D., 1994, "Decision making during a phase III randomized controlled trial," Controlled Clinical Trials, 15, 360-78.

[4] Brenner, D., Fraser, D.A.S., and McDunnough, P., 1982, "On asymptotic normality of likelihood and conditional analysis," can. J. Statist., 10, 163-172.

[5] Dankmar BoÈhning, 1998, "Zero-Inflated Poisson Models and C.A.MAN: A Tutorial Collection of Evidence", Biometrical Journal, 40 (7), 833±843.

[6] Deuchert E, Brody S., 2007, "Plausible and implausible parameters for mathematical modeling of nominal heterosexual HIV transmission," Ann Epidemiol., 17(3), 237-44.

[7] Elena E Giorgi, Bob Funkhouser, Gayathri Athreya, et al. (2010). "Estimating time since infection in early homogeneous HIV-1 samples using a poisson model," BMC Bioinformatics, 11, 532.

[8] Fraser, D.A.S., and McDunnough, P. (1984). "Further remarks on asymptotic normality of likelihood and conditional analysis," can. J. Statist., 12, 183-190.

[9] Gurprit Grover, Dulumoni Das and Rupak Gupta (2011). "On the Estimation of Expected Proportion of AIDS Cases Using Bayesian Approach," The IUP Journal of Computational Mathematics, Vol. IV, No. 2, 37-48.

[10] H. Jeffreys., 1946, Methods of Mathematical Physics, Cambridge University Press.

[11] http://www.icmr.nic.in/icmrsql/insprofile.asp?insno1=000134 (20th April 2013)

[12] http://www.mapsofindia.com/maps/india/indiastateandunion.htm (25th June, 2013)

[13] http://www.nacoonline.org/NACO/About_NACO/Vision_and_Values/ (28th April 2013)

[14] Nel A, Louw C, Hellstrom E, Braunstein SL, Treadwell I, et al. (2011). "HIV Prevalence and Incidence among Sexually Active Females in Two Districts of South Africa to Determine Microbicide Trial Feasibility," PLoS ONE 6(8): e21528. doi:10.1371/journal.pone.0021528.

[15] Pandey A, Sahu D, Bakkali T, et al. (2012), "Estimate of HIV prevalence and number of people living with HIV in India 2008-2009," BMJ open 2012;2:e000926. doi:10.1136/bmjopen-2012-000926.

[16] Richard G. White, S. Cooper Ben et al. (2007). "Quantifying HIV-1 transmission due to contaminated injections," PNAS vol. 104 no. 23, 9794-9799.

[17] Robbins, H., 1955, "An Empirical Bayes Approach to Statistics," in 119-127. Proceedings of the 3rd Berke-ley Symposium on MathematicaI Statistics and Probability (Vol. l), Berkeley, CA: University of California Press, 157-163

[18] Stephen R. Cole, Haitao Chu and Ron Brookmeyer (2007). Confidence Intervals for Biomarker-based Human Immunodeficiency Virus Incidence Estimates and Differences using Prevalent Data, American Journal of Epidemiology, 165, 94-100.

[19] W. Y. Tan, Tan Wai-Yuan (2000). Stochastic Modeling of AIDS Epidemiology And HIV Pathogenesis, World Scientific Publication, Singapore.

[20] Yinglin Xia, Dianne Morrison-Beedy, Jingming Ma, Changyong Feng, Wendi Cross, and Xin Tu (2012). "Modeling Count Outcomes from HIV Risk Reduction Interventions: A Comparison of Competing Statistical Models for Count Responses, AIDS Research and Treatment," Article ID 593569.