Empirical Bayesian Analysis of the Poisson Intervention and Incidence Parameters

A. A. Bartolucci [a], **K. P. Singh**[b] **and R. Shanmugam**[c]

[a]*Department of Biostatistics, University of Alabama at Birmingham, Birmingham, Alabama, 35294 USA (Albartol@uab.edu)*

[b]*Department of Biostatistics and Epidemiology, University of North Texas ,Forth Worth, Texas, 76107 USA (Ksingh@hsc.unt.edu)*

[c]*Mathematics Department, University of Colorado, Denver, Colorado, 80217 USA (Rshanmugam@castle.cudenver.edu)*

Abstract: In some population the AIDS/HIV incidence rate 2 0 (0, 4) is altered in the middle of a data collection period due to preventive treatments imposed by the health service agencies. The intervened Poisson [IP] model is appropriate to analyze data of this type. However, the classical approach leading to the maximum likelihood [ML], moment [M] or minimum variance unbiased [MVU] estimator of 2 is mathematically formidable and practically inconvenient as far as sequentially updating the estimate when new data arrive. Previous subjective Bayesian work has been done to overcome these issues. Hence, there is a need to devise a more practical Empirical Bayesian technique to estimate 2, and it is done in this article. The results are illustrated using a data on AIDS/HIV incidence data from a state health department. Advantages in the Bayesian Intervened approach are cited.

Key Words: Bayesian; Empirical; Poisson; Intervened; Incidence

1. BACKGROUND AND MOTIVATION

As in Shanmugam et al. [1997] we consider a random sample $O_{(n)} = (O_1, O_2,...,O_n)$ on the number of AIDS/HIV cases that have occurred in n∃ 1 units or locations. An underlying model for $O_{(n)}$ is to be chosen depending on how the data collection apparatus is set up. Had a random sample of n units been chosen a priori, then the model for $O_{(n)}$ is Poisson (P):

$$p(x\#\ 2,P) = Pr(O = x\#\ 2, P) = e^{-2}\,2^x/x!$$
$$x = 0,1,2... ;2>0 \qquad (1)$$

If the data are collected from a list of units which are reported to have at least one case of AIDS/HIV incidence, then the model for $O_{(n)}$ is positive Poisson (PP):

$$p(x\#\ 2,PP) = Pr(O = x\#\ 2, PP) = (e^2 - 1)^{-1}2^x/x!,$$
$$x = 1,2 ... ;\ 2 > 0. \qquad (2)$$

because a zero event is unlikely.

In addition to the reality that O = 0 is not observable, suppose a medical intervention of some sort takes place. That is, due to the seriousness of the epidemic, the health officials may impose various preventive treatments such as educating the public about the ways of avoiding AIDS/HIV or enacting quarantine on the inflicted persons even in the middle of a data collection period, and these actions cause the incidence rate to be different from the time of such medical intervention. To study a health chance mechanism of this type, Shanmugam [1985] introduced a model, and named it intervened Poisson (IP) model. That is :

$$p(x*2IP) = P_r[O = x\#\ 2,IP] =$$
$$[e^{[\Delta+1]2}(1 - e^{-2})]^{-1}[(1 + \Delta)^z - \Delta^x]2^x/x! \qquad (3)$$

where $x = 1,2,....; 2 > 0, 0 \# \rho < 4$ is called the intervention parameter.

The PP model stated in (2) is obtained as a special case of (3) by substituting $\Delta = 0$. What does this special case imply? A zero value for Δ is indicative of completely successful preventive treatments [Shanmugam, 1985] , whereas $\Delta = 1$ is to be interpreted as a status quo in the incidence rate even after the preventive treatments are applied. Streit [1987] proposed a locally most powerful statistic to test a hypothesis $H_0 : \Delta = 1$ versus $H_1 : \Delta < 1$. Recently, Shanmugam [1992] derived a c(∀) test statistic as an alternate to Streit's statistic.

The discussions in this article pertain to estimating the incidence rate, 2 as this estimate is a basis for health officials in making future decisions on whether to further impose a stronger preventive action.

Unfortunately, the estimation of 2 is a formidable task [Shanmugam, 1985] as it involves solving a nonlinear estimating function :

$$\bar{\theta} = [\rho + 1 - (e^{\theta} - 1)] = \overline{X}$$

whether the ML or M method is applied, where $\bar{x} = \sum x_i / n$ denotes the sample average. On the contrary, the MVU method yields a simple. expression: 0=nS(n-1,n,-np)/S(n,n,-np) for estimating 2, but its difficulty lies in computing the generalized Stirling number S(·), as the number is known to be extremely large even for moderate values of its arguments (see Shanmugum [1985] for details on the generalized Stirling numbers).

Furthermore in the discussions leading to $\bar{\theta}$ or 0 the incidence rate is assumed to be stable and fixed although unknown. In an environment of AIDS/HIV epidemic, this assumption is trivially unreal. Rather, the incidence rate 2 should be treated as a random variable, although it is still fixed over the design period. This notion is Bayesian. See Zellner [1988] for the importance of Bayes concepts in many aspects of life.

The aim of this article is to chart out a Bayesian technique to estimate 2. Advantages of the Bayesian approach in estimating 2 are pointed out. Using the data on AIDS/HIV incidence in a particular state, the results are illustrated.

2. PRIOR AND POSTERIOR KNOWLEDGE OF θ

It is needless to state that in a Bayesian framework the prior distribution which quantifies the knowledge gained thus far about the parameter plays a crucial role. Different ideas have been entertained by statisticians to reach a consensus in making up a prior distribution (For details see Zellner 1971). Among others, the three major guiding criteria in making up the prior have been Data Dominance (DD), Invariance [I] and Conjugacy [C].

Under DD criterion, the data likelihood $l(x_{(n)}\# 2)$ ought to be dominant over the prior knowledge, and it is captured by the Jeffreys' [1961] non informational [vague] prior density:

$$J(2) = + [M^2{}_{22} \ln l (x (_{no})\# 2]^{-1/2} \qquad (4)$$

Where no, E, $\vartheta^2{}_{\theta\theta}$ and $l (x_{n0}|2)$ denote respectively the prior sample size, the conditional expectation for a fixed 2, the second order derivative with respect 2, and the conditional likelihood function (for a fixed 2). For the IP model stated in (3), note that:

$$l = l(x_{n_o} | \theta, IP) = [e^{[\rho-1]\theta} (1 - e^{\theta})^{-n_o}] * \qquad (5)$$
$$[(1-\rho)^x - \rho^x]^{n_o} \theta^{n_o \overline{x_o}} / \prod x_i!$$

Using (5) in (4) we obtain, after simplifications, the Jeffrey's non informational prior density,

$$J(\theta) = [\ \theta / n_o \partial_\theta \mu\ ]^{\frac{1}{2}} \qquad (6)$$

where $\mu = 2[p + 1 + (e^2 - 1)^{-1}]$ denotes the population mean. Under the conjugate criterion, both the prior and the conditional data likelihood $l(x_{(no)}|2)$ are to be "compatible" with each other. By compatible, we mean that under such a prior, the posterior will also be a member of the same distribution family. Such a prior is conjugate, and it is also versatile. In other words, a conjugate prior is a convenient building block in the Bayesian analysis. The natural choice for the conjugate prior of 2 in (3) is:

$$C(\theta) = n_o e^{n_o \theta} (n_o \theta)^{n\overline{x_o}-1} / \Gamma(n_o \overline{x_o}) \qquad (7)$$

where the hyper parameters n_0 and $\overline{x_o}$ denote respectively the prior sample size and mean. The

conjugate prior in (7) is versatile enough to accommodate both the exponential (with $n_0x_0 = 1$) and other skewed patterns of prior knowledge of the incidence rate 2. Unlike in the cases of vague priors, the conjugate prior contains the prior sample mean.

The posterior distribution $B(2|x_{(n)})$ of the incidence rate depends on which prior is employed. If $B(2)$ indicates the chosen prior then $B(2|x_{(n)}) = B(2)\, l(x_{(n)} \mid (2) / M[x_{(n)}]$ where the normalizer:

$$M[x_{(n)}] = \int_{\infty}^{\infty} \pi(\theta) l(x_{(n)} \mid \theta) d\theta$$

denotes the marginal distribution of the data $x_{(n)}$.

With the non informative prior in (4), the exact expression for the posterior density $J(2|x_{(n)})$ is too cumbersome to be of much use. Hence, we proceed to find a simpler but useful approximate posterior distribution as follows. Since $(1 - e^2)^{-n} = 1$ for large n, and the variance:

$$\sigma^2=\mu-e^{\theta}[\theta/(e^{\theta}-1)]^2$$

of the IP model is approximately equal to $(\Delta + 1)$ 2, we approximate the posterior density as:

$$J(\theta \mid x_{(n)} = [n(\rho+1)]^{\bar{nx}-[n_o-1]} e^{-n(\rho+1)\theta} \theta^{\bar{nx}} / \Gamma(\bar{nx}+1) \tag{8}$$

If we did as in the case of non informational prior using $[1-e^{-\theta}]^{-n}=1$ for large n, the approximate version of the posterior distribution becomes:

$$H(\theta \mid x_{(n)}) = [n(\rho+1)]^{\bar{nx}-[n_o-1]} e^{-n(\rho+1)\theta} \bullet$$
$$\theta^{\bar{nx}-n_o} / \Gamma(\bar{nx}-[n_o-1]). \tag{9}$$

When the prior sample size $n_o=0$, the posterior distribution in (9) is analogous to the posterior distribution in (8). Thus we will not consider (9) further as it is a special case. With the conjugate prior density $c(\theta)$ in (7), the exact expression for the posterior distribution becomes

$$c(\theta \mid x_{(n_o)}) = e^{-[n(\rho+1)]\theta} (1-e^{-\theta})^{-n} \bullet$$
$$\theta^{nx+n_ox_o-1} / N_c(\rho, n, \bar{x}, n_o, \bar{x_o})$$
for $\theta>0$ and $\rho>0$, (10)

where the normalizing constant is:

$$N_c(\rho, n, \bar{x}, n_o, \bar{x_o}).$$

By considering the posterior density of the form:

$$\pi(\theta \mid x_{(n)}, IP) = [n(\rho+1)+\alpha]^{\bar{nx}+\beta} e^{-n(\rho+1+\alpha)\theta} \bullet$$
$$\theta^{\bar{nx}+\beta-1} / \Gamma(\bar{nx}+\beta). \tag{11}$$

With some $\alpha>0$ and $\beta>0$ in a general situation one can address both the posterior conjugate distributions in (10) with $\alpha = n_o, \beta = n_o\bar{x}$ and with $\alpha=0$, $\beta=n_o$ in (12) below as special cases.

If the underlying model were to be Poisson in (1) then the posterior distribution of the incidence rate, θ, would have also been a gamma type of the form:

$$\pi(\theta \mid x_{(n)}, P) = [n+\alpha]^{\bar{nx}+\beta} e^{-(n+\alpha)\theta} \bullet$$
$$\theta^{\bar{nx}+\beta-1} / \Gamma(\bar{nx}+\beta) \tag{12}$$

with some suitable empirically derived values for the hyerparameters $\alpha>0$ and $\beta>0$.

3. EMPIRICAL APPROACH TO ESTIMATION

The empirical approach to estimating the hyperparameters as well as the intervention parameter, ρ, uses the marginal distribution of the data, M[x], defined earlier which in turn can be extended to a posterior predictive distribution of an observation, y, as follows:

For the Poisson distribution we have:

$$p(y \mid x, P) = \frac{\Gamma(\bar{nx}+\beta+y)}{y!\Gamma(\bar{nx}+\beta)} \left(\frac{n+\alpha}{n+\alpha+1}\right)^{\bar{nx}+\beta} \bullet$$
$$\left(\frac{1}{n+\alpha+1}\right)^{y}, \quad y = 0,1,2,..... \tag{13}$$

Likewise for the IP distribution we have:

$$p(y \mid x, IP) = \left[1 - \left\{\frac{n(\rho+1)+\alpha}{n(\rho+1)+\alpha+1}\right\}^{n\bar{x}+\beta}\right]^{-1} *$$

$$\left[\left\{\frac{n(\rho+1)+\alpha}{n(\rho+1)(n+1)+\alpha+1}\right\}\right]^{n\bar{x}+\beta} \frac{\Gamma(n\bar{x}+\beta+1)}{y!\Gamma(n\bar{x}+\beta)} *$$

$$\left[\left\{\frac{[\rho+1]}{(n+1)(\rho+1)+\alpha}\right\}^{y}\right] - \left[\left\{\frac{\rho}{(n+1)(\rho+1)+\alpha}\right\}^{y}\right]$$

$y = 0,1,2.....$ (14)

Note that empirically one can use (13) to get initial estimates of the hyperparamters and use these values in (14) to derive an estimate of the intervention parameter, ρ, and in turn use these values in (11) to derive a final estimate of the incidence parameter, θ.

4. ALTERNATIVE BAYES ESTIMATION OF θ

In Bayesian analysis the parameter is estimated such that it provides a minimum risk which is expected loss with respect to the posterior distribution. Various functional losses such as quadratic or non linear are often considered.

Irregardless of the loss function one can proceed as follows for estimation of the incidence parameter.

Using the posterior IP distribution we obtain the modal estimate,

$$\hat{\theta}_m(IP) = (n\bar{x} - \beta - 1)/[n(\rho+1)+\alpha] \qquad (15)$$

for the intervened Poisson data. The modal estimate of the incidence rate with the usual Poisson data is

$$\hat{\theta}_m(P) = \left[1 + \frac{n\rho}{n+\alpha}\right]\hat{\theta}_m(IP)$$

which yields the relation:

$$\hat{Q}_m(P) \geq \hat{Q}_m(IP) \qquad (16)$$

implying the incidence rate is over estimated when the regular Poisson data are used.

It is well known [Zellner, 1971] that the Bayes estimate , $\hat{\theta}_Q$, is the posterior mean, $E[\theta|X_{(n)}]$ under a squared error loss: $L_Q = (\theta - \hat{\theta}_Q)^2$ and the posterior variance ,$E[(\theta - \hat{\theta}_Q)^2|X_n]$ is the Bayes Risk, R(Q). With an intervened Poisson situation and posterior distribution in (11), we note that the estimate with respect to squared error loss is

$$\hat{\theta}_Q(IP) = E[\theta \mid X_{(n)}] = \left[\frac{n\bar{x}+\beta}{n(\rho+1)+\alpha}\right]$$

with the Bayes Risk

$(Q,IP) = \hat{\theta}_Q(IP)/[n(\rho+1)+\alpha$. Using a Poisson sample the Bayes estimate with respect to a squared error loss is:

$$\hat{\theta}_Q(P) = \left[1 + \frac{n\rho}{n+\alpha}\right]\hat{\theta}_Q(IP)$$

with a Bayes Risk

$R(Q,P) = [1+(n\rho/(n+\alpha))]^2 R(Q,IP)$.

Notice that the Bayes Risk in estimating the incidence rate using Poisson data is much more than the Bayes Risk in estimation using the IP data implying that the incidence rate is overestimated using Poisson data in comparison to using the intervened Poisson data.

5. ILLUSTRATION

In this section, the results are illustrated using data on quarterly incicdence of AIDS/HIV which were supplied by the AIDS/HIV surveillance office of a State Department of Public Health. Our sample consists of intravenous drug users. Our data are assumed to follow a Poisson model with parameter 2 portraying the incidence rate of AIDS/HIV in the state. Our data yields :

$\bar{X}$ = 1.21 with a sample size n = 29.

The hyper parameters are estimated using the gamma [conjugate] prior distribution of 2 and the marginal distributions in (13) and (14) The empirical estimates of the marginals are $\forall$=0.9 and $\exists$=1.55. The intervention parameter from (14) is derived as Δ=0.1 which indicates nearly successful intervention. The incidence parameter from the posterior likelihood in (11) is 2=1.110.

Table 1 summarizes the data results for the alternative Bayes estimation of 2 from section 4.0 as well as the Bayes Risk under the intervened and non intervened cases.

Table 1. Estimates of the incidence parameter and the Bayes Risks for the Non Intervened and Intervened Poisson Data

Statistic	Value
Q(P)	1.0882
Q(IP)	0.9920
R (Q,P)	0.0332
RQ,IP)	0.0302

6. DISCUSSION

Clearly from the empirical approach one can easily derive the values of the hyperparameters in the marginal Poisson case and use these values at least as initial estimates in the Intervened Poisson case to derive the estimate of the intervention parameter. One can then use the posterior likelihood of 2 to derive an estimate of the incidence parameter. All the functional forms are well behaved and convergence occurs in 4 to 5 iterations.

From Table 1 we see the advantage of the use of the non-zero intervened parameter. The estimates of 2 as well as the Bayes risks are reduced with the value of the intervened parameter , Δ. Our data example is a further refinement of the study of Shanmugam et al. [1997] which used the subjective Bayesian approach. Here we have the added advantage of the empirical approach using only the data at hand.
The approach has wide application to many situations requiring analysis of incidence rates following the Poisson model.

7. REFERENCES

Jeffreys , H. *The Theory of Probability*, London: Oxford University Press. 1961.

Shanmugam, R., An Intervened Poisson distribution and its medical applications, *Biometrics.* 41,1025-1029, 1985.

Shanmugam, R., An inferential procedure for the Poisson intervention parameter, *Biometrics* 48 , 559-565, 1992.

Shanmugan, R., A.A.Bartolucci, and K.P.Singh, .Bayesian analysis of intervened AIDS/HIV rate, Modelling and Simulation: Epidemiology, Health, and Medical Research. A.D., MacDonald and M. Mc Aleer (eds.), Modelling and Simulation Society of Australia, Inc. 3: 1051-1056, 1997.

Streit, F. Testing whether effective intervention takes place, *Biometrics, 43*, 999-1000, 1987.

Zellner, A. *An Introduction to the Bayesian Inference and Decision,* Wiley Press, New York, 1971.

Zellner, A. A Bayesian Era . *In Bayesian Statistics*, edited by J.M. Bernardo et. al. Vol 3, 509-516, 1988.