Journal of Modern Mathematics and Statistics 3 (3): 56-59, 2009
ISSN: 1994-5388

Bayesian Estimation of the Proportion of Subscribers to Nigeria's National Health Insurance Scheme

[1]I.A. Adeleke, [1]Ibiwoye Ade and [2]R.K. Ogundeji
[1]Department of Actuarial Science and Insurance,
[2]Department of Mathematics, Lagos University, Nigeria

Abstract: Since, the National Health Insurance Scheme (NHIS) came on stream in Nigeria in 1999, it had been assumed that employees, particularly those in the organized formal sector would voluntarily participate in the scheme, given the inconvenience many had experienced with the previous arrangements. However, participation rate appears to be below expectation. What is even more critical is that there is no idea what this rate of participation might be since, there has been no empirical study to determine it. This tends to make any plan towards mobilizing the populace for embracing the NHIS option a difficult task. In response to this need, this study uses Bayesian approach to estimate the proportion of subscribers to the scheme and makes policy recommendations.

Key words: NHIS, Bayesian, subscription, proportion, estimation, empirical

INTRODUCTION

The National Health Insurance Scheme (NHIS) was introduced in Nigeria with the promulgation of decree No. 35 of 1999. The broad objective of the scheme is to ensure that every Nigerian has access to good health care services at affordable costs. Participants are expected to pay capitation fees to licensed HMOs, which would allow the subscriber to have access to registered health care providers. Given the inefficiencies experienced with public hospitals and the rather expensive costs of private hospitals, it was expected that the populace would readily embrace the scheme. For some reasons this does not appear to be the reality. Worse still, it is unclear the level of participation. Thus, it is difficult to know what the scale should be of say, an enlightenment programme or some other strategy that could boost participation, which need to be put in place.

This study uses Bayesian methods of inference to estimate the rate of participation in the NHIS scheme based on the outcome of a survey conducted in an earlier study. It derives motivation from Rubin's (1983) illustrations of the Bayesian method as an extremely powerful tool for the applied statistician especially in the way, it could provide sensible answers in a straightforward manner in problems where, sampling theories approaches appear awkward. Further support is derived from Link and Sauer (1996), who argued that Empirical Bayes methods provide alternative approaches that incorporate the structural advantages of Bayesian models while, requiring less stringent specification of prior knowledge.

MATERIALS AND METHODS

Bayes methods: Empirical Bayes methods have received considerable attention in the statistical study since, their introduction in the early part of the 20th century (Robbins, 1956). Morris (1983) argued that when considered in the context of a group of related parameters the Bayesian procedures yield improved estimates of individual parameters. Also in what appears to be a bold attempt at making a case for the adoption of the method Rubin (1983) compared the frequentist approach to the Bayesian method and concluded that the frequentist perspective appears to be devoid of easily followable principles that would lead to the construction of good inferences whereas, the Bayesian perspective naturally lead to the construction of such inferences.

Empirical Bayes methods have been applied in a number of contexts. Some recent examples include applications to problems in forest science (Burk and Ek, 1982) to monitoring of air pollution (Suggs and Curran, 1983) and in a variety of medical applications (Stijnen and van Houwelingen, 1990). It would seem that applications of empirical Bayes methodology in nonmedical biological settings are few. In the ecological study where, some application had been recorded for instance, Johnson (1989) noted that in spite of the theoretical justification of empirical Bayes methods, their use has not been widespread. Within a period of one decade since this apparent disinterest in the methods, there seems to be an awakening. Recent applications include estimating population sizes from survey data (Johnson, 1989), estimating numbers of species

Corresponding Author: I.A. Adeleke, Department of Actuarial Science and Insurance, Lagos University, Nigeria

(Mingoti and Meeden, 1992), capture-recapture data (Smith, 1991), toxicity data (Piegorsch, 1994), summary analysis of avian trends (Link and Sauer, 1995) and identification of extremes in collections of parameter estimates (Link and Sauer, 1996).

Link and Sauer (1996) advanced that a Bayesian analysis can be thought of as a combination of existing knowledge with new knowledge, the two being synthesized in such a way as to account for the amount of confidence that is placed in each source of knowledge.

Usually, the existing knowledge base and updated knowledge base are summarized by probability distributions describing the likely range of values for each unknown parameter, which are referred to as the prior and posterior distributions, respectively. Informally, the prior distribution can be thought as an approximation to the histogram of the true, unknown values of the parameters under investigation.

Application of empirical Bayes analysis to NHIS data: In the original data, the respondents were structured along eight categories, viz.: educational qualification, marital status, number of children, number of children under 18 years, income per annum, occupation, awareness of NHIS and gender. Therefore, for a class of subscribers we let X denote the number of persons registered with HMO in a random sample of N subscribers. Given that the true proportion is P, X is assumed to be a binomial random variable with mean (NP). This second estimate is the observed proportion given as:

$$\hat{P} = X/N$$

Observed proportion P has variance and standard error given, respectively as:

$$\sigma^2 = [P(1-P)/N] \text{ and } \sigma = \sqrt{P(1-P)/N}$$

Then, for a single observation the marginal likelihood of the number of subscribers can be represented by the β-Binomial distribution (Leonard and Hsu, 1999) thus:

$$\begin{aligned} p(X_i = k/\alpha,\beta) &= \int p(X_i = k/\theta_i)p(\theta_i/\alpha,\beta)d\theta_i \\ &= \int Bin(k/n_i,\theta_i)Be(\theta_i/\alpha,\beta)d\theta_i \\ &= \binom{n_i}{k}\frac{B(k+\alpha, n_i-k+\beta)}{B(\alpha,\beta)} \\ &= Bb(k/\alpha,\beta,n_i) \end{aligned}$$

Using the moments, the expectation and variance follow:

$$E[X\ ] = n_i\frac{\alpha}{\alpha+\beta}$$

and

$$Var[X_i] = \frac{n_i\alpha\beta}{(\alpha+\beta)^2}\frac{\alpha+\beta+n_i}{\alpha+\beta+1}$$

so,

$$\begin{aligned} E[X_i^2] &= Var[X_i] + (E[X_i])^2 \\ &= \frac{n_i^2\alpha^2(\alpha+\beta+1)+n_i\alpha\beta(\alpha+\beta+n_i)}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= \frac{n_i\alpha(\alpha+\beta)+(n_i\alpha+n_i+\beta)}{(\alpha+\beta)(\alpha+\beta)(\alpha+\beta+1)} \\ &= \frac{n_i\alpha(n_i\alpha+n_i+\beta)}{(\alpha+\beta)(\alpha+\beta+1)} \end{aligned}$$

We equate these to the empirical moments:

$$a = \frac{1}{n}\sum_{i=1}^{n}X_i$$

and

$$b = \frac{1}{n}\sum_{i=1}^{n}X_i^2$$

and taking $n_i = t$,

$$\hat{\alpha} = \frac{a(b-ta)}{a((t-1)a+t)-tb}$$

$$\hat{\beta} = \frac{(t-a)(b-ta)}{a(4a+t)-ta}$$

We now re-parameterize the posterior distribution of P by setting:

$$\pi = \frac{\alpha}{(\alpha+\beta)}$$

and

$$\theta = (\alpha+\beta)$$

The mean of the posterior distribution easily follows:

$$\begin{aligned} E\ (P/\hat{P}) &= \frac{\alpha+X}{\alpha+\beta+N} \\ &= \pi\left(\frac{\theta}{\theta+N}\right)+\hat{P}\left(\frac{N}{\theta+N}\right) \end{aligned}$$

RESULTS AND DISCUSSION

The result of the re-parametized hyper parameters from the NHIS data are shown in Table 1.

To obtain empirical Bayes estimates, we need only substitute estimates of the hyper parameters π and θ in the formula for the Bayes estimator. Following Link and Sauer (1995), we define an empirical Bayes estimate of a proportion P by:

$$\hat{P}_{EBAYES} = \hat{\pi}\left(\frac{\hat{\theta}}{\hat{\theta}+N}\right) + \hat{P}\left(\frac{N}{\hat{\theta}+N}\right)$$

We note also that the estimated hyper parameters can be used to estimate the posterior distribution for P, from which, we can construct confidence intervals for likely range of values for P. The true value for the formula in the posterior distribution is then obtained by substituting the hyper parameter estimates computed in Table 1. The results are shown in Table 2.

The empirical Bayes analysis of proportions of NHIS subscribers, assuming an underlying β-binomial model, show that the percentage of male subscribers ranges from about 8.8-10.2%, while that of female subscribers ranges from about 25.2-27.1%. For the occupational group, the result reveals that the proportion of Civil servants, who subscribed to the NHIS scheme as well as that of professionals is about 20%. The proportion of the respondents who are aware of the benefits of subscribing to the NHIS scheme is between 24 and 26%. For employees earning <₦100,000 annum^{-1} the subscription rate is 7.6-10%. The respective subscription proportions for those earning about ₦240,000 annum^{-1} is 23.5% between 25.5%, for those earning between ₦500,000-1,000,000 annum^{-1}, it is 7.6%, for those earning between ₦1,000,000-2,400,000 annum^{-1}, it is between 6.4 and 9%, while for employees earning >₦2,400,000 annum^{-1} the subscription is between 37 and 43%.

The proportion of subscribers among families that have only one child is between 16 and 20% that of families with 2 children is between 28 and 32%, families with three children have a proportion between 19 and 22%. Families

Table 1: Estimated re-parameterised hyper parameters

Parameters	$\pi = \alpha/(\alpha+\beta)$	$\theta = (\alpha+\beta)$
Educational qualification	0.20485209	11.2565400
Marital status	0.25779154	2.4279943
Number of children	0.28525253	4.0299010
Number of children under 18 years	0.21184063	6.8713730
Income per annum	0.17929003	6.3233100
Occupation	0.20833330	4462.3500000
Awareness of NHIS	0.13710902	3.8839810
Gender	0.17838828	9.6135160

Table 2: Computed proportion of participation in NHIS

Parameters	X_i	N_i	P_i (observed)	P_{EB}	Bayesian (95% LCI)	95% CI (95% UCI)	P_{EB} rank
Educational qualification							
Secondary school certificate	819	3.627	0.225806	0.225742	0.212136	0.239348	16
OND/NCE	117	2.808	0.041667	0.042318	0.034872	0.049764	3
HND/B.A/B.Sc	936	6.084	0.153846	0.15394	0.144872	0.163009	10
M.A./M.Sc/MPHE	702	1.989	0.352941	0.352108	0.331117	0.373099	23
Ph.D	117	468	0.25	0.24894	0.209764	0.288115	20
Marital status							
Single	117	4.446	0.026316	0.026442	0.021726	0.031158	1
Married	2,457	9.945	0.247059	0.247061	0.238585	0.255538	18
Widow	117	234	0.5	0.497513	0.433449	0.561577	26
No. of children							
<1	234	1.287	0.181818	0.182141	0.161054	0.203228	12
2	702	2.340	0.3	0.299975	0.281407	0.318542	22
3	585	2.808	0.208333	0.208444	0.193419	0.223468	14
4	234	2.106	0.111111	0.111444	0.098004	0.124884	9
>6	585	936	0.625	0.623544	0.592504	0.654583	27
Number of children under 18 years							
<1	468	2.574	0.181818	0.181898	0.166995	0.196801	11
2	1,170	3.159	0.37037	0.370026	0.35319	0.386863	24
3	117	1.404	0.083333	0.083959	0.069453	0.098466	5
Income per annum							
<100,000	234	2.691	0.086957	0.087173	0.076515	0.097831	7
Between 240,000 and 500,000	1,638	6.669	0.245614	0.245551	0.235221	0.255881	17
Between 500,000 and 100,0000	234	2.691	0.086957	0.087173	0.076515	0.097831	7
Between 1000,000 and 2400,000	117	1.521	0.076923	0.077347	0.063921	0.090772	4
≥2400,000	468	1.170	0.4	0.398814	0.370756	0.426871	25
Occupation							
Civil servant	1,989	9.828	0.202381	0.20424	0.196269	0.21221	13
Professional	702	3.276	0.214286	0.210853	0.196885	0.224822	15
Awareness of NHIS							
Yes	2,574	10.413	0.247191	0.24715	0.238865	0.255435	19
No	117	4.329	0.027027	0.027126	0.022286	0.031965	2
Gender							
Male	702	7.371	0.095238	0.09526	0.088558	0.101962	8
Female	1,989	7.605	0.261538	0.261475	0.251598	0.271351	21

with 4 children have a subscription proportion between 10-12% while, the last category in this group, families with 6 or more children record a proportion of between 59-65%. When, a family has only one child who is under 18 years the proportion is between 17-20%. For a family with two of the children <18 years, the proportion is between 35-39%; for a family with 3 children the proportion is between 7-10%. The result for the level of education category shows that the proportion of subscribers among holders of general certificate of education is between 21.2-23.9% that for holders of OND is between 3-5% that for holders of HND or Bachelor of Science is between 14-16%; for 2nd degree, it is between 33 and 37%, while the proportion for holders of Ph.D is between 21 and 23.9%. The proportion of subscribers among single respondents is between 2 and 3%, while that of couples is between 24 and 26%.

Overall, the computed confidence interval at the 95% level of significance provides evidence of a very narrow width (maximum width: 0.05 or 5%), thereby justifying the choice of Bayes method of estimation for the analysis.

CONCLUSION

The overaching objective of the study is to update previous knowledge about subscription to the NHIS scheme. As expected, the Bayesian analysis has saliently brought out the structure of the subscription to the NHIS scheme and inevitably underscore the need to raise the level of awareness of the populace about the NHIS scheme. It has presented, an opportunity to craft a marketing plan targeted at specific subgroup. For instance, since the proportion of women subscribers is about three times the proportion of male subscribers, it would suggest that more awareness campaign need to be done on men than women. Also, since the least subscribers to the scheme as indicated in the 'rank' column of Table 2 are the single individuals, efforts should be made at creating more awareness of the benefit of the scheme for this group. Similarly, since there does not appear to be much difference in the proportion of subscribers between civil servants and professionals, uniform marketing programmes can be mounted for them.

The study also, opens a vista for further research. Specifically, it raises questions about the cause of low subscription rates among some categories. This deserve to be investigated further.

REFERENCES

Burk, T.E. and A.R. Ek, 1982. Application of empirical Bayes James Stein procedures to simultaneous estimation problems in forest inventory. Forest Sci., 28: 753-771. http://www.forestry.umn.edu/publications/staffpapers/staff papers10.pdf.

Johnson, D.H., 1989. An empirical Bayes approach to analyzing recurring animal surveys. Ecology, 70(4): 945-952. DOI: 10.2307/1941361. http://www.esajournals.org/doi/abs/10.2307/1941361.

Leonard, T. and J.S.J. Hsu, 1999. Bayesian Methods, Cambridge Series in Statistical and Probabilistic Mathematics. ISBN: 13:9780521004145, 10: 0521004144. http://www.cambridge.org/catalogue/catalogue-asp?isbn=9780521004145.

Link, W.A. and J.R. Sauer, 1995. Estimation and confidence intervals for empirical mixing distributions. Biometrics, 51 (3): 810-821. PMID: 07/06/200913:19. http://www.jstor.org/stable/2532983.

Link, W.A. and J.R. Sauer, 1996. Extremes in ecology: On avoiding the misleading effects of sampling variation in summary analysis. Ecology, 77: 1633-1640. DOI: 10.2307/2265557. http://www.jstor.org/stable/2265557.

Mingoti, S.A. and G. Meeden, 1992. Estimating the total number of distinct species using presence and absence data. Biometrics, 48 (3): 863-875. http://www.jstor.org/stable/2532351.

Morris, C.N., 1983. Parametric empirical Bayes inference: Theory and applications (with comments). J. Am. Stat. Assoc., 78(381): 47-65. PMID: 07/06/2009 14:06. http://www.jstor.org/stable/2287098.

Piegorsch, W.W., 1994. Empirical Bayes calculations of concordance between endpoints in environmental toxicity experiments. Environ. Ecol. Stat., 1: 153-164. DOI: 10.1007/Bf02426659. http://www.springerlink.com/content/u61n55u70m6703g.

Robbins, H., 1956. An empirical Bayes approach to statistics. Proc. 3rd Berkeley Symposium, 1: 157-163. http://www.biostat.jhsph.edu/~yonchen/10_Empirical_Bayes_Approach.pdf.

Rubin, D.B., 1983. Some applications of Bayesian statistics to educational data. J. Inst. Statistician, 32 (1 and 2): 55-68. http://www.jstor.org/stable/2987592.

Smith, P.J., 1991. Bayesian analysis for a multiple capture-recapture model. Biometrika, 78: 399-407. DOI: 10.1093/bionet/78.2.399. http://www.bionet.oxf-ordjournals.org/cg/content/abstract/78/2/399.

Stijnen, T. and J.C. van Houwelingen, 1990. Empirical Bayes methods in clinical trials meta-analysis. Biometrical. J., 32 (3): 335-346. DOI: 10.1002/bimj.4710320316.http://www.interscience-willey.com/journal/114076371/abstract?cretry=1&SRETRY=0.

Suggs, J.C. and T.C. Curran, 1983. An empirical Bayes method for comparing air pollution data to air quality standards. Atmospheric Environ., 17: 837-842. DOI:10.1016/0004-6981(83)90435-3.http://www.sciencedirect.com/science?-ob=B757C-4888D5C-R&.