Poisson Versus Gamma Sampling for

Estimation in the Poisson Process

by

D. L. Solomon and R. R. Davidson

BU-440-M December, 1972

ABSTRACT

In observing a Poisson Process for the purpose of estimating the process parameter λ, two sampling schemes of interest are 1) Poisson sampling in which we observe the number of events occurring in a predetermined period of time t, and 2) Gamma sampling in which we observe the time required for a fixed number k of events to occur. This paper studies conditions under which each of the schemes is to be preferred.

We assume that prior information can be adequately represented by a member of the natural conjugate family and that the terminal loss associated with estimating λ by ℓ is proportional to $\lambda^a(\ell-\lambda)^2$. We also postulate sampling costs to be associated with each unit of time elapsed and each event observed.

We first determine for each of the two schemes the optimal (i.e., minimizes total Bayes risk) design parameter (t or k) and then compare optimal Poisson sampling with optimal Gamma sampling to characterize the set of sampling costs, loss functions and prior parameters for which Poisson (Gamma) sampling is preferred.

In a recent paper, G. M. El-Sayyad (J.R.S.S. Ser. B, 34, 1972) has independently obtained the same results as obtained here. We nevertheless chose to prepare this manuscript for several reasons. We feel that although the results are identical, the design optimality criterion used here is better motivated than that of El-Sayyad and represents a special case of a criterion which can be applied in the most general

(Bayesian decision - theoretic) problems of optimal design. Secondly, the technique of proof used here differs from El-Sayyad's and is more likely to be useful in other applications in that, as noted in this manuscript, El-Sayyad's proof, although more elegant than ours, turns on a point peculiar to this problem

Poisson Versus Gamma Sampling for

Estimation in the Poisson Process

by

D. L. Solomon and R. R. Davidson

BU-440-M December, 1972

SUMMARY

In observing a Poisson Process for the purpose of estimating the process parameter λ, two sampling schemes of interest are 1) Poisson sampling in which we observe the number of events occurring in a predetermined period of time t, and 2) Gamma sampling in which we observe the time required for a fixed number k of events to occur. This paper studies conditions under which each of the schemes is to be preferred.

We assume that prior information can be adequately represented by a member of the natural conjugate family and that the terminal loss associated with estimating λ by ℓ is proportional to $\lambda^a(\ell-\lambda)^2$. We also postulate sampling costs to be associated with each unit of time elapsed and each event observed.

We first determine for each of the two schemes the optimal (i.e. minimizes total Bayes risk) design parameter (t or k) and then compare optimal Poisson sampling with optimal Gamma sampling to characterize the set of sampling costs, loss functions and prior parameters for which Poisson (Gamma) sampling is preferred.

In a recent paper G. M. El-Sayyad [1972] has independently obtained the same results as obtained here. We nevertheless chose to prepare this manuscript for several reasons. We feel that although the results are identical, the design optimality criterion used here is better motivated than that of El-Sayyad and

represents a special case of a criterion which can be applied in the most general (Bayesian decision - theoretic) problems of optimal design. Secondly, the technique of proof used here differs from El-Sayyad's and is more likely to be useful in other applications in that, as noted in this manuscript, El-Sayyad's proof, although more elegant than ours, turns on a point peculiar to this problem.

1. INTRODUCTION

Suppose that events are occurring according to a Poisson process at an unknown rate λ, and that the purpose of observing the process is the estimation of λ. Either of the following two fixed (non-sequential) sampling schemes can be used.

P(t): Poisson sampling. Specify a time t and observe K_t, the number of events in time t. The distribution of K_t is Poisson with probability function

$$p(k|\lambda,t) = (\lambda t)^k e^{-\lambda t}/k! \qquad k = 0,1,2,\ldots$$

G(k): Gamma sampling. Specify a number of events k and observe T_k, the time until the k^{th} occurrence. The distribution of T_k is Gamma with probability density

$$g(t|\lambda,k) = \lambda^k t^{k-1} e^{-\lambda t}/\Gamma(k) \qquad t > 0 .$$

The object of this study is to compare the sampling schemes P(t) and G(k) for the purpose of establishing in advance of sampling which of the schemes is more appropriate. The approach taken will depend on the distribution used to represent the prior knowledge of the investigator, the structure of the loss incurred in the estimation of λ and on the costs of sampling.

2. THE ESTIMATION OF λ

Under either Poisson or Gamma sampling, the likelihood is proportional to $\lambda^k e^{-\lambda t}$, and (t,k) is a sufficient statistic. For Poisson sampling t is specified and k is the realization of K_t; for Gamma sampling k is specified and t is the realization of T_k. The maximum likelihood estimator $\hat{\lambda} = k/t$ is of course the same for either $P(t)$ or $G(k)$.

We now assume that the loss which results when we estimate λ by ℓ is of the form

$$L(\ell, \lambda) = \lambda^a (\ell - \lambda)^2 \tag{1}$$

where $\underline{a}$ is a specified constant. Furthermore, we assume that λ is the realization of a random variable Λ having probability density

$$t_o^{k_o} \lambda^{k_o - 1} e^{-t_o \lambda} / \Gamma(k_o) \qquad \lambda > 0,$$

that is Λ is distributed as Gamma (t_o, k_o). This family of distributions is conjugate (closed under combination with the likelihood) for either of the sampling schemes $P(t)$ or $G(k)$. The resulting posterior density for Λ is

$$f(\lambda | t,k) = (t_o + t)^{k_o + k} \lambda^{(k_o + k) - 1} e^{-(t_o + t)\lambda} / \Gamma(k_o + k) \qquad \lambda > 0,$$

that is, the conditional distribution of Λ given $K_t = k$ or given $T_k = t$ is Gamma $(t_o + t,\ k_o + k)$.

For the loss function (1), the Bayes estimator for λ is given by

$$\lambda^* = \frac{E(\Lambda^{a+1}|t,k)}{E(\Lambda^a|t,k)}$$

where expectation is taken with respect to the posterior distribution of Λ given either $K_t = k$ or $T_k = t$. The estimator λ^* is defined for each (t,k) if and only if $t_o > 0$ and $(a + k_o) > 0$, in which case

$$\lambda^* = (a + k_o + k)/(t_o + t) \ .$$

That λ^* is increasing in the loss parameter $\underline{a}$ is appropriate inasmuch as mis-estimation of large values of λ is costlier for large values of $\underline{a}$.

The posterior risk associated with the Bayes estimator λ^* is

$$R(t,k) \equiv E[L(\lambda^*,\Lambda)|t,k] = \frac{\Gamma(a + k_o + k + 1)}{\Gamma(k_o + k)(t_o + t)^{a+2}} \ .$$

Suppose now that there are specified costs of sampling; c_1 per unit time and c_2 per event. The cost of data having sufficient statistic (t,k) is then

$$C(t,k) = c_1 t + c_2 k \ .$$

The Bayes risk is found by taking the expected value of $R(t,k)$ with respect to the appropriate marginal distribution. For the sampling scheme $P(t)$, the marginal distribution of K_t is the negative binomial distribution with probability function

$$\frac{\Gamma(k_o + k)}{\Gamma(k_o)\Gamma(k + 1)} \frac{t_o^{k_o} t^k}{(t_o + t)^{k_o + k}}, \qquad k = 0,1,\ldots$$

and the corresponding Bayes risk is

$$r_P(t) = \frac{\Gamma(a + k_o + 1)}{\Gamma(k_o)t_o^{a + 1}(t_o + t)} .$$

For the sampling scheme G(k), the marginal distribution of T_k is the inverted beta distribution with probability density

$$\frac{\Gamma(k_o + k)}{\Gamma(k_o)\Gamma(k)} \frac{t_o^{k_o} t^{k-1}}{(t_o + t)^{k_o + k}}, \qquad t \geq 0$$

and the corresponding Bayes risk is

$$r_G(k) = \frac{\Gamma(a + k_o + 2)}{\Gamma(k_o)t_o^{a + 2}(a + k_o + k + 1)} .$$

In a similar way, the expected cost of sampling is found by taking the expected value of C(t,k) with respect to the appropriate marginal distribution. For the sampling scheme P(t), the expected cost of sampling is

$$c_P(t) = (c_1 + c_2 k_o/t_o)t \equiv c_P t$$

while for the sampling scheme $G(k)$, the expected cost is

$$c_G(k) = (c_1 t_o/(k_o - 1) + c_2)k \equiv c_G k$$

provided $k_o > 1$. Under Poisson sampling, the sampling costs (c_1, c_2) and $(c_P, 0)$ produce the same expected cost of sampling, while under Gamma sampling $(0, c_G)$ and (c_1, c_2) are equivalent sampling costs. Note for future reference that $c_P t_o = c_G(k_o - 1) + c_2$, so that

$$\frac{(k_o - 1)}{t_o} \leq \frac{c_P}{c_G} \leq \frac{k_o}{t_o} \tag{2}$$

with equality on the left when $c_2 = 0$ and equality on the right when $c_1 = 0$.

We now explore how the Bayes risks and the costs of sampling can be used to determine which of the two sampling schemes should be employed.

3. COMPARISON OF THE SAMPLING SCHEMES

When the sampling costs c_1 and c_2 are viewed as costs per unit loss, the Bayes risk and the expected cost of sampling can be added to give the total Bayes risk, and this can then be used as a measure of the total liability incurred in estimating λ. For the sampling scheme $P(t)$ the total Bayes risk is $b_P(t) = r_P(t) + c_P t$, while for the sampling scheme $G(k)$ the total Bayes risk is $b_G(k) = r_G(k) + c_G k$.

For each of the two schemes the optimal design parameter can be found by minimizing the total Bayes risk. This is a routine exercise in calculus. For Poisson sampling the optimal t^* is the value of t which minimizes $b_P(t)$, namely,

$$t^* = \max\left[0, \sqrt{\Gamma(a + k_o + 1)/\Gamma(k_o) c_P t_o^{a+1}} - t_o\right]$$

and the total Bayes risk associated with t^* is

$$b_P(t^*) = \begin{cases} 2\sqrt{\Gamma(a+k_o+1)c_P/\Gamma(k_o)t_o^{a+1}} - c_P t_o & \text{if } t^* > 0 \\ \Gamma(a+k_o+1)/\Gamma(k_o)t_o^{a+2} & \text{if } t^* = 0 \end{cases}.$$

For Gamma sampling the optimal k^* is the value of k which minimizes $b_G(k)$, namely

$$k^* = \max\left[0, \sqrt{\Gamma(a+k_o+2)/\Gamma(k_o)c_G t_o^{a+2}} - (a+k_o+1)\right]$$

and the total Bayes risk associated with k^* is

$$b_G(k^*) = \begin{cases} 2\sqrt{\Gamma(a+k_o+2)c_G/\Gamma(k_o)t_o^{a+2}} - (a+k_o+1)c_G & \\ & \text{if } k^* > 0 \\ \Gamma(a+k_o+1)/\Gamma(k_o)t_o^{a+2} & \text{if } k^* = 0 \end{cases}.$$

In this case, of course, the value of k^* need not be an integer, in which case one would use the integer value adjacent to k^* which gives the smaller total Bayes risk. It is clear from the above that there are situations in which non-sampling is optimal. This arises for certain configurations of the parameters (t_o, k_o) and a specified by the experimenter, and of the costs (c_1, c_2) of sampling. For instance, the parameters (t_o, k_o) may correspond to strong prior information while the costs (c_1, c_2) of sampling are high.

The total Bayes risk for the optimal sampling design of each scheme can now be used as the basis for selecting the better of the two sampling schemes. This selection can be made prior to any sampling. Specifically, we would perform Poisson sampling with $t = t^*$ if $b_P(t^*) < b_G(k^*)$, and would perform Gamma sampling with $k = k^*$ (or the appropriate nearest integer) if $b_P(t^*) > b_G(k^*)$. Let us now examine the way that this selection criterion depends on the parameters $\underline{a}$ and (t_o, k_o) and on the costs (c_1, c_2).

Consider the function

$$h(u) = \begin{cases} 2t_o\sqrt{u_o u} - t_o u & \text{if } u < u_o \\ u_o t_o & \text{if } u \geq u_o \end{cases}$$

where $u_o = \Gamma(a + k_o + 1)/\Gamma(k_o)t_o^{a+3}$. Note that for $u_P = c_P$, $h(u_P) = b_P(t^*)$ and that for $u_G = c_G(a + k_o + 1)/t_o$, $h(u_G) = b_G(k^*)$. When either u_P or u_G exceeds u_o the optimal sampling design of the corresponding scheme is no sampling. Moreover,

$$h'(u) = \begin{cases} t_o\sqrt{u_o/u} - t_o & \text{if } u < u_o \\ 0 & \text{if } u \geq u_o \end{cases}$$

so that $h(u)$ is strictly monotone in u for $u < u_o$. Thus provided that at least one of u_P and u_G is less than u_o, we have

$$b_P(t^*) \begin{pmatrix} < \\ = \\ > \end{pmatrix} b_G(k^*) \text{ if and only if } u_P \begin{pmatrix} < \\ = \\ > \end{pmatrix} u_G. \tag{3}$$

Now

$$\frac{u_P}{u_G} = \frac{c_P}{c_G} \cdot \frac{t_o}{(a + k_o + 1)}. \tag{4}$$

It then follows from (2) that

$$u_P/u_G \le k_o/(a + k_o + 1) < 1 \text{ when } a > -1$$

and

$$u_P/u_G \ge (k_o - 1)/(a + k_o + 1) > 1 \text{ when } a < -2 \quad .$$

Thus regardless of the costs of sampling and the parameters of the prior, Poisson sampling is preferred when $a > -1$ and Gamma sampling is preferred when $a < -2$. By substituting $a = -1$ and $a = -2$ into (4) and then appealing to (3) it can be seen that Poisson sampling is preferred for $a = -1$ and Gamma sampling for $a = -2$. Only when $a \in (-2, -1)$ is the selection of the better sampling scheme dependent on (t_o, k_o) and (c_1, c_2). In this event it follows from (3) and (4) that

$$b_P(t^*) \begin{pmatrix} < \\ = \\ > \end{pmatrix} b_G(k^*) \text{ if and only if } a \begin{pmatrix} > \\ = \\ < \end{pmatrix} A$$

where

$$A = \left(\frac{c_P}{c_G}\right) t_o - (k_o + 1) = -\left[1 + \frac{c_1 t_o}{c_1 t_o + c_2(k_o - 1)}\right] .$$

In El-Sayyad [1972], the above result is established using a different technique. In his equation (18), El-Sayyad puts the parameters of the two schemes into correspondence through the equation

$$\frac{t}{k} = \frac{t_o}{a + k_o + 1} \quad . \tag{5}$$

When $a < A$ we can improve on the sampling scheme $P(t)$ by using the Gamma scheme $G(k)$ where k is related to t through (5). Similarly, when $a > A$ we can improve on the sampling scheme $G(k)$ by using the Poisson scheme $P(t)$ where t is related to k through (5). It is then clear that the criterion for selecting the better scheme depends on (t_o, k_o) and (c_1, c_2) only through A and its relation to the loss parameter $\underset{\sim}{a}$.

El-Sayyad fails to provide any motivation for using the correspondence (5), and it appears that he has used this strictly as a technique of proof. It is interesting to note, however, that when t and k are related through (5), the sampling schemes P(t) and G(k) have the same Bayes risk, namely $r_P(t) = r_G(k)$. Moreover, the ratio of the expected costs of sampling depends on (t,k) only through (t/k), the ratio which is set fixed by (5). Thus the comparison of the expected costs of sampling schemes having the same Bayes risk depends only on the specified parameters $\underset{\sim}{a}$, (t_o, k_o) and (c_1, c_2). For any two sampling schemes having the same Bayes risk, it is clear that the scheme having the smaller expected cost of sampling would be preferred. In the present context this criterion leads to the same conclusion as the criterion of selecting the scheme with the smaller total Bayes risk.

It should be emphasized that the situation discussed in the preceding paragraph is specific to this problem and is not likely to arise in other contexts. Specifically, the comparison of expected sampling costs of schemes having the same Bayes risk will, in general still depend on the value of t or k as well as on the specified parameters. That this has not happened in the present context is strictly a coincidence. More importantly, the criterion of equating Bayes risks and comparing expected sampling costs is not a sensible approach in that in most instances neither of the schemes being compared is the optimal one.

REFERENCES

El-Sayyad, G. M. [1972]. A comparison of methods of sampling from a Poisson process. J. Royal Statist. Soc. Ser. B, 34, 427-430.