AN ANALYSIS OF INTERVENED AIDS/HIV RATE

Ramalingam Shanmugam, Mathematics Department, University of Colorado,
Alfred A. Bartolucci and Karan P. Singh, Biostatistics Department, University of Alabama at Birmingham

Abstract In some population the AIDS/HIV incidence rate $\theta \in (0, \infty)$ is altered in the middle of a data collection period due to preventive treatments imposed by the health service agencies. Shanmugam [1985] introduced the intervened Poisson [IP] model which 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 θ is mathematically formidable and practically inconvenient as far as sequentially updating the estimate when new data arrive. Hence, there is a need to devise a Bayesian technique to estimate θ, and it is done in this article. The results are illustrated using a data on AIDS/HIV incidence in the state of Alabama. Advantages in the Bayesian approach are cited.

1.1 BACKGROUND AND MOTIVATION:

Consider a random sample $X_{(n)} = (X_1, X_2,...,X_n)$ on the number of AIDS/HIV cases that have occurred in $n \geq 1$ houses. An underlying model for $X_{(n)}$ is to be chosen depending on how the data collection apparatus is set up. Had a random sample of n houses been chosen a priori, then the model for $X_{(n)}$ is Poisson (P):

$$p(x \mid \theta, P) = Pr(X = x \mid \theta, P) = e^{-\theta} \theta^x / x! \quad x = 0,1,2... ; \theta > 0 \tag{1}$$

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

$$p(x \mid \theta, PP) = Pr(X = x \mid \theta, PP) = (e^{\theta} - 1)^{-1} \theta^x / x!, \quad x = 1,2 ... ; \theta > 0. \tag{2}$$

because a zero event is unlikely.

In addition to the reality that X = 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

$$p(x \mid \theta IP) = P_r[X = x \mid \theta, IP] = [e^{[\rho+1]\theta}(1 - e^{-\theta})]^{-1}[(1 + \rho)^z - \rho^x]\theta^x / x! \tag{3}$$

where $x = 1,2,....; \theta > 0, 0 \leq \rho < \infty$ is called the intervention parameter.

The PP model stated in (2) is obtained as a special case of (3) by substituting $\rho = 0$. What does this special case imply? A zero value for ρ is indicative of completely successful preventive treatments (see Shanmugam [1985] for details), whereas $\rho = 1$ 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 $H_0 : \rho = 1$ versus $H_1 : \rho < 1$. Recently, Shanmugam [1992] derived a $c(\alpha)$ test statistic as an alternate to Streit's statistic.

The discussions in this article pertain to estimating the incidence rate, θ 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 (see Shanmugam [1985] for details) of θ is a formidable task, as it involves solving a nonlinear estimating function $\bar{\theta} [\rho + 1 + (e^{\theta} - 1)^{-1}] = \bar{x}$ whether the ML or M method is applied, where $\bar{x} = \Sigma x_i / n$ denotes the sample average.

On the contrary, the MVU method yields a simple. expression: $\bar{\bar{\theta}} = n\, S(n-1, n, -np)/S(n, n, -np)$ for estimating θ, but its difficulty lies in computing the generalized Stirling number $S(\cdot)$, as the number is known to be extremely large even for moderate values of its arguments (see Shanmugum [1984] for details on the generalized Stirling numbers).

Furthermore, in the discussions leading to $\bar{\theta}$ or $\bar{\bar{\theta}}$, 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 θ 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 θ. Advantages of the Bayesian approach in estimating θ are pointed out. Using the data on AIDS/HIV incidence in the state of Alabama, 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 tossed around by statisticians to reach a consensus in making up a prior distribution (see Zellner [1971] for details). 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 $l(x_{(n)} \mid \theta)$ ought to be dominant over the prior knowledge, and it is captured by the Jeffreys' [1961] non informational [vague] prior density:

$$J(\theta) = E\,[\partial^2{}_{\theta\theta}\, \ln l\,(x\,(_{no}\,)\mid \theta]^{-1/2} \qquad (4)$$

where n_0, E, $\partial^2{}_{\theta\theta}$ and $l\,(x_{n0}\mid\theta)$ denote respectively the prior sample size, the conditional expectation for a fixed θ, the second order derivative with respect θ, and the conditional likelihood function (for a fixed θ). For the IP model stated in (3), note that

$$l = l(x_{n_0})\mid\theta,\mathbb{IP}) = [e^{\,[p\cdot 1]\theta}(1 - e^{\,-\theta})]^{-n_0}[(1+p)^x - p^x]^{n_0}\theta^{\,n_0\bar{x}_0}/\underset{i\cdot 1}{\overset{n}{\pi}}(x_i!)$$

$$(5)$$

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

$$J(\theta) = [\theta/n_0 \partial_\theta \mu]^{\frac{1}{2}} \qquad (6)$$

where $\mu = \theta[p + 1 + (e^\theta - 1)^{-1}]$ denotes the population mean.

Under the conjugate criterion, both the prior and the conditional data likelihood $l(x_{(no)}\mid\theta)$ 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 θ in (3) is

$$C(\theta) = n_0 e^{\,-n_0\theta}(n_0\theta)^{n_0\bar{x}_0 - 1}/\Gamma(n_0\bar{x}_0), \qquad (7)$$

where the hyper parameters n_0 and $\bar{x}_0$ denote respectively the prior sample size and mean. The conjugate prior in (7) is versatile enough to accommodate both the exponential (with $n_0 x_0 = 1$) and other skewed patterns of prior knowledge of the incidence rate θ. Unlike in the cases of vague priors, the conjugate prior contains the prior sample mean.

The posterior distribution $\pi(\theta\mid x_{(n)})$ of the incidence rate depends on which prior is employed. If $\pi(\theta)$ indicates the chosen prior then $\pi(\theta\mid x_{(n)}) = \pi\,(\theta)\, l\,(x_{(n)} \mid (\theta) / M\,[x_{(n)}]$ 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 $J\,(\theta\mid x_{(n)})$ is too cumbersome to be of much use. Hence, we proceed to find a simpler but useful approximate posterior distribution as follows. Since $(1 - e^\theta)^{-n} \approx 1$ for large n, and the variance

$$\sigma^2 = \mu - e^{\,\theta}\left(\frac{\theta}{e^{\,\theta} - 1}\right)^2$$

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

$$J(\theta\mid x_{(n)}) \approx [n(p+1)]^{n\bar{x}+1} e^{\,-n(p-1)\theta}\theta^{n\bar{x}}/\Gamma(n\bar{x}+1). \qquad (8)$$

1052

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

$$H(\theta|x_{(n)}) \approx [n(\rho+1)]^{\overline{nx}-[n_0-1]} e^{-n(\rho+1)\theta} \theta^{\overline{nx}-n_0} \Gamma(\overline{nx}-[n_0-1]).$$

$$(9)$$

When the prior sample size $n_0=0$, the posterior distribution in (9) coincides with the posterior distribution in (8), and hence we will not pursue (9) for further discussion as an individual item since it is a special case. With the conjugate prior density $c(\theta)$ in (7), the exact expression for the posterior distribution is easily obtained, and it is

$$c(\theta|x_{(n)}) = e^{-[n(\rho+1)]\cdot n_0]\theta} (1-e^{-\theta})^{-n} \theta^{\overline{nx}+n_0\overline{x}_0-1} / N_c(\rho,n,\overline{x},n_0,\overline{x}_0),$$

$$(10)$$

where the normalizing constant is $N_c(\rho,n,\overline{x},n_0,\overline{x}_0)$.

By considering the posterior density of the form:

$$\pi(\theta|x_{(n)},IP) = [n(\rho+1)+\alpha]^{\overline{nx}+\beta} e^{-[n(\rho+1)+\alpha]\theta} \theta^{\overline{nx}+\beta-1} / \Gamma(\overline{nx}+\beta)$$

$$(11)$$

with some $\alpha \geq 0$ and $\beta \geq 0$ in a general set-up, we will be able to address both the posterior conjugate distributions in (10) with ($\alpha = n_0$, $\beta = n_0 \overline{x}_0$) and in (12) with ($\alpha = 0$, $\beta = n_0 - 1$) as a special cases of non informative prior.

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

$$\pi(\theta|x_{(n)},P) = (n+\alpha)^{\overline{nx}+\beta} e^{-(n+\alpha)\theta} \theta^{\overline{nx}+\beta-1} / \Gamma(\overline{nx}+\beta) \quad (12)$$

with some suitable values for the hyper parameters $\alpha \geq 0$ and $\beta \geq 0$.

3   Shift of Information About θ.

The literature contains different methods of measuring the information in a given data about the parameter. Some measures are parametric while others are nonparametric. See a recent book by Sengupta [1993] for a complete bibliography on this subject. Among non parametric types, Shannon's

information measure is popular, and it is defined by

$$S(\theta) = -E[\ln\pi(\theta)] = -\int_{-\infty}^{\infty} \pi(\theta)\ln\pi(\theta)d\theta \quad (13)$$

in which $\pi(\theta)$ portrays a probability density of the parameter θ. By substituting n = 0 in (11) and (12), we obtain the same prior density.

$$\pi(\theta) = \alpha^{\beta} e^{-\alpha\theta} \theta^{\beta-1} / \Gamma(\beta) \quad (14)$$

for the incidence rate θ whether the model for the data yet to be collected is Poisson or intervened Poisson type, and this makes sense as it pertains to the period before the data collection. Using (13) in (14), we notice that the prior information about the incidence rate θ is

$$S(\theta) = \ln[\Gamma(\beta)/\alpha] + (1-\beta)\partial_\beta\Gamma(\beta) + \beta, \quad (15)$$

before the data collection apparatus is set up.

With the data collection apparatus is as described for the IP model in (3), the Shannon's information for θ changes to

$$S(\theta|x_{(n)},IP) = \ln(\Gamma(\overline{nx}+\beta)/[n(\rho+1)+\alpha])$$
$$+[1-(\overline{nx}+\beta)]\partial_\beta\Gamma(\overline{x}+\beta)+(\overline{nx}+\beta) \quad (16)$$

in the aftermath of data collection. The gain in information for θ due to the IP data is then

$$S(\theta|x_{(n)},IP)-S(\theta) =$$
$$\ln[\frac{\Gamma(\overline{nx}+\beta)}{\Gamma(\beta)}] + \ln(1-\frac{n[\rho+1]}{n[\rho+1]+2} +$$
$$[1-(\overline{nx}+\beta)]\partial_\beta\Gamma(\overline{nx}+\beta)-[1-\beta]\partial_\beta\Gamma(\beta)+\overline{nx} \quad (17)$$

which due to approximations,

$$\ln\Gamma(z) \approx \ln\sqrt{(2\pi)} - z + (z-\frac{1}{2})\ln z \quad (18)$$

and

$$-\ln(1-z) \approx z \text{ for } |z| < 1, \quad (19)$$

simplifies to

$$S(\theta|x_{(n)},IP)-S(\theta) \approx \frac{\overline{nx}}{\overline{nx}+\beta} - \frac{n(\rho+1)}{n(\rho+1)+\alpha}. \quad (20)$$

It is easy to see that with the data collection as described for the Poisson model in (1), the gain in Shannon's information after the data collection would have been

$$S(\theta|x_{(n)},P)-S(\theta) \approx \frac{\overline{nx}}{\overline{nx}+\beta} - \frac{n}{n+\alpha}. \quad (21)$$

1053

Hence, we obtain the shift A (n, α, ρ) in Shannon's information for θ due to the intervention, and it is

$$A(n,\alpha,\rho)=S(\theta|x_{(n)},IP)-S(\theta|x_{(n)},P)=(\frac{\alpha}{n+\alpha})(\frac{n}{n[\rho+1]+\alpha}).$$
(22)

This amount is a value in the scale: [0. 1]. Notice that the shift is a function of the current sample size n, the hyper parameters α and the intervention parameter ρ. Because $0 \leq \rho < \infty$, we note that

$$0 < A(n,\alpha,\rho) \leq (\frac{\alpha}{n+\alpha})(\frac{n}{n+\alpha}).$$
(23)

Furthermore, the shift is monotically increasing to its upper bound as the intervention level, ρ is decreasing. Recall that ρ is smaller when the intervention efforts were more effective (Shanmugan [1985]). However, the shift in the Shannon's information between Poisson and intervened Poisson becomes insignificant for an extremely large sample size n.

4 Bayes Estimation of θ

In Bayesian analysis, the parameter is estimated such that it provides a minimum risk (which is an expected loss) with respect to the posterior distribution. Of course, there is not a consensus opinion on defining the loss, although the quadratic loss is popularly used.

While it is difficult to select a loss function, the mode of the posterior distribution could be chosen as an estimate of θ. Using the posterior distribution in (13), we obtain the modal estimate:

$$\hat{\theta}_m(IP)=(n\bar{x}+\beta-1)/[n(\rho+1)+\alpha]$$
(24)

with an intervened Poisson data. The modal estimate of the incidence rate with a usual Poisson data is $\hat{\theta}_m(P)=[1+\frac{n\rho}{n+\alpha}]\hat{\theta}(IP)$ using the posterior

distribution in (14). Notice that

$$\hat{\theta}_m(P) \geq \hat{\theta}(IP),$$
(25)

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

It is well known (see 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 its Bayes risk, R(Q). With an intervened Poisson data, and posterior distribution in (11), we note that the estimate with respect to a squared error loss is

$$\hat{\theta}_Q(IP) = E[\theta|x_{(n)}] = \left[\frac{n\bar{x}+\beta}{n(\rho+1)+\alpha}\right]$$
(26)

with the Bayes risk

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

$$\hat{\theta}_Q(IP) = \left[1 + \frac{n\rho}{n+\alpha}\right]\hat{\theta}_Q(IP)$$
(27)

with a Bayes risk

$$R(Q,P) = \left[1+\frac{n\rho}{n+\alpha}\right]^2 R(Q,IP)$$
(28)

Notice that the Bayes risk in estimating the incidence rate using Poisson data is much more than the Bayes risk in estimating it using IP data. Also, $\hat{\theta}_Q(P) \geq \hat{\theta}_Q(IP)$, implying that the incidence rate is overestimated using Poisson data in comparison to such estimate, using intervened Poisson data.

5 Illustration

In this section, the results are illustrated using data on quarterly incidence of AIDS/HIV in the state of Alabama which were supplied by the AIDS/HIV surveillance office of the Alabama Department of Public Health. Our sample consists of intravenous drug users. Our data are assumed to follow a Poisson model with parameter θ portraying the incidence rate of AIDS/HIV in the state. Our data yields $\bar{x} = 1.0689$ with a sample size n = 29.

The hyper parameters are estimated using the gamma [conjugate] prior distribution of θ and the posterior distribution in (12) is estimated. In Table 1, we assumed α = 0 for the non informative prior and α = 1 for the simple exponential. Although chosen for illustration purposes here, the hyperparameters can also be chosen by the method of Birch and Bartolucci (1983). Values of ρ varied from 0.0 for completely successful preventive treatments to ρ = 1 for the status quo in the incidence rate. The values of β maximizing the expected likelihood function of θ weighted by the prior density of θ varied from 1.53 to 1.58 for respective values of α and ρ.

In Table 2, the shift in Shannon's information is seen to monotonically increasing to it's upper bound of 0.03222. For $\alpha = 0$, according to (22) the shift is naturally zero. In Table 3, we have the results for the estimation of θ for the intervened Poisson data versus estimation using the usual Poisson data. The Bayes risk estimates are listed in Table 4. Clearly, for $\rho = 1$ and prior value of α, we see that the incidence rate is not over estimated using the intervened Poisson data. Again for $\rho = 0$, the results of the intervened and usual Poisson are equivalent. For the status quo or $\rho = 1$, the risk is reduced considerably for the intervened Poisson model. The Bayes application with the Poisson model clearly reinforces the previous work done with this data. However, in all cases the intervened Poisson model prevent an over inflated estimate of the incidence rate and allows for the reduced risk.

References

Birch, R., Bartolucci, A.A. Determination of the hyperparameters of a prior probability model in survival analysis. *Computer Programs in Biomedicine.* 17. 89 - 94. 1983.

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

Sengupta, J. K. Economics of Information and Efficiency, *New York: Kluwer Academic Press.* 1993.

Shanmugan, R. On central versus factorial momements, *South Africian Stat. J.*, 18, 97 - 110. 1984.

Shanmugan, R. An intervened Poisson distribution and its medical applications, *Biometrics, 48, 559 - 565.*

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

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

Zellner, A. Bayesian estimation and prediction usig asymmetric loss functions, *J. Amer. Stat. Assoc.*, 81, 446 - 451. 1986.

Zellner, A. (1988) A Bayesian Era, in Bayesian Statistics, *edited by J. M. Benardo et. al.*, Vol. 3, 509 - 516. 1988.

ρ	β	α
0.0	1.57	0.0
0.5	1.58	0.0
1.0	1.57	1.0
0.0	1.53	1.0
0.5	1.54	1.0

Table 1: Estimation of the Hyperparameters

ρ	α	R(Q,IP)	R(Q,P)
0	0	0.0375	0.0375
0	1	0.0350	0.0350
1	0	0.0094	0.0375
1	1	0.0091	0.0350

Table 4: Estimates of the Bayes Risk for the Intervened and Poisson Data

$\alpha = 1$

ρ	$A(n,\alpha,\rho)$
0.0	0.0322
0.5	0.0217
1.0	0.0164

Table 2: Shift in Shannon's Information

ρ	α	θ(IP)	θ(P)
0	0	1.0885	1.0885
0	1	1.0513	1.0513
1	0	0.5443	1.0886
1	1	0.0534	1.0516

Table 3: Estimates of θ for the intervened Poisson data versus the usual Poisson data