The Annals of Statistics
1980, Vol. 8, No. 1, 171–182

DYNKIN'S IDENTITY APPLIED TO BAYES SEQUENTIAL ESTIMATION OF A POISSON PROCESS RATE

BY C. P. SHAPIRO AND ROBERT L. WARDROP[1]

Michigan State University and University of Wisconsin-Madison

Conditional on the value of $\theta, \theta > 0$, let $X(t), t > 0$, be a homogeneous Poisson process. Sequential estimation procedures of the form $(\sigma, \hat{\theta}(\sigma))$ are considered. To measure loss due to estimation, a family of functions, indexed by p, is used: $L(\theta, \hat{\theta}) = \theta^{-p}(\theta - \hat{\theta})^2$, and the cost of sampling involves cost per arrival and cost per unit time. The notion of "monotone case" for total cost functions of a continuous time process is defined in terms of the characteristic operator of the process at the total cost function. The Bayes sequential procedure is then derived for those cost functions in the monotone case with optimality proven using extensions of Dynkin's identity for the characteristic operator. Finally, the sampling theory properties of these procedures are studied as sampling costs tend to zero.

1. Introduction. Conditional on the value of $\theta, \theta > 0$, suppose that $X(t), t \geqslant 0$, is homogeneous Poisson process. Sequential estimation procedures of the form $(\sigma, \hat{\theta}(\sigma))$ will be considered where σ is a stopping time with respect to $\{\mathcal{F}(t), t \geqslant 0\}$, with $\mathcal{F}(t)$ the sigma algebra of events generated by $\{X(s), 0 \leqslant s \leqslant t\}$, and $\hat{\theta}(\sigma)$ is an $\mathcal{F}(\sigma)$ measurable function, with $\mathcal{F}(\sigma)$ the sigma algebra of events prior to σ.

A family of functions, indexed by $p, 0 \leqslant p \leqslant 3$, will measure loss due to estimation:

$$L_p(\theta, \hat{\theta}) = \theta^{-p}(\theta - \hat{\theta})^2. \tag{1.1}$$

The effect of the p is to measure loss in units of θ^{-p}. The case of $p = 0$, squared error loss, has received the usual attention. Dvoretsky, Kiefer, and Wolfowitz (1953), and Hodges and Lehman (1951) have suggested $p = 1$. Using $p = 2$ causes the decision problem to be invariant under the group of scale transformations (Ferguson, 1967). The restriction of p to $[0, 3]$ is for mathematical tractibility, and the inclusion of noninteger p values usually increases the computational difficulty only slightly.

The cost of sampling will involve two components: c_A, the cost of observing one arrival, and c_T, the cost of observing the process for one unit of time.

In Section 2, the notion of "monotone case" for continuous time problems is defined in terms of the characteristic operator of $X(t)$ at the total cost function. Then with a gamma prior distribution on θ, the Bayes sequential procedure is

Received July 1977; revised June 1978.

[1]This research was sponsored, in part, by the United States Army under Contract No. DAAG29-75-C-0024.

AMS 1970 *subject classifications.* Primary 62L12, 62F15; secondary 62L15.

Key words and phrases. Dynkin's identity, characteristic operator, Poisson process, Bayes sequential estimation.

derived when the total cost is in the monotone case with optimality proven by using extensions of Dynkin's identity. As will be seen in the section, two familiar stopping times (types I and II censoring) are Bayes for appropriate choices of p and the sampling costs. The results obtained for $p = 1$ or 2 indicate that stopping rules suggested by El-Sayyad and Freeman (1973) are in fact optimal. Also, Novic (1977) has derived these stopping rules using a different technique.

In Section 3, the asymptotic properties of the Bayes sequential procedures are studied.

2. The Bayes sequential procedures. Assume that θ has prior density

$$\lambda(\theta;\, \alpha_0, \beta_0) = \Gamma(\alpha_0)^{-1}\beta_0^{\alpha_0}\,\theta^{\alpha_0 - 1}e^{-\beta_0\theta}$$

for $\theta > 0$, where $\beta_0 > 0$ and $\alpha_0 > p$, p the loss function index in (1.1). For $t > 0$ fixed, there exists a version of the posterior density of θ given $\mathcal{F}(t)$ of the form $\lambda(\theta;\, \alpha_0 + X(t), \beta_0 + t)$. It is easy to show that for all countably valued stopping times σ, the posterior density of θ given $\mathcal{F}(\sigma)$ has version $\lambda(\theta;\, \alpha_0 + X(\sigma), \beta_0 + \sigma)$. To extend this result to arbitrary stopping times, note that $X(t)$ is right continuous, and, for each θ, $\lambda(\theta;\, x, t)$ is continuous in t and x. Since an arbitrary stopping time can be approximated by a decreasing sequence of countably valued stopping times, the continuity above and the countability result imply that the posterior density of θ given $\mathcal{F}(\sigma)$ has version $\lambda(\theta;\, \alpha_\sigma, \beta_\sigma)$, with $\alpha_\sigma = \alpha_0 + X(\sigma)$ and $\beta_\sigma = \beta_0 + \sigma$, for an arbitrary stopping time σ.

The Bayes estimator of θ given $\mathcal{F}(\sigma)$ is

$$\hat{\theta}_p(\sigma) = (\alpha_\sigma - p)\beta_\sigma^{-1}, \tag{2.1}$$

and the posterior expected loss is

$$E\big[L_p(\theta, \hat{\theta}_p(\sigma))|\mathcal{F}(\sigma)\big) = \beta_\sigma^{p-2}\Gamma(\alpha_\sigma + 1 - p)\Gamma(\alpha_\sigma)^{-1}. \tag{2.2}$$

The total cost of observing the process for σ units of time is defined to be

$$\mathcal{C}_p(\sigma, X(\sigma)) = \beta_\sigma^{p-2}\Gamma(\alpha_\sigma + 1 - p)\Gamma(\alpha_\sigma)^{-1} + c_A X(\sigma) + c_T\sigma. \tag{2.3}$$

The Bayes sequential procedure (BSP) minimizes $E(\mathcal{C}_p(\sigma, X(\sigma)))$ over all stopping times σ.

Candidates for the BSP will be infinitesimal look-ahead rules derived from the characteristic operator of $X(t)$ at $\mathcal{C}_p(t, x)$. Such rules can be shown to be Bayes if the characteristic operator changes sign at most once and if the expected cost of using procedure $(\sigma, \hat{\theta}_p(\sigma))$ can be expressed in terms of the operator for a large class of stopping times σ.

Let $f(t, x)$ be measurable and finite on $[0, \infty)^2$. Define

$$D_f(t, x, h) = \frac{E\big[f(t + h, X(t + h))|(t, X(t)) = (t, x)\big] - f(t, x)}{h} \tag{2.4}$$

and consider only those functions f such that $D_f(t, x, h)$ is uniformly bounded in $h > 0$ and the limit of D_f exists as h decreases to zero. For such f define the characteristic operator at f as

$$Af(t, x) = \lim_{h\downarrow 0} D_f(t, x, h). \tag{2.5}$$

Of particular interest is the characteristic operator at $\mathcal{C}_p(t, x)$ which is given in the lemma below.

LEMMA 2.1. $A\ \mathcal{C}_p(t, x) = -\ \beta_t^{p-3}\Gamma(\alpha_0 + x - p + 1)\Gamma(\alpha_0 + x)^{-1} + c_A(\alpha_0 + x)\beta_t^{-1} + c_T$.

PROOF. See appendix.

For $t \geqslant 0$, define $B_t = B_t(p) = \{A\mathcal{C}_p(t, X(t)) \geqslant 0\}$. The cost $\mathcal{C}_p$ is in the *monotone case* if and only if, for all $t < s$, $B_t \subset B_s$ and $\lim_{t\to\infty}P(B_t) = 1$. This definition is modeled after the definition of the monotone case given by Chow, Robbins, and Siegmund (1971) for discrete time problems, and is interpreted in a similar way. Namely, if $A\mathcal{C}_p(t, x) > 0$, then the "infinitesimal" prospect for the future (proceeding from state (t, x)) is bad since by (2.4) the expected value of the incremental change in $\mathcal{C}_p$ is positive. If the cost sequence is in the monotone case, then once the infinitesimal prospect becomes bad, it remains bad. Thus, if $\mathcal{C}_p$ is well behaved, the infinitesimal look-ahead rule which stops the first time $A\mathcal{C}_p(t, X(t))$ is nonnegative, should be optimal. Before giving sufficient conditions for $\mathcal{C}_p$ to be in the monotone case, a result about the ratio of gamma functions is needed.

LEMMA 2.2. *Suppose* $w > 0$. *Then* $\Gamma(w - b)\Gamma(w)^{-1}$ *is increasing in* w *for* $b < 0$ *and decreasing in* w *for* $0 < b < w$.

PROOF. Take $b > 0$, and note that $\Gamma(w - b)\Gamma(w)^{-1}$ is decreasing in w since $\log \Gamma(w)$ is convex (Berk (1972), equation (2.1)).

LEMMA 2.3. *In each of the cases* (i)-(iii) *below*, $\mathcal{C}_p$ *is in the monotone case*:

(i) $c_T = 0, c_A > 0, 0 \leqslant p < 1$;
(ii) $1 \leqslant p \leqslant 2$;
(iii) $c_A = 0, c_T > 0, 2 < p \leqslant 3$.

PROOF. In each of the cases (i)–(iii), there exists functions K_{1j}, K_{2j} such that $A\mathcal{C}_p \geqslant 0$ if and only if $K_{1j}(\alpha_t, \beta_t) \geqslant K_{2j}(\alpha_t, \beta_t)$ where K_{1j} is nondecreasing in t and K_{2j} is nonincreasing in t.

Henceforth, only cases (i)–(iii) of Lemma 2.3 will be considered. For such p, c_A, and c_T, define the stopping time

$$\tau_p = \text{first } t \geqslant 0 \quad \text{such that } A\mathcal{C}_p(t, X(t)) \geqslant 0. \tag{2.6}$$

Another property of ratios of gamma functions is needed to derive bounds on τ_p and $X(\tau_p)$.

LEMMA 2.4. *If* $w > p \geqslant 0$, *then* $\Gamma(w - p)\Gamma(w)^{-1} \geqslant w^{-p}$.
If $0 \leqslant p < 1$ *and* $w > p + 1$, *then* $\Gamma(w - p)\Gamma(w)^{-1} \leqslant (w - (p + 1))^{-p}$.
If $p \geqslant 1$ *and* $w > p$, *then* $\Gamma(w - p)\Gamma(w)^{-1} \leqslant (w - p)^{-p}$.

PROOF. Let a be in $[0, \infty)$ and define $h(w, a) = (w - a)^p \Gamma(w - p)\Gamma(w)^{-1}$, for $w > a$. Then $h(w, a) \leqslant h(w + 1, a)$ if and only if

$$\log w + p \log(w - a) \leqslant p \log(w + 1 - a) + \log(w - p). \tag{2.7}$$

If $p \geqslant 1$, then using the concavity of logs, (2.7) holds with $a = p$; thus, $h(w, p) \leqslant h(w + 1, p)$. If $p < 1$, then (2.7) holds with $a = p + 1$, and hence $h(w, p + 1) \leqslant h(w + 1, p + 1)$. Noting that $h(w, a) \to 1$ as $w \to \infty$ for each a yields the upper bounds in the lemma. Choosing $a = 0$ reverses the inequality in (2.7) and gives the lower bound, with only $w > p$ required.

LEMMA 2.5. *For the cases of Lemma* 2.3, *the following bounds on* τ_p *and* $X(\tau_p)$ *hold*:

(i) $0 \leqslant p < 2$, $\tau_p \leqslant c_A^{1/(p-2)}$,
(ii) $1 \leqslant p < 3$, $\tau_p \leqslant c_T^{1/(p-3)}$,
(iii) $0 < p \leqslant 2$, $X(\tau_p) \leqslant c_A^{-1/p} \beta_0^{(p-2)/p} + 1$,
(iv) $1 < p \leqslant 3$, $X(\tau_p) \leqslant c_T^{-1/(p-1)} \beta_0^{(p-3)/(p-1)} + 1$,

where the bounds are infinite if the cost involved is zero.

PROOF. First, note that any stopping rule τ_p with both costs positive is bounded by the corresponding τ_p with only one cost positive. Thus, it suffices to bound the τ_p and $X(\tau_p)$ defined with one nonzero cost. For (i), $A\mathcal{C}_p < 0$ implies that

$$\begin{aligned} \beta_t^{2-p} &< c_A^{-1}\Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t + 1)^{-1} \\ &\leqslant c_A^{-1}. \end{aligned}$$

This implies the bound claimed. (ii) is proven in a similar fashion.

For (iii), $A\mathcal{C}_p < 0$ implies that

$$\begin{aligned} c_A\Gamma(\alpha_t + 1) &< \beta_t^{p-2}\Gamma(\alpha_t + 1 - p) \\ &\leqslant \beta_0^{p-2}\Gamma(\alpha_t + 1 - p). \end{aligned}$$

Next, if $p > 1$, then

$$\begin{aligned} c_A\beta_0^{2-p} &< \Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t + 1)^{-1} \\ &\leqslant (\alpha_t + 1 - p)^{-p} \qquad \text{by Lemma 2.4.} \end{aligned}$$

If $0 < p \leqslant 1$, then $c_A\beta_0^{(2-p)} < (\alpha_t - p)^{-p}$ by Lemma 2.4. Recalling that $\alpha_t = X(t) + \alpha_0$ yields (iii). Part (iv) is proven by a similar argument.

To deduce the optimality of rules τ_p, $E\mathcal{C}_p(\sigma, X(\sigma))$ will be expressed in terms of $A\mathcal{C}_p$. Referring back to $f(t, x)$ and $Af(t, x)$ defined earlier in the section, let σ denote a stopping time with $E\sigma < \infty$. Also note that $X(t)$ is marginally a strong Markov process. If f is bounded and $f(t, X(t))$ is a.s. right continuous in t, then the following identity due to Dynkin is well known:

$$Ef(\sigma, X(\sigma)) = E\textstyle\int_0^\sigma Af(t, X(t))dt + f(0, 0). \tag{2.8}$$

For details, see Athreya and Kurtz (1973). In the present situation $\mathcal{C}_p(t, x)$ is not a bounded function, and hence some extension of (2.8) which will include $\mathcal{C}_p(t, x)$ is needed.

LEMMA 2.6. *Suppose f is nonnegative and continuous in (t, x). Let σ be a stopping time such that $E\int_0^\sigma |Af(t, X(t))| dt < \infty$.*

(i) *If $EX(\sigma) < \infty$ and if f is nonincreasing in t and nondecreasing in x such that $Ef(0, X(\sigma)) < \infty$, then* (2.8) *holds.*

(ii) *If $E\sigma < \infty$ and if f is nondecreasing in t and nonincreasing in x such that $Ef(\sigma, 0) < \infty$, then* (2.8) *holds.*

PROOF. (i) Define the following sets of stopping times:

$$S_1^* = \{\sigma: EX(\sigma) < \infty, Ef(0, X(\sigma)) < \infty, E\int_0^\sigma |Af(t, X(t))| dt < \infty\},$$

$$S_2^* = \{\sigma: \sigma \text{ is in } S_1^* \text{ and } E\sigma < \infty\},$$

$$S_3^* = \{\sigma: \sigma \text{ is in } S_2^* \text{ and } X(\sigma) \text{ is bounded}\}.$$

First, (2.8), will be shown to hold for σ in S_3^*, then for σ in S_2^*, and finally for σ in S_1^*.

Let σ be in S_3^*. Then there exists $m < \infty$ such that $X(\sigma) < m$. Define $f_m(t, x) = f(t, x)$ on $[0, \infty) \times [0, m]$, and $= f(t, m)$ otherwise. Then (2.8) holds for f_m and σ. But $X(\sigma) < m$ implies that f_m can be replaced by f in the expression of (2.8). Thus, (2.8) holds for all σ in S_3^*.

Now take σ in S_2^*, and sequence $m_k \uparrow \infty$. Define stopping times $\sigma_k = \sigma$ if $X(\sigma) < m_k$ and $= t_k$ if $X(\sigma) \geq m_k$, where t_k is the first time that $X(t) = m_k$. Then σ_k is in S_3^* and (2.8) holds for f and σ_k. However, $Ef(\sigma_k, X(\sigma_k)) = Ef(\sigma, X(\sigma)) [X(\sigma) < m_k] + Ef(t_k, X(t_k)) [X(\sigma) \geq m_k]$. The first term tends to $Ef(\sigma, X(\sigma))$ and the second term tends to 0 since $f(\sigma, X(\sigma))$ and $f(t_k, X(t_k))$ are both bounded by $f(0, X(\sigma))$, which is integrable, and $EX(\sigma) < \infty$. In like manner $E\int_0^{\sigma_k} Af(t, X(t)) dt = E\int_0^\sigma Af(t, X(t)) dt [X(\sigma) < m_k] + E\int_0^{t_k} Af(t, X(t)) dt [X(\sigma) \geq m_k]$, where the first term tends to $E\int_0^\sigma Af(t, X(t)) dt$ and the second term is bounded by $E\int_0^\sigma |Af| [X(\sigma) \geq m_k]$ which tends to 0.

Now take σ in S_1^*, and sequence $t_k \uparrow \infty$. The truncation is now done on σ. Define $\sigma_k = \min(\sigma, t_k)$. Then σ_k is in S_2^* for all k and (2.8) holds for σ_k and f. As before, $Ef(\sigma_k, X(\sigma_k)) = Ef(\sigma, X(\sigma)) [\sigma \leq t_k] + Ef(t_k, X(t_k)) [\sigma > t_k]$. Note that σ in S_1^* implies $EX(\sigma) < \infty$, and hence $P(\sigma < \infty) = 1$. Using this and arguments similar to those above imply that $Ef(\sigma_k, X(\sigma_k))$ tends to $Ef(\sigma, X(\sigma))$. Finally, using similar arguments, $E\int_0^{\sigma_k} Af(t, X(t)) dt$ can be shown to tend to $E\int_0^\sigma Af(t, X(t)) dt$.

(ii) The proof is parallel to that of (i) with the roles of $X(\sigma)$ and σ reversed.

The lemma below makes precise the roles of the characteristic operator and Dynkin's identity in proving the optimality of τ_p.

LEMMA 2.7. *Consider only those $\mathcal{C}_p$ in the monotone case, and define $S_p = \{\sigma$: either* (2.8) *holds with $f(t, x) = \mathcal{C}_p(t, x)$ or $E\mathcal{C}_p(\sigma, X(\sigma)) = \infty\}$. If τ_p is in S_p and $E\mathcal{C}_p(\tau_p, X(\tau_p)) < \infty$, then $E\mathcal{C}_p(\tau_p, X(\tau_p)) \leq E\mathcal{C}_p(\sigma, X(\sigma))$, for all σ in S_p.*

PROOF. Let σ be a stopping time such that (2.8) holds and $E\mathcal{C}_p(\sigma, X(\sigma))$ is finite, and recall the definition of τ_p. Then

$$E\mathcal{C}_p(\sigma, X(\sigma)) - E\mathcal{C}_p(\tau_p, X(\tau_p)) = E\int_0^\sigma A\mathcal{C}_p - E\int_0^{\tau_p} A\mathcal{C}_p$$
$$= E[\sigma \geqslant \tau_p]\int_{\tau_p}^\sigma A\mathcal{C}_p - E[\sigma < \tau_p]\int_\sigma^{\tau_p} A\mathcal{C}_p$$

which is nonnegative by definition of τ_p and the monotone property of $\mathcal{C}_p$.

THEOREM 2.1. *For the cases of Lemma* 2.3, τ_p *optimal in that* $E\mathcal{C}_p(\tau_p, X(\tau_p)) \leqslant E\mathcal{C}_p(\sigma, X(\sigma))$ *for all stopping times* σ.

PROOF. In each of the cases (i)–(iii) of Lemma 2.3, the class S_p defined in Lemma 2.7 will be shown to include all stopping times σ. In case (i), $\mathcal{C}_p(t, x)$ is decreasing in t and increasing in x. Suppose σ is a stopping time with $EX(\sigma) < \infty$ and note that if $EX(\sigma) = \infty$ then the expected cost is infinite. Next, $\mathcal{C}_p(0, x)$ is bounded above by $\beta_0^{p-2}(\alpha_0 + x + 1 - p)^{1-p} + c_A x$ using the lower bound in Lemma 2.4. Thus, $EX(\sigma) < \infty$ implies that $E\mathcal{C}_p(0, X(\sigma)) < \infty$.

To show that $E\int_0^\sigma |A\mathcal{C}_p|dt < \infty$, consider a sequence of integers $m_k \uparrow \infty$, and define $\sigma_k = \sigma$ if $X(\sigma) < m_k$ and $\sigma_k = t_k$ if $X(\sigma) \geqslant m_k$, where t_k is the first time $X(t) = m_k$. Then it follows that $E(X(\sigma_k)) = E\int_0^{\sigma_k} \alpha_t \beta_t^{-1} dt$ (see Shapiro and Wardrop (1977), Lemma 2.4). By Fatou's lemma,

$$E\int_0^\sigma \alpha_t \beta_t^{-1} dt = E \liminf \int_0^{\sigma_k} \alpha_t \beta_t^{-1} dt \leqslant \liminf E(X(\sigma_k)) \leqslant E(X(\sigma)) < \infty.$$

But from the expression for $A\mathcal{C}_p$ in Lemma 2.1, $E\int_0^\sigma \alpha_t \beta_t^{-1} dt < \infty$ implies that $E\int_0^\sigma |A\mathcal{C}_p|dt < \infty$. Hence, Lemma 2.6 implies that S_p contains all σ and part (i) is completed by noting that $E(\mathcal{C}_p(\tau_p, X(\tau_p))) < \infty$.

For case (ii), first note that $E\mathcal{C}_p(\tau_p, X(\tau_p)) < \infty$. The cost is $\mathcal{C}_p(t, x) = (\beta_0 + t)^{p-2}\Gamma(\alpha_0 + x + 1 - p)\Gamma(\alpha_0 + x)^{-1} + c_A x + c_T t$, where the first term is decreasing in both t and x. Using techniques similar to those in Lemma 2.6, it can be shown that (2.8) holds with $f(t, x) = (\beta_0 + t)^{p-2}\Gamma(\alpha_0 + x + 1 - p)\Gamma(\alpha_0 + x)^{-1}$ and σ a stopping time such that $P(\sigma < \infty) = 1$. Then Lemma 2.6 can be applied directly to the last two terms of the cost depending on which costs are nonzero to yield S_p containing all σ.

For (iii), again note that $E\mathcal{C}_p(\tau_p, X(\tau_p)) < \infty$, and that the cost is now increasing in t and decreasing in x. Consider σ with $E\sigma < \infty$ since the expected cost is infinite if $E\sigma = \infty$. Then $E\mathcal{C}_p(\sigma, 0) < \infty$ since $E\sigma < \infty$ implies that $E\sigma^{p-2} < \infty$ for $2 < p \leqslant 3$. Also $|A\mathcal{C}_p|$ is bounded which implies that $E\int_0^\sigma |A\mathcal{C}_p| < \infty$, and Lemma 2.6 (ii) can be used to imply that S_p contains all σ.

If either $c_T = p = 0$, or $c_A = 0$ and $p = 1$, then stopping rule τ_p is type I censoring with $X(\tau_p)$ unbounded while τ_p is bounded. If either $c_T = 0$ and $p = 2$, or $c_A = 0$ and $p = 3$, then τ_p is type II censoring and τ_p is unbounded while $X(\tau_p)$ is bounded. For all other τ_p, both $X(\tau_p)$ and τ_p are bounded (this may be an important consideration in applications). Finally, the stopping times τ_p are (computationally) easy to use, especially for integer p.

3. Asymptotic properties of τ_p. In this section the Bayes sequential procedures $(\tau_p, \hat{\theta}(\tau_p))$, $0 \leqslant p \leqslant 3$, are examined from the sampling theory perspective. Attention is first restricted to the case when exactly one of the pair (c_A, c_T) is positive and the other is zero. The case of both costs positive is considered at the end of this section. The parameter θ is considered fixed but unknown and all probabilities and expectations are conditional on θ and denoted by P_θ and E_θ, respectively. Of interest are the asymptotic behaviors of the procedures.

Write $\tau(p, c_A, c_T, \alpha_0, \beta_0) = \tau_p$ to make explicit the dependence of τ_p on the various design parameters. It is easy to verify that, for $0 \leqslant p \leqslant 2$,

$$\tau(p, c, 0, \alpha_0, \beta_0) = \tau(p + 1, 0, c, \alpha_0 + 1, \beta_0). \tag{3.1}$$

In the results that follow, (3.1) will be useful in simplifying proofs.

For $c > 0$, and $0 \leqslant p \leqslant 3$, define

$$t_p^* = t_p^*(\theta, c_A, c_T) = (c_A\theta + c_T)^{-\frac{1}{2}}\theta^{(1-p)/2}. \tag{3.2}$$

LEMMA 3.1. (i) *Let $c_A = 0$, $c_T = c > 0$. For every $\varepsilon > 0$ and $1 \leqslant p \leqslant 3$,*

$$P_\theta\left[|\tau_p/t_p^* - 1| > \varepsilon\right] \leqslant k \exp\left[c^{-\frac{1}{2}}D(c, \varepsilon, \theta, p)\right],$$

for k a constant and $D(c, \varepsilon, \theta, p) \to D(\varepsilon, \theta, p) < 0$ and finite as $c \to 0$.

(ii) *Let $c_T = 0$, $c_A = c > 0$. For every $\varepsilon > 0$ and $0 \leqslant p \leqslant 2$,*

$$P_\theta\left[|\tau_p/t_p^* - 1| > \varepsilon\right] \leqslant k \exp\left[c^{-\frac{1}{2}}D(c, \varepsilon, \theta, p + 1)\right],$$

for k and D given in (i).

PROOF. (i) The result is easy to obtain for $p = 1$. For $p > 1$, write $\tau = \tau_p$, and $t^* = t_p^*$. Then $P_\theta[|\tau/t^* - 1| > \varepsilon]$ equals

$$P_\theta\left[\tau/t^* > 1 + \varepsilon\right] + P_\theta\left[\tau/t^* < 1 - \varepsilon\right]. \tag{3.3}$$

The first term in (3.3) equals $P_\theta[\tau > s]$ with $s = t^*b$ and $b = 1 + \varepsilon$. By Lemma 2.4, $P_\theta[\tau > s]$ does not exceed

$$P_\theta\left[X(s) < (2 - \alpha_0) + c^{1/(1-p)}\beta_s^{(p-3)/(p-1)}\right]. \tag{3.4}$$

By Bernstein's inequality for $0 < u \leqslant 1$, (3.4) is bounded above by

$$e^2 \exp\left[c^{-\frac{1}{2}}D^+(c, \varepsilon, \theta, p, u)\right]$$

with

$$D^+(c, \varepsilon, \theta, p, u) = ub\theta^{(3-p)/2}\left[(\beta_s/s)^{(p-3)/(p-1)}b^{2/(1-p)} + (e^{-u} - 1)u^{-1}\right].$$

Note that $(\beta_s/s)^{(p-3)/(p-1)}\uparrow 1$ as $c\downarrow 0$; $b^{2/(1-p)} < 1$; and $(e^{-u} - 1)u^{-1} \to -1$ as $u \to 0$. Thus, there exists $u_0^+ > 0$ such that for all u, $0 < u \leqslant u_0^+ D^+(c, \varepsilon, \theta, p, u) <$ 0 for all c.

The second term in (3.3) equals $P_\theta(\tau < s)$ with $s = t^* b$ and $b = 1 - \varepsilon$. A similar argument to the one above shows that $P_\theta(\tau < s)$ is bounded above by $e^{\alpha_0} \exp[c^{-\frac{1}{2}} D^-(c, \varepsilon, \theta, p, u)]$ with

$$D^-(c, \varepsilon, \theta, p, u) = ub\theta^{(3-p)/2}\left[(e^u - 1)u^{-1} - (\beta_s/s)^{(p-3)/(p-1)} b^{2/(1-p)}\right].$$

Thus, there exists u_0^- such that for all $u, 0 < u \leqslant u_0^-$, D^- is negative for all $c \leqslant c_0$ for some $c_0 > 0$. Let $u_0 = \min(u_0^+, u_0^-)$ and let $D(c, \varepsilon, \theta, p) = \max[D^+(c, \varepsilon, \theta, p, u_0), D^-(c, \varepsilon, \theta, p, u_0)]$.

(ii) may be proven by the above argument and the identity (3.1).

LEMMA 3.2. (i) *Let* $c_A = 0, c_T = c > 0$. *For* $1 \leqslant p < 3$, τ_p / t_p^* *is uniformly integrable as* $c \to 0$.

(ii) *Let* $c_T = 0, c_A = c > 0$. *For* $0 \leqslant p < 2$, τ_p / t_p^* *is uniformly integrable as* $c \to 0$.

PROOF. (i) Let $Y_c = |\tau_p / t_p^* - 1|$. It suffices to show that for fixed $a > 0$, $\int_{[Y_c > a]} Y_c dP_\theta$ tends to 0 as $c \to 0$. By Lemmas 2.5 and 3.1,

$$\int_{[Y_c > a]} Y_c \, dP_\theta \leqslant \theta^{(p-1)/2}\left[c^{-\frac{1}{2}}\right]^{(p-1)/(3-p)} k^* \exp\left[c^{-\frac{1}{2}} D(c, a, \theta, p)\right]$$

which tends to 0 as $c \to 0$. Also, (ii) can be proven with the same argument.

THEOREM 3.1. (i) *Let* $c_A = 0, c_T = c > 0$. *For* $1 \leqslant p \leqslant 3$, $\lim_{c \to 0} \tau_p / t_p^* = 1$ (*in* P_θ *probability*) *and* $\lim_{c \to 0} E_\theta(\tau_p / t_p^*) = 1$.

(ii) *Let* $c_T = 0, c_A = c > 0$. *For* $0 \leqslant p \leqslant 2$, $\lim_{c \to 0} \tau_p / t_p^* = 1$ (*in* P_θ *probability*) *and* $\lim_{c \to 0} E_\theta(\tau_p / t_p^*) = 1$.

PROOF. (i) The first limit follows immediately from Lemma 3.1. If $1 \leqslant p < 3$, then the second limit follows from Lemmas 3.1 and 3.2. If $p = 3$, then the definition of τ_3 implies that

$$c^{-\frac{1}{2}} - \alpha_0 \leqslant X(\tau_3) \leqslant c^{-\frac{1}{2}} + (3 - \alpha_0).$$

Thus,

$$t_3^* - \alpha_0 \theta^{-1} \leqslant \theta^{-1} X(\tau_3) \leqslant t_3^* + \theta^{-1}(3 - \alpha_0)$$

and,

$$1 - (t_3^*)^{-1} \alpha_0 \theta^{-1} \leqslant (t_3^*)^{-1} E_\theta \tau_3 \leqslant 1 + (t_3^*)^{-1} \theta^{-1}(3 - \alpha_0).$$

Noting that $t_3^* \to \infty$ as $c \to 0$ completes the proof. The proof of (ii) is similar.

Once the limiting form of τ_p is derived, the asymptotic normality of $\tau_p^{\frac{1}{2}}(\hat{\theta}_p(\tau_p) - \theta)$ follows from standard sequential methods. The theorem below gives the limiting form of the expected cost.

THEOREM 3.2. (i) *Let* $c_A = 0, c_T = c > 0$. *For* $1 \leqslant p \leqslant 3$,

$$c_T^{-\frac{1}{2}} \theta^{(p-1)/2} E_\theta \mathcal{C}_p(\tau_p, X(\tau_p)) \to 2 \quad \textit{as } c \to 0.$$

(ii) *Let* $c_T = 0$, $c_A = c > 0$. *For* $0 \leqslant p \leqslant 2$,

$$(\theta c_A)^{-\frac{1}{2}}\theta^{(p-1)/2}E_\theta\mathcal{C}_p(\tau_p, X(\tau_p)) \to 2 \quad as\ c \to 0.$$

PROOF. (i) Write $\tau = \tau_p$, $t^* = t_p^*$. From the definition of τ,

$$c \geqslant \beta_\tau^{p-3}\Gamma(\alpha_\tau + 1 - p)\Gamma(\alpha_\tau)^{-1} \tag{3.5}$$

which implies that $c^{-\frac{1}{2}}E_\theta\mathcal{C}_p \leqslant c^{\frac{1}{2}}E_\theta\beta_\tau + c^{\frac{1}{2}}E_\theta\tau$. This upper bound tends to $2\theta^{(1-p)/2}$ by Theorem 3.1. For a lower bound, let $\tau - \varepsilon$ be the time of the $X(\tau) -$ 1st arrival. Then the reverse inequality of (3.5) is satisfied at $\tau - \varepsilon$ and $\alpha_\tau - 1$ to yield

$$c < \beta_{\tau-\varepsilon}^{p-3}\Gamma(\alpha_\tau - p)\Gamma(\alpha_\tau - 1)^{-1}.$$

This implies that

(3.6)

$$c^{-\frac{1}{2}}E_\theta\mathcal{C}_p \geqslant Ec^{\frac{1}{2}}\beta_\tau\left[1 + \varepsilon(\beta_\tau - \varepsilon)^{-1}\right]^{p-3}\left[1 + (1-p)(\alpha_\tau - 1)^{-1}\right] + c^{\frac{1}{2}}E_\theta\tau.$$

The function inside the expectation of expression (3.6) is uniformly integrable and tends to $\theta^{(1-p)/2}$ in probability by Theorem 3.1 and the result that both τ and α_τ tend to ∞ as c tends to 0.

(ii) First note that $c^{\frac{1}{2}}E_\theta X(\tau) = c^{\frac{1}{2}}\theta E_\theta\tau$, (by Doob's optional stopping theorem), which tends to $\theta^{(2-p)/2}$ by Theorem 3.1. Now, from the definition of τ,

$$c\alpha_\tau \geqslant \beta_\tau^{p-2}\Gamma(\alpha_\tau + 1 - p)\Gamma(\alpha_\tau)^{-1}$$

which implies that

$$c^{-\frac{1}{2}}E_\theta\mathcal{C}_p \leqslant c^{\frac{1}{2}}E_\theta\alpha_\tau + c^{\frac{1}{2}}E_\theta X(\tau).$$

This upper bound tends to $2\theta^{(2-p)/2}$. Using the same techniques as in (i) yields

$$c^{-\frac{1}{2}}E_\theta\mathcal{C}_p \geqslant E_\theta c^{\frac{1}{2}}(\alpha_\tau - p)\left[\beta_\tau(\beta_\tau - \varepsilon)^{-1}\right]^{p-2} + c^{\frac{1}{2}}E_\theta X(\tau)$$

which tends to $2\theta^{(2-p)/2}$.

For the remainder of this section, assume that both costs are positive and write $\underline{c} = (c_A, c_T)$. Recall that $1 \leqslant p \leqslant 2$ is required for the monotone case. The limiting form of τ_p is easily obtained.

THEOREM 3.3. *If* $c_A, c_T \to 0$, *then*

$$\tau_p / t_p^* \to 1 \quad \text{a.s. } (P_\theta).$$

PROOF. The definition of τ_p is used to obtain upper and lower bounds on τ_p/t_p^*, both bounds tending to 1 a.s. (P_θ).

The theorem above does not give exponential rates which play a major role in deriving the limiting form of the expected cost. When $c_A, c_T \to 0$ such that $c_A c_T^{-1} \to c_0 < \infty$, such rates are obtained for $p = 1$ and 2 to yield the following result.

THEOREM 3.4. *Suppose* $p = 1$ *or* 2, *and that* c_A, $c_T \to 0$ *such that* $c_A c_T^{-1} \to c_0 \leqslant \infty$. *Then*

$$(c_A\theta + c_T)^{-\frac{1}{2}}\theta^{(p-1)/2}E_\theta \mathcal{C}_p(\tau_p, X(\tau_p)) \to 2.$$

PROOF. The techniques used in the proof of Lemma 3.1 yield the following exponential rates:

If $c_0 > 0$, then $P_\theta[|\tau_p/t_p^* - 1| > \varepsilon] \leqslant 2\exp[c_A^{-\frac{1}{2}}D]$;

if $c_0 = 0$, then $P_\theta[|\tau_p/t_p^* - 1| > \varepsilon] \leqslant 2\exp[c_T^{\frac{1}{2}}c_A^{-1}D]$,

where $D = D(c, \varepsilon, \theta, p) \to D(\varepsilon, \theta, p) < 0$ and finite. The result then follows using methods in the proof of Theorem 3.2. Details for $p = 1$ may be found in Shapiro and Wardrop (1977).

4. Concluding remark. In section 3, the Bayes sequential procedure τ_p is shown to be asymptotically equivalent to $t_p^* = t_p^*(\theta)$. This t_p^* can be motivated as follows. Let

$$L_t(\theta, \hat{\theta}) = \theta^{-p}(\theta - \hat{\theta})^2 + c_A X(t) + c_T t.$$

If sampling continues up to time t, and if estimator $\hat{\theta} = X(t)t^{-1}$ is used, then $E_\theta L_t(\theta, \hat{\theta}) = \theta^{1-p}t^{-1} + c_A\theta t + c_T t$ is minimized at $t = t_p^*$.

APPENDIX

PROOF OF LEMMA 2.1. Write

$$\mathcal{C}_p(t, x) = H_p(t, x) + c_A x + c_T t, \quad \text{where}$$
$$H_p(t, x) = \beta_t^{p-2}\Gamma(\alpha_0 + x + 1 - p)\Gamma(\alpha_0 + x)^{-1},$$

the cost due to estimation. Then, $A\mathcal{C}_p(t, x) = AH_p(t, x) + c_A(\alpha_0 + x)\beta_t^{-1} + c_T$, with the last two terms being easily computed since $Ax = (\alpha_0 + x)\beta_t^{-1}$. To derive $AH_p(t, x)$, write

$$\begin{aligned} E\big(H_p(t + h, X(t + h)) &- H_p(t, x)|(t, X(t)) = (t, x), \theta\big) \\ &= E\big(\beta_{t+h}^{p-2}\Gamma(\alpha_0 + x + Y + 1 - p)\Gamma(\alpha_0 + x + Y)^{-1} \\ &\quad - H_p(t, x)|(t, X(t)) = (t, x), \theta\big) \end{aligned}$$

where Y is Poisson (θh) given θ. Express this expectation as $r_{p,0}(\theta, h, t) + r_{p,1}(\theta, h, t) + R_{p,2}(\theta, h, t)$, with

$$r_{p,0}(\theta, h, t) = \big(\beta_{t+h}^{p-2} - \beta_t^{p-2}\big)\Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t)^{-1}e^{-\theta h},$$

$$\begin{aligned} r_{p,1}(\theta, h, t) = \big[&\beta_{t+h}^{p-2}\Gamma(\alpha_t + 2 - p)\Gamma(\alpha_t + 1)^{-1} \\ &- \beta_t^{p-2}\Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t)^{-1}\big]\theta h e^{-\theta h}, \end{aligned}$$

$$\begin{aligned} R_{p,2}(\theta, h, t) = \Sigma_{K=2}^{\infty}\big[&\beta_{t+h}^{p-2}\Gamma(\alpha_t + K + 1 - p)\Gamma(\alpha_t + K)^{-1} \\ &- \beta_t^{p-2}\Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t)^{-1}\big]\frac{(\theta h)^K}{K!}e^{-\theta h}. \end{aligned}$$

$E[R_{p,2}(\theta, h, t)|(t, X(t)) = (t, x)]$ will be shown to be $o(h)$. Define

$$d_K(t) = \Pi_{i=1}^{K}(\alpha_t - p + i)(\alpha_t - 1 + i)^{-1},$$

and note that $d_K(t) \leqslant 1$ for $p \geqslant 1$, and that $d_K(t) \leqslant (\alpha_t - p + 1)^K \alpha_t^{-K}$ for $0 \leqslant p < 1$. In terms of $d_K(t)$,

$$R_{p,2}(\theta, h, t) = \Sigma_{K=2}^{\infty}(\beta_{t+h}^{p-2} - \beta_t^{p-2})\Gamma(\alpha_t - p + 1)\Gamma(\alpha_t)^{-1} d_K(t) \frac{(\theta h)^K}{K!} e^{-\theta h}.$$

Thus, for $p \geqslant 1$,

$$|R_{p,2}(\theta, h, t)| \leqslant \Gamma(\alpha_t - p + 1)\Gamma(\alpha_t)^{-1}|\beta_{t+h}^{p-2} - \beta_t^{p-2}|e^{-\theta h}|e^{\theta h} - 1 - \theta h|$$

which is $o(h)$ as $h \downarrow 0$, and is also dominated by an integrable function of θ for h in a neighborhood of zero. Thus, $E(R_{p,2}(\theta, h, t)|(t, X(t)) = (t, x)) = o(h)$. Likewise, for $0 \leqslant p < 1$,

$$|R_{p,2}(\theta, h, t)| \leqslant \Gamma(\alpha_t - p + 1)\Gamma(\alpha_t)^{-1}|\beta_{t+h}^{p-2} - \beta_t^{p-2}|e^{-\theta h}\big[|e^{\theta h} - 1 - \theta h| + |e^{\theta h(\alpha_t - p + 1)\alpha_t^{-1}} - 1 - \theta h(\alpha_t - p + 1)\alpha_t^{-1}|\big]$$

which yields $E(R_{p,2}(\theta, h, t)|(t, X(t)) = (t, x)) = o(h)$.

Finally, as $h \downarrow 0$,

$$\frac{r_{p,0}(\theta, h, t)}{h} \to (p - 2)\beta_t^{p-3}\Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t)^{-1},$$

$$\frac{r_{p,1}(\theta, h, t)}{h} \to \beta_t^{p-2}\big(\Gamma(\alpha_t + 2 - p)\Gamma(\alpha_t + 1)^{-1} - \Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t)^{-1}\big)\theta.$$

Using $E[\theta|(t, X(t)) = (t, x)] = (\alpha_0 + x)\beta_t^{-1}$, and noting that the limit and expectation can be interchanged yields

$$h^{-1}E(r_{p,0} + r_{p,1}|(t, X(t)) = (t, x))$$

$$\to -\beta_t^{p-3}\Gamma(\alpha_t + 1 - p)\Gamma(\alpha_t)^{-1} \quad \text{as } h \downarrow 0$$

completing the proof.

Acknowledgments. Thanks to the referee and Associate Editor for their careful reading of our manuscript and also for several helpful suggestions. Section 4, and Theorem 3.3, in particular, appear because of the referees report.

REFERENCES

[1] Athreya, K. B. and Kurtz, T. G. (1973). A generalization of Dynkin's identity and some applications. *Ann. Probability* **4** 570–579.

[2] Berk, R. H. (1972). Consistency and asymptotic normality of MLE's for exponential models. *Ann. Math. Statist.* **43** 193–204.

[3] Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton-Mifflin, Boston.

[4] Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1953). Sequential decision problems for processes with continuous time parameter. Problems of estimation. *Ann. Math. Statist.* **24** 403–415.

[5] EL-SAYYAD, G. M. AND FREEMAN, P. R. (1973). Bayesian sequential estimation of a Poisson process rate. *Biometrika* **60** 289–296.

[6] FERGUSON, T. S. (1967). *Mathematical Statistics a Decision Theoretic Approach.* Academic Press, New York.

[7] HODGES, J. L. AND LEHMANN, E. L. (1951). Some applications of the Cramér-Rao inequality. *Proc. Second Berkeley Symp. Math. Statist. Prob.,* University of California Press.

[8] NOVIC, BRADLEY. (1977). Bayes sequential estimation of a Poisson rate. Technical Report #134, Department of Statistics, Carnegie-Mellon University, Pittsburg, Pa.

[9] SHAPIRO, C. P., AND WARDROP, R. L. (1977). The continuous time Bayes sequential procedure for estimating the arrival rate of a Poisson process and large sample properties. Technical Report No. 1753, Mathematics Research Center, University of Wisconsin, Madison.

DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
1210 W. DAYTON STREET
MADISON, WISCONSIN 53706