EPH - International Journal of Mathematics and Statistics
ISSN (Online): 2208-2212
Volume 3 Issue 2 October 2017

DOI:https://doi.org/10.53555/eijms.v3i2.9

BAYESIAN SEQUENTIAL ESTIMATION OF PROPORTION OF ORTHOPEDIC SURGERY OF TYPE 2 DIABETIC PATIENTS AMONG DIFFERENT AGE GROUPS –A CASE STUDY OF GOVERNMENT MEDICAL COLLEGE, JAMMU-INDIA

Roohi Gupta[1], Priyanka Anand[2] and Rahul Gupta[3]*
*[1,2,3]*Department of Statistics, University of Jammu,*

Corresponding Author:-
E-mail:- rahulgupta68@yahoo.com

Abstract:-
A study to determine the proportion of diabetic cases handled by the hospitals in Jammu regarding orthopedic surgeries, especially among different age group, by employing a Bayesian model called Beta-binomial conjugate model using Bayesian sequential estimation method. From the hospital record of Government Medical College (GMC), Jammu. Data was collected and collated over a period of 3 years from 2014-2016 in terms of 4 quarters (Jan-March, April-June, July-Sep and OctDec). The study also compares the computed sample and Empirical Bayes estimates over the three year period and also the variances of the computed estimates and it has been strongly advocated that Empirical Bayes estimators are better estimators on the basis of efficiency and consistency and the proposed method can be employed successfully in any suitable location globally.

Key words:-*Bayesian, Empirical Bayes, Beta-Binomial conjugate model, Estimation of proportion.*

1. INTRODUCTION AND BACKGROUND

Bayesian solutions provide many instinctive and reasonable conclusions than likelihood based inferences, which is simply because Bayesian analysis includes prior information as well as likelihood function. In empirical Bayes(EB) approach, data is used to help determination of the prior through estimation of the so called hyperparameters. The considerable impact of Empirical Bayes (EB) on statistical applications continues to obtain increasing popularity since its introduction about four decades ago. The EB structure combines information from several but similar sources. The primary interest in EB analysis is in the hyperparameters (□) rather than the parameters, from individual studies (□). More generally, the hierarchical structure allows for appraisal of heterogeneity both within and between groups (Carlin and Louis(1)). Thus Bayesian approach to parameter estimation, when conditions of data allow, is to estimate the posterior distribution of the parameter(s) in question (□) so that inference on □ is then based on the posterior distribution. A prior distribution for □ is needed to derive the posterior and, in some cases, the prior may have its own parameters, called hyperparameter(s) (□). Quite often, □ is unknown to the analyst, in which case the prior is not completely specified. One way of resolving this problem is through empirical Bayes (EB) analysis. More importantly, EB can lead to more precise estimates than sampling theory approaches (Rubin(2). EB analysis also provides a more dependable ranking of parameters and aids in the identification of extreme values in the group(Link and Hahn(3). These properties of EB derive from the fact that it uses related supplementary data which frequentist inference ignores(Okafor(4)). The EB concept was first proposed by Robbins(5) in a non-parametric setting. Significant works in this regard has been done by Deely and Lindley(6), Hui and Berger(7), Morris(8), Wong(9), Smith(10), Raghunathan(11), Altman and Casella(12), Efron(13) and Carlin and Louis(14). Significant work on this field include Ogundeji and Okafor(15), Okafor,et al(16) and Okafor and Mbata(17).

By Type 2 diabetes, the study refers to the branch of diabetes concerned with conditions involving the endocrine system. Type 2 diabetes is a progressive condition, meaning that the longer someone has it, the more "help" they will need to manage blood glucose levels. This will require more medications & eventually, injected insulin will be needed. People with type 2 diabetes produce insulin, but their bodies don't use it correctly, this is referred to as being insulin resistant. People with Type 2 diabetes may also be unable to produce enough insulin to handle the glucose in their body. In these instances, insulin is needed to allow the glucose to travel from the bloodstream into our cells, where it's used to create energy.

In an attempt to determine the proportion of diabetic cases handled by the hospital regarding orthopedic surgeries, especially among different age group, we employed a Bayesian model called Beta-binomial conjugate model using Bayesian sequential estimation method, on the lines of Ogundeji et.al(18). From the hospital record of Government Medical College(GMC), Jammu, data were collected and collated over a period of 3 years from 2014-2016 in terms of 4 quarters (Jan-March, April-June, July-Sep and Oct-Dec). The study also compares the computed sample and EB estimates over the three year period and also the variances of the computed estimates.

2.THE BAYESIAN SEQUENTIAL PROCEDURE

If a set of observations x_1 , x_2, x_3 , . , . , . , x_n generates a posterior division and, in a comparable situation, supplementary data are collected beyond these observations, then the posterior distribution found with earlier observations becomes the new prior distribution and the additional observation give a new posterior distribution and conclusion can be made from the second posterior distribution. This process can go on with even more observation. In other words, the second posterior becomes the new prior and the next set of observations give the next posterior from which the inference can be made. This is the principle of Bayesian sequential methodology that we applied to estimate the proportion of counted data obtained from the hospital.

Based on the Bayesian approach described above, data were collected monthly and collated yearly for three years (2014, 2015 and 2016) from the hospital records of Government Medical College(GMC), Jammu. Tie population proportion of diabetic patients admitted for orthopedic surgery in GMC is denoted by P_0 while the proportion of diabetic patients admitted for orthopedic surgery in age group j is P_j (j = 1, 2, . . ., 5). X_0 represents a random outcome of patient examined in age group j.

$$Y_{ijk} = \begin{cases} 1, \text{if ith diabetic Patient is admitted} \\ \text{for orthopedic surgery in} \\ \text{age group j and in year k} \\ 0 \text{, otherwise} \end{cases}$$

$$Y_{jk} = \sum_{i=1}^{n_{ijk}} Y_{ijk} = \text{ the total number of diabetic patients admitted for orthopedic surgeries in}$$

$$\text{age group j in the year k ,}$$

n_{jk} = the total number of diabetic patients admitted for treatments (both orthopedic and nonorthopedic surgeries) in GMC Jammu in age group j in year k.

$P_{jk} = \frac{Y_{jk}}{n_{jk}} =$ the proportion of diabetic patients admitted for orthopedic surgery in age group j and year k.

For each year in each age group, we computed sample proportions P_{jk} as follows:

In 2014 and age group j: $P_{j1} = \frac{Y_{j1}}{n_{j1}}$

In 2015 and age group j: $P_{j2} = \frac{Y_{j2}}{n_{j2}}$

In 2016 and age group j: $P_{j3} = \frac{Y_{j3}}{n_{j3}}$

Estimators of sample proportions: $\hat{P}_{jk} = \frac{Y_{jk}}{n_{jk}}$ *and*

$$Var(\hat{P}) = \frac{\hat{P}_{jk}(1-\hat{P}_{jk})}{n_{jk}}$$

3. THE BETA – BINOMIAL CONJUGATE MODEL

The Empirical Bayesian model to be realistic is a conjugate beta – binomial model where the binomial distribution represents the likelihood of the observed data likelihood and the beta distribution taken as the prior distribution of the binomial parameter. The posterior mean is

$$\tilde{P}_{jk} = \int P_{jk} f(P_{jk} \mid Y_{jk}, \eta) dp_{jk} \qquad (1)$$

A key component of this integral is f (P_{jk} |Y_{jk} ,□). The posterior distribution of which is P_{jk}. Under the general Bayesian framework and using the beta conjugate prior plus the binomial likelihood, the posterior distribution of P_{jk} is:

$$f(P_{jk} \mid Y_{jk}, \eta) = \frac{\binom{n_{jk}}{Y_{jk}} P_{jk}^{Y_{jk}} (1-P_{jk})^{n_{jk}-Y_{jk}} \frac{1}{B(r,s)} P_{jk}^{r-1}(1-P_{jk})^{s-1}}{\int P_{jk}^{Y_{jk}} (1-P_{jk})^{n_{jk}-Y_{jk}} \frac{1}{B(r,s)} P_{jk}^{r-1}(1-P_{jk})^{s-1} dP_{jk}}, \eta = (r.s) \qquad (2)$$

There is need to estimate the hyper-parameters r and s of the beta distribution in order to completely specify the prior. This can be achieved easily through reparameterisation of $f(P_{jk} \mid \eta)$, and using moment estimator. Letting $P_0 = \frac{r}{r+s}; M = r+s$ and using the prior distribution of $P_{jk}; E(P_{jk}) = P_0$ and $Var(P_{jk}) = \frac{rs}{(r+s+1)(r+s)^2} = \frac{P_o(1-P_0)}{M+1}$. These are known as prior mean and variance respectively. Consequently,

$$f(P_{jk} \mid Y_{jk}, \hat{\mu}, \hat{M}) = \frac{1}{B(\alpha,\beta)} P_{jk}^{\alpha-1}(1-P_{jk})^{\beta-1} \quad (3)$$

where,

$$\hat{\alpha} = Y_{jk} + \hat{M}\hat{P}_0;$$
$$\hat{\beta} = n_i - Y_{jk} + \hat{M}(1-\hat{P}_0),$$

$$\hat{P}_0 = \frac{\sum Y_{jk}}{\sum n_{jk}}$$

and

$$\hat{M} = \frac{\hat{P}_0(1-\hat{P}_0) - S_P^2}{d}$$

$$S_P^2 \frac{\hat{P}_0(1-\hat{P}_0)\sum \frac{1}{n_{jk}}}{N}$$

where $S_P^2 = \frac{N\sum n_{jk}(\hat{P}_{jk} - \hat{P}_{jk})^2}{(N-1)\sum ni}$ with M and P_0 estimated, then,

$$\widetilde{P}_{EB} = E(P_{jk} / Y_{jk}, \hat{P}_0, \hat{M})$$
$$= \left(\frac{\hat{M}}{n_{jk}+\hat{M}}\right)P_0 + \left(\frac{n_{jk}}{n_{jk}+\hat{M}}\right)\frac{Y_{jk}}{n_{jk}} \, and \quad (4)$$

$$Var(\widetilde{P}_{EB}) = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}. \quad (5)$$

Consequently, $\hat{\lambda} = \frac{\hat{M}}{n_{jk}+\hat{M}}$ and it can be easily seen that where $\hat{M}$ (the scale factor) is large relative to n_{jk}, λ is large and $\hat{P}_0$ receives a larger weight than $\frac{Y_{jk}}{n_{jk}}$. But large $\hat{M}$ implies small prior variance. Thus, the estimate which is associated with smaller variance receives larger weight in determining the posterior mean $\widetilde{P}_{EB}$. On the other hand, if $\hat{M}$ is small relative to n_{jk}, the sample mean receives more weight. Note that the posterior density for the overall age group proportion P_0 is obtained by replacing Y_{jk} and n_{jk} in equation (3) with Y and N, respectively.

Under conjugacy, the EB estimator of a proportion $\hat{P}_{if}$ is a weighted mean of two estimators, the mean of the prior density P_0 and the sample proportion estimator $\hat{P}_{if}$. Thus,

$$\tilde{P}_{EB} = \lambda P_0 + (1-\lambda)\hat{P}_{if} \quad (6)$$

$\tilde{P}_{EB}$ is the empirical Bayes Estimators with λ as the shrinkage factor. λ is a function of the prior and sample estimator variances such that, if variance of sample estimator is large, the weight of $\hat{P}_0 (i.e. \lambda)$ will be large and $\tilde{P}_{EB}$ will shrink towards $\hat{P}_0$. Two components of the above model $\lambda \, and \, \hat{P}_0$ are derived from the Empirical Bayesian process, proposed by Carlin and Louis(2000).

4. RESULTS

The empirical results of the application of Beta Binomial model and Bayesian sequential approach to the data of different age group observations (patients) for the three years (2014, 2015 and 2016) are presented in Table 1 and 2. The hyper parameters □ and M are anticipated using sample information. These were consequently used to establish the statistical constants of the posterior distributions □and□ thereby entirely specifying them.

In this analysis, comparisons are made on yearly basis estimated sample proportions and Empirical Bayesian proportions as well as variances of estimated sample proportions and Empirical Bayesian proportions.

Table 1
(Comparative Analysis of Estimated Sample proportions & EB Proportions)

Age group	Jan-March		April-June		July-Sept		Oct-Dec	
	Pji	PEB	Pj2	PEB	Pj3	PEB	Pj4	PEB
35-44	0.375	0.411	0.227	0.284	0.217	0.184	0.161	0.137
45-54	0.208	0.226	0.265	0.331	0.391	0.503	0.217	0.189
55-64	0.083	0.088	0.189	0.237	0.313	0.269	0.354	0.311
65-74	0.25	0.272	0.113	0.144	0.144	0.208	0.411	0.346

Table 2
(Comparative Analysis of variances of Estimated Sample proportions & EB Proportions)

Age group	Quarter 1		Quarter 2		Quarter 3		Quarter 4	
	Var(P_{ji})	Var(P_{EB})	Var(P_{j2})	Var(P_{EB})	Var(P_{j3})	Var(P_{EB})	Var(P_{j4})	Var(P_{EB})
35-44	0.097	0.0021	0.369	0.0019	0.304	0.00004	0.406	0.0008
45-54	0.319	0.0015	0.230	0.0020	0.081	0.0013	0.304	0.0010
55-64	0.569	0.0007	0.352	0.0017	0.470	0.0014	0.117	0.0014
65-74	0.449	0.0017	0.161	0.0011	0.261	0.0012	0.065	0.0014

Over the years under deliberation, the results show that the highest number of patients attended to at the hospital is among the age group 15 to 44 years but this age group also has the smallest proportion of diabetic surgeries. Similarly, the smallest number of patients can be found among the two age groups 'less than one Year' and 'greater than 64'. Also the highest proportion of diabetic surgeries is among patients whose ages are greater than 64 years followed by those who are less than one year old.

The overall EB proportion of patients admitted for diabetic surgeries in the hospital across the age groups enlarged steadily (0.4629, 0.5228 and 0.5591) over the" respective years of study.

Furthermore, the computed variances of the sample and EB estimates are smallest amid the age groups '15-44 years' and '44 - 64 years' while highest variances are noted in age group 'less than one' and 'greater than 64 years' (see Table 2).

5. SUMMARY AND CONCLUSIONS

This effort has been able to demonstrate how Bayesian sequential estimation of proportion can be applied to statistical progression control for different ages of the patients and years. The consequence of the investigation was compared both of yearly estimated sample proportions and EB proportions plus the variances. It was found that overall EB proportion of diabetic patients admitted for orthopedic surgeries in the hospital across the age groups increased gradually. Similarly, the overall variances of the proportions be inclined more to zero over the three years under review. Thus, the results show that the Empirical Bayesian estimators are better estimators on the basis of efficiency and consistency properties of good estimators.

REFERENCES

[1]. **Carlin, B. P. and Louis, T. A**. "Empirical Bayes: Past, Present and Future", Journal of the American Statistical Association, 2000, 95(452): 1286-1289.

[2]. **Rubin,D.** Using empirical Bayes techniques in the few school validity studies. Journal of the American Statistical Association, 1980, 75: 801-816.

[3]. **Link, W. A. and Hahn, D. C.** "Empirical Bayes Estimation of Proportions with Application to Cowbird Parasitism Rates", Journal of Ecological Society of America, 1986, 77(8): 2528-2537.

[4]. **Okafor, R.** , "Using an Empirical Bayes Model to estimate currency exchange rate", Journal of Applied Statistics, 1999, 269(80): 973-983.

[5]. **Robbins, H.** "An empirical Bayes approach to Statistics", Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1955, Vol. 1.

[6]. **Deely, J. and Lindley, V**. "Bayes empirical Bayes", Journal of the American Statistical Association, 1981, 76: 833-841.

[7]. **Hui, S. and Berger, J.** "Empirical Bayes estimation of rates in longitudinal studies", Journal of the American Statistical Association, 1983, 78: 753-760.

[8]. **Morris, C**. "Parametric empirical Bayes inference: Theory and applications", Journal of the American Statistical Association, 1983, 78: 47-65.

[9]. **Wong, G.** "Bayesian models for directed graphs", Journal of the American Statistical Association, 1987, 82: 140-148.

[10]. **Smith, P. J.** "Bayesian analyses for a multiple capture-recapture model", Biometrika, 1991, 78:399-407.

[11]. **Raghunathan, T. E**. "A Quasi-Empirical Bayes methods for Small Area Estimation" Journal of the American Statistical Association, 1993, 88: 1444-1448.

[12]. **Altman, N. S. and Casella, G**. "Nonparametric Empirical Bayes Growth Curve Analysis", Journal of the American Statistical Association, 1995, 90: 508-515.

[13]. **Efron, B.** "Empirical Bayes Methods for Combining Likelihoods (with discussion)", Journal of the American Statistical Association, 1996, 91: 538-565.

[14]. **Carlin, B. P. and Louis, T. A.** Bayes and Empirical Bayes Methods for Data Analysis, Boca Raton, Florida: Chapman and Hall/CRC Press, 2000.

[15]. **Ogundeji, R. K. and Okafor, R. O.** "Empirical Bayes Approach to Estimate Mean CGPA of University Graduates", International Journal of Applied Mathematics & Statistics 2010, 17(J10): 77-89.

[16]. **Okafor R.O., Opara A.I., Ogundeji R.K., Mbata U.A., and Olalude G. A.** "Empirical Bayes Model for Inference on Vehicular Traffic Density at the Main Campus of University of Lagos, Lagos, Nigeria", International Journal of Statistics and Systems, 2010, 5 (1): 53–62.

[17]. **Okafor R.O and Mbata U.A.** " A Bayesian Model for Estimation of Population Proportions", Section on Bayesian Statistical Science –Joint Statistical Meeting 2011,USA.

[18]. **Ogundeji R. K., Adewara, A.J. and Nurudeen T.S.** Bayesian Sequential Estimation of Proportion of Orthopedic Surgery among Different Age Groups: A Case Study of National Orthopedic Hospital, Igbobi-Nigeria; International Journal of Statistics and Applications, 2012, 2(6):108-113.