The Annals of Probability
1990, Vol. 18, No. 2, 877–886

CONDITIONS FOR QUASI-STATIONARITY OF THE BAYES RULE IN SELECTION PROBLEMS WITH AN UNKNOWN NUMBER OF RANKABLE OPTIONS

By F. Thomas Bruss and Stephen M. Samuels

University of California, Los Angeles, and Purdue University

In the so called *secretary problem*, if an unknown number N of options arrive at i.i.d. times with a known continuous distribution, then only the geometric, among proper distributions on N, has the property that the *stopping risk* depends just on the elapsed time and not on the number of arrivals so far. But even with such a prior, the *optimal rule* may, in general, depend on the number of arrivals so far. The optimal rule is closely related to the optimal policy in the Gianini and Samuels *infinite secretary problem*, except for a linear change in the time scale which depends only on the parameter of the prior, and not on the loss function.

1. Introduction. An unknown random number N of *options*, will arrive at times $Z_1, \dots, Z_N$, where $Z_1, Z_2, \dots$ are i.i.d. random variables with some known continuous distribution F on an interval $(0, T)$, possibly infinite. The options can be ranked from best (rank 1) to worst, and Z_i is the arrival time of the ith best. At any time $t \in (0, T)$ only the relative ranks of those options which have arrived so far can be observed. The object is to find a stopping rule τ based only on the observed relative ranks, which minimizes some risk function $Eq(R_\tau)$, where R_τ is the rank of the option selected by τ and $\{q(i)\colon i = 1, 2, \dots\}$ is a prescribed nondecreasing, nonnegative loss function. We need to allow τ's which, with positive probability, fail to select any option (it is convenient to set $\tau = T$ on this event) and to prescribe a loss $Q(N)$ for "not stopping." $Q(\cdot)$ should be nondecreasing, with $Q(0) = 0$.

This is the problem we considered in Bruss and Samuels (1987). In that paper we took $Q(N) = q(N)$ [but we could just as well have chosen $Q(N) = q(N+1)$] and showed that a single stopping rule, say τ^*, depending only on the loss function $q(\cdot)$ is nearly optimal for all stochastically large N. Specifically, letting $v^{(N)}$ denote the minimal risk, we have

$$v^{(N)} \le E^{(N)} q(R_{\tau^*}) \le v$$

and

$$N \uparrow \infty \text{ in distribution} \Rightarrow v^{(N)} \uparrow v,$$

where $v = v(q(\cdot))$ is the minimal risk and τ^* is the optimal rule in the infinite secretary problem of Gianini and Samuels (1976). (The result applies whenever v is finite.)

Received May 1988; revised April 1989.
AMS 1980 *subject classification*. 60G40.
Key words and phrases. Best choice problem, secretary problem, Bayes strategy, noninformative prior.

For loss functions that are eventually constant, as in Frank and Samuels (1980), where the goal is to maximize the probability of selecting one of the m best [i.e., $q(i) = 0$ for $1 \le i \le m$; $q(i) = 1$ for $i > m$], a more logical loss for not stopping is $Q(N) \equiv c$ (a constant). If $c \le \max q(\cdot)$, the above results must be modified slightly; they apply only to N's for which $P(q(N+1) = \max q(\cdot) | N > 0) = 1$. This, by the way, is guaranteed in the *best choice problem* $m = 1$. [However, conceivably one might want the loss for not stopping to be strictly greater than the loss for accepting some option, even the worst one. In this case, the results of Bruss and Samuels (1987) do not apply; hence this case will not be considered here.]

In Section 3 of Bruss and Samuels (1987) we looked at a Bayesian version of the problem in which N is given a prior distribution. In particular, we showed that when N has the improper, so called *noninformative* prior $P(N = n) = 1$ for all n, the posteriors are Pascal and the risk for stopping depends only on the relative rank of the current arrival and the elapsed time t, and not on the number of arrivals by time t. (Such cases we call *quasi-stationary*.) It is, in fact, precisely the stopping risk in the *infinite secretary problem* of Gianini and Samuels (1976). Moreover, in the special case of the best choice problem, results of Stewart (1981) imply that the formal Bayes *stopping rule* itself coincides with the optimal *infinite secretary problem* policy. Some questions raised by these results are the following:

1. Are there *proper* priors for which the same result holds for the stopping risk? If so, what are these priors?
2. For such priors is the optimal stopping rule also independent of the number of observed arrivals (i.e., does quasi-stationarity imply *stationarity*?) and, if so, does the optimal rule coincide with the optimal *infinite secretary problem* rule? In other words, does Stewart's 1981 result generalize beyond the case of the noninformative prior and the best-choice problem? [In Bruss and Samuels (1987), we overzealously remarked that it does.]

We show in Section 2.4 that the answer to question 1 is "yes," provided the priors are geometric. It is convenient to parameterize these priors by $p = \theta/(\theta + 1)$.

Let us mention that one well known way to get a geometric pair is for N to have a Poisson distribution with a random parameter which itself has an exponential distribution. If the exponential distribution has parameter θ, then the resulting geometric distribution will have parameter $p = \theta/(\theta + 1)$.

The posteriors are then Pascal with one parameter equal to $(\theta + t)/(\theta + 1)$. For $t > 0$ these make perfectly good sense even when $\theta = 0$; they are, in fact, the posteriors corresponding to the improper, noninformative prior. So the noninformative prior may be regarded as "included" among the geometric priors.

The answer to question 2 is not so simple and is somewhat surprising. In general, quasi-stationarity does NOT imply stationarity except for a few very special loss functions, such as Stewart's. So Stewart's result is quite anomalous. To answer question 2, we first need to obtain a general expression for

Bayes rules (in Section 2.1) and specialize it to the quasi-stationary (geometric prior) case (in Section 2.2). Then Theorem 2.1 says essentially, that it becomes optimal to disregard the number of arrivals so far when that number—call it k —is large enough so that (i) the loss for not stopping will not change, $Q(k) = \sup Q(\cdot)$, and (ii) the loss for stopping remains constant for all ranks greater than $k+1$, $q(k+2) = \sup q(\cdot)$.

For the problem of selecting one of the m best, we note that these conditions are satisfied for all $k \geq 1$ *only* for the best-choice problem $m = 1$ and for $m = 2$. An example at the end of Section 2.2 serves to illustrate the nature of the nonstationarity for $m = 3$. Bruss (1988; Theorem 3 and Application 3) sheds a different light and additional insight on why the cases $m = 1$ and $m = 2$ are special.

2. Results. Without loss of generality, we may take the arrival interval to be (0, 1) and the arrival-time distribution to be uniform. Let $T_1, T_2, \dots,$ be the arrival times, in chronological order,

$$N(t) = \text{number of arrivals by time } t,$$

$$R(t) = \text{rank of last arrival up to time } t \text{ among all } N,$$

$$r(t) = \text{relative rank of last arrival up to time } t \text{ among first } N(t) \text{ arrivals},$$

$$\rho_{j,k}(t) = \text{the stopping risk: } E(q(R(t))|r(t) = j, N(t) = k),$$

$$f_k(t) = \inf_{\tau > t} E(q(R(\tau))|N(t) = k),$$

where the inf is over all stopping rules τ which do not stop before time t.

2.1. *Partial solution by backward induction.* In the infinite secretary problem $[N(t) \equiv \infty]$ of Gianini and Samuels (1976) and Gianini (1977),

$$f(t) \equiv \inf_{\tau > t} E(q(R(\tau))|\text{the past up to time } t)$$

was shown to be constant (not depending on the past) for each t, satisfying the piecewise differential equation

$$f'(t) = t^{-1} \sum_{j=1}^{\infty} \left[f(t) - \rho_j(t) \right]^+, \qquad 0 \leq t < 1, \tag{1}$$

with boundary condition $f(1) = \sup q(\cdot)$, where

$$\rho_j(t) = \sum_{i=j}^{\infty} q(i) \binom{i-1}{j-1} t^j (1-t)^{i-j}. \tag{2}$$

The optimal policy has risk $f(0)$ and stops at the first arrival time, say x, for which $\rho_{r(x)}(x) \leq f(x)$, where $r(x)$ is the relative rank of the arrival at time x. This rule can be expressed in terms of so-called *cutoff points* $s_1 < s_2 < \dots,$ where $f(s_k) = \rho_k(s_k)$. Hence $f(0) = f(s_1) = \rho_1(s_1)$ and s_k is the earliest time at which an arrival of relative rank k is acceptable.

In contrast with the infinite secretary problem, the *unknown number of options problem*, which we consider in this paper, presents less technical difficulties, as one would expect, but the result is not nearly as neat, in general.

LEMMA 2.1. *For* $0 \le t < 1$,

$$f_k(t) = f_k(t+\delta) + o(\delta)$$
$$(3)\qquad - \sum_{n=k+1}^{\infty}\left([f_k(t) - f_{k+1}(t)] + \frac{1}{k+1}\sum_{j=1}^{k+1}[f_{k+1}(t) - \rho_{j,k+1}(t)]^+\right)$$
$$\times \frac{(n-k)\,\delta}{1-t} P(N = n|N(t) = k),$$

with boundary conditions $f_k(1) = Q(k)$.

PROOF. We first note the fact that the posterior distribution of N, at time t, depends only on $N(t)$, and not on the times or relative ranks of the arrivals up to time t, i.e.,

$$P\big(N = n|N(t), T_1, T_2, \dots, T_{N(t)}, r(T_1), r(T_2), \dots, r(T_{N(t)})\big)$$
$$= P(N = n|N(t)).$$

Indeed, this conditional independence of N and the arrival times and ranks, given $N(t)$, is an immediate consequence of the fact that, for every $n \ge k$, the conditional distribution of $T_1, T_2, \dots, T_k$, given $N = n$ and $N(t) = k$, is just the order statistics of k i.i.d. uniforms on $[0, t]$.

Now let $f_k(t)$ be defined above. Then, by standard backward induction,

$$f_k(t) = P(\text{no arrival in } [t, t+\delta]|N(t) = k)\cdot f_k(t+\delta)$$
$$(4)\qquad + E\big[\min\big(f_{k+1}(T_{k+1}), \rho_{r(T_{k+1}),k+1}(T_{k+1})\big)I_{\{t<T_{k+1}<t+\delta\}}|N(t) = k\big]$$
$$+ o(\delta).$$

The middle term on the right side of (4) is just

$$\min\big[f_{k+1}(t), \rho_{r(t),k+1}(t)\big]P(N(t+\delta) > k|N(t) = k) + o(\delta).$$

Conditioning on N and the uniformly distributed relative rank of the $k+1$st option yields (3) after some rearrangement. □

We now note that if there are k arrivals by time t, then the posterior distribution of N is

$$(5)\qquad P(N = n|N(t) = k) = \frac{\binom{n}{k}(1-t)^{n-k}t^kP(N = n)}{\sum_{r=k}^{\infty}\binom{r}{k}(1-t)^{r-k}t^kP(N = r)}.$$

The posterior distribution, at time t, of the actual rank of an option—which has relative rank, say j, and is one of, say k, arrivals by time t—clearly depends only on t, j and k, and not on the order of arrivals or the arrival times. Hence

$$(6)\quad \begin{aligned} &P(R(t)=i|N(t)=k, r(t)=j)\\ &\quad=\sum_{n=k}^{\infty} P(R(t)=i|N=n, N(t)=k, r(t)=j)P(N=n|N(t)=k), \end{aligned}$$

where

$$(7)\quad P(R(t)=i|N=n, N(t)=k, r(t)=j)=\frac{\binom{i-1}{j-1}\binom{n-i}{k-j}}{\binom{n}{k}}.$$

2.2. *Geometric prior*, *Pascal posterior*. If N is geometric with parameter $\theta/(\theta+1)$, then, from (5), $N+1$ has, conditionally, given $N(t)=k$, a Pascal distribution with parameters $k+1$ and $(\theta+\mathrm{t})/(\theta+1)$, and (6) also becomes Pascal, with parameters j and $(\theta+t)/(\theta+1)$. Notice that this latter distribution does not depend on k. When we substitute these Pascal distributions into (3), it simplifies to

$$(8)\quad \begin{aligned} \frac{f_k^{(\theta)}(t+\delta)-f_k^{(\theta)}(t)}{\delta/(\theta+1)}=\frac{\theta+1}{\theta+t}\Bigg((k+1)\big[f_k^{(\theta)}(t)-f_{k+1}^{(\theta)}(t)\big]\\ +\sum_{j=1}^{k+1}\bigg[f_{k+1}^{(\theta)}(t)-\rho_j\bigg(\frac{\theta+t}{\theta+1}\bigg)\bigg]^+\Bigg)+O(\delta),\\ 0\le t<1, \end{aligned}$$

where $\rho_j(\cdot)$ is given by (2). If we let $f_k^*((\theta+t)/(\theta+1))=f_k^{(\theta)}(t)$ and substitute into (8), we get

PROPOSITION 2.1.

$$(9)\quad \begin{aligned} (f_k^*)'(s)=s^{-1}\Bigg((k+1)\big[f_k^*(s)-f_{k+1}^*(s)\big]+\sum_{j=1}^{k+1}\big[f_{k+1}^*(s)-\rho_j(s)\big]^+\Bigg),\\ \theta/(\theta+1)\le s<1. \end{aligned}$$

The optimal policy stops at the first arrival time, say x, for which $\rho_{r(x)}(y)\le f^*_{N(x)}(y)$, where $r(x)$ is the relative rank of the arrival at time x, $N(x)$ is the number of arrivals by time x and $y=(\theta+x)/(\theta+1)$.

It is tempting to think that because the stopping risks ρ_j do not depend on k, the number of arrivals by time t, we ought to have $f_k^*(\cdot)=f_{k+1}^*(\cdot)$. If this were so, and if $Q(k)=Q(k+1)=q(k+2)=\sup q(\cdot)$, then (9) would reduce to precisely (1), restricted to the subinterval $[\theta/(\theta+1),1]$. [Note that if $q(k+2)=\sup q(\cdot)$, then, for all $j>k+1$, $\rho_j(t)=\sup q(\cdot)\ge f(t)$, so all but the first $k+1$ terms of the sum in (1) vanish.] But, as we shall see, equality does not hold in general, though it does hold in certain important special cases. The basic results are the following.

THEOREM 2.1. *For N geometric and any $s \in [0,1]$,*

$$Q(k) = \sup Q(\cdot) \quad \Rightarrow \quad f_{k+1}^*(s) \le f_k^*(s),$$
$$q(k+2) = \sup q(\cdot) \quad \Rightarrow \quad f_{k+1}^*(s) \ge f_k^*(s).$$

For the case $Q(n) \equiv q(n+1)$,

$$f_{k+1}^*(s) = f_k^*(s) \quad \Leftrightarrow \quad q(k+1) = \sup q(\cdot),$$

whereas for the case $Q(n) \equiv \sup q(\cdot)$,

$$f_{k+1}^*(s) = f_k^*(s) \quad \Leftrightarrow \quad q(k+2) = \sup q(\cdot).$$

The proof of this theorem is given in Section 2.3.

Applied to the problem of maximizing the problem of choosing one of the m best options, the theorem yields the following.

COROLLARY 2.1. *If $q(i) = 0$ for $i \le m$, $q(i) = 1$ for $i > m$ and $Q(n) \equiv 1$ $\forall\ n \ge 1$, then for any $\theta/(\theta+1) \le s \le 1$,*

$$f_k^*(s) = f_{k+1}^*(s) = f(s) \quad \text{if } k \ge m-1,$$
$$f_k^*(s) > f_{k+1}^*(s) \quad \text{if } k < m-1,$$

where $f(\cdot)$ is defined in (1).

Thus, for $m = 1$, the best choice problem, and for $m = 2$ as well, the Bayes rule coincides completely with the optimal infinite secretary problem rule, with time advanced from t to $(\theta + t)/(\theta + 1)$. Stewart's (1981) result, which was mentioned in the introduction, may be regarded as covering the case $m = 1$, $\theta = 0$. Notice also that, in the best choice problem, if $\theta > 1/(e-1)$, then the optimal rule simply selects the first arrival.

For $m \ge 3$, however, when there are fewer than $m - 1$ arrivals by time t, then the smaller k is, the greater $f_k^*(s)$ is; hence the greater the temptation to stop. We illustrate this phenomenon for the case $m = 3$.

LEMMA 2.2. *If $f_k^*(\cdot) = f_{k+1}^*(\cdot)$, then*

$$(f_{k-1}^*)'(s) = s^{-1}k[f_{k-1}^*(s) - f_k^*(s)] + (f_k^*)'(s) - s^{-1}[f_k^*(s) - \rho_{k+1}(s)]^+. \tag{10}$$

Let $h(s) = f_{k-1}^*(s) - f_k^*(s)$. Then

$$\frac{h'(s)}{s^k} - \frac{kh(s)}{s^{k+1}} = -\frac{[f_k^*(u) - \rho_{k+1}(s)]^+}{s^{k+1}}. \tag{11}$$

The solution is

$$h(1) - \frac{h(s)}{s^k} = -\int_s^1 \frac{[f_k^*(u) - \rho_{k+1}(u)]^+}{u^{k+1}}\,du. \tag{12}$$

[*Note that $h(1) = Q(k-1) - Q(k)$.*]

If, in addition, $Q(n) \equiv \sup q(\cdot)$ *and* $q(k+1) < q(k+2) = \sup q(\cdot)$*, then*

$$f_{k-1}^*(s) = f(s) + (s/s_{k+1})^k h(s_{k+1}), \qquad \forall s\colon \theta/(\theta+1) \le s \le s_{k+1}, \tag{13}$$

where $f(\cdot)$ *is the solution to* (1) *and* s_{k+1} *is the cutoff point such that* $f_k(s) > \rho_{k+1}(s)$ *if and only if* $s > s_{k+1}$.

PROOF. From (9), we have

$$(f_k^*)'(s) = s^{-1}\left([f_k^*(s) - \rho_{k+1}(s)]^+ + \sum_{j=1}^{k} [f_k^*(s) - \rho_j(s)]^+\right)$$

because $f_k^*(\cdot) = f_{k+1}^*(\cdot)$. Replace k by $k-1$ in (9) and use the above expression to substitute for the sum; this gives (10).

Rearrange (10) and divide by s^k to get (11). Notice that the left side is the derivative of $h(s)/s^k$; hence (12) follows immediately.

The additional conditions insure that $f_k^*(\cdot) = f(\cdot)$; existence of the cutoff point is then well known from Gianini and Samuels (1976). Hence the right side of (11) is zero for $\theta/(\theta+1) \le s \le s_{k+1}$ and (13) follows immediately. □

EXAMPLE. The above conditions are all satisfied in the *choose one of the m best* problem with $k = m - 1$. When $m = 3$, the stopping risks $\rho_1(s) = (1-s)^3$, $\rho_2(s) = (1-s)^2(1+2s)$ and $\rho_3(s) = 1 - s^3$. Hence $\rho_1(s) + \rho_2(s) + p_3(s) = 3(1-s)$, so, on $[s_3, 1]$,

$$f'(s) = s^{-1}[3f(s) - 3(1-s)],$$

$$f(s) = 1 - \tfrac{3}{2}s + \tfrac{3}{2}s^3.$$

Hence $s_3 = \sqrt{3/5}$ and

$$h(s) = -\tfrac{5}{2}s^3 - \tfrac{3}{2}s + 4s^2 = s(1-s)\left(\tfrac{5}{2}s - \tfrac{3}{2}\right), \qquad s_3 \le s \le 1.$$

The net effect of this is to advance the earliest effective time, $s = (\theta + t)/(\theta + 1)$, at which the optimal rule simply selects the first arrival, from s_1: $f(s_1) = \rho_1(s_1)$ to s_1^*: $f(s_1^*) + (s_1^*/s_3)^2 h(s_3) = \rho_1(s_1^*) = (1 - s_1^*)^3$.

We can solve for s_1^* as follows: First, note that $f(s_1^*) = f(0)$ which was evaluated in Frank and Samuels (1980) to be 0.2918. Next, substituting $s_3 = \sqrt{3/5}$ into the formula for $h(s)$ above, we get $h(s_3) = 0.0762$. Substituting these values yields $s_1^* = 0.3266$, *which should be compared with* $s_1 = 0.3367$, as given in Frank and Samuels (1980).

As the above example illustrates, for loss functions that are eventually constant, there is a starting point for calculating the f_k^*'s. For other loss functions, if the infinite problem optimal risk is finite [see Gianini and Samuels (1976) for sufficient conditions], then a sequence of truncations will eventually lead to the solution—at least in principle.

2.3. *Proof of Theorem* 2.1. The essence of the proof is a version of the censoring argument used in Bruss and Samuels (1987).

Fix t and $k + 1 = N(t)$. For our family of Pascal posteriors, those arrivals after time t which are better than the worst arrival before time t have the same distribution as would *all* of the post-t arrivals if $N(t)$ were k. Hence to every stopping rule, $\tau > t$, which can be used when $N(t) = k$, there is a corresponding rule, say τ', which, if used when $N(t) = k + 1$, does not select any option worse than the worst pre-t arrival and has the same risk as τ used when $N(t) = k$, except that τ' may incur a higher loss for not stopping, because $Q(n)$ is nondecreasing in n. But if $Q(k) = Q(k+1) = \dots$, then the losses for not stopping are certain to be equal.

On the other hand, if $q(k+2) = q(k+3) = \dots$, then an optimal rule, when $N(t) = k + 1$, will never stop with an option worse than the worst pre-t arrival. Hence this optimal rule could also be used when $N(t) = k$.

Sufficiency, in the two special cases, is immediate. To prove necessity, we first note that if $q(k+2) < \sup q(\cdot)$, there is positive probability that it will be better to accept an option worse than the worst pre-t arrival than not to. For example: Say $q(k+2) < q(l)$. Suppose the $k+1$ best options all arrive by time t and the next $l - k - 2$ best all arrive after the lth best, which itself (having relative rank $k + 2$) is preceded by at least $l - k - 3$ worse options and arrives so close to time $T = 1$ that it cannot be turned down.

The fact that this scenario has positive probability guarantees that, in both cases, $f_k^*(t) \neq f_{k+1}^*(t)$ if $q(k+2) < \sup q(\cdot)$. To complete the necessity argument we need only consider the case $Q(n) \equiv q(n+1)$ when $q(k+1) < q(k+2) = \sup q(\cdot)$. Here the optimal policy is the same whether $N(t) = k$ or $k + 1$, and the loss is the same unless there are no post-t, arrivals better than the worst pre-t arrival. There may, of course, be no post-t arrivals at all; hence

$$\begin{aligned} f_{k+1}^*(s) - f_k^*(s) &> [Q(k+1) - Q(k)] P_\theta(N = k+1 \mid N(t) = k+1) \\ &= [q(k+2) - q(k+1)] s^{k+1}, \end{aligned}$$

where $s = (\theta + t)/(\theta + 1)$.

2.4. *Necessary and sufficient conditions for quasi-stationarity.*

THEOREM 2.2. *The modified geometric prior distribution on N is the only prior for which, for all loss functions, the stopping risks $\rho_{j,k}(t)$ do not depend on k.*

PROOF. Take $i = j = 1$ and substitute (5) and (7) into the right side of (6); call it $g(t)$, since it does not depend on k. Then

$$(14) \quad g(t) \equiv \frac{\sum_{n=k}^{\infty} \binom{n-1}{k-1} (1-t)^{n-k} t^k P(N = n)}{\sum_{r=k}^{\infty} \binom{r}{k} (1-t)^{r-k} t^k P(N = r)}, \qquad k = 1, 2, \dots .$$

If we now let

$$h(t) = (1 - g(t))/(1 - t),$$

then (14) becomes, after some rewriting,

$$h(t) \equiv \frac{\sum_{n=k+1}^{\infty}(n-1)_k(1-t)^{n-k}P(N=n)}{\sum_{n=k}^{\infty}(n)_k(1-t)^{n-k+1}P(N=n)}, \tag{15}$$

which, after further rearrangement, becomes

$$\frac{h(t)}{P(N>0)} \equiv \frac{\sum_{n=k}^{\infty}(n)_k(1-t)^{n-k}[P(N-1=n)/P(N>0)]}{\sum_{n=k}^{\infty}(n)_k(1-t)^{n-k}P(N=n)}. \tag{16}$$

The denominator and numerator of the right side of (16) are the kth derivatives of the probability generating functions of N and of $(N-1)|\{N>0\}$, respectively. So, letting $t \to 1$, we see that $P(N=n+1)/P(N=n)$ is a constant for all $n = 1, 2, \ldots,$ so N must have a *modified* geometric distribution [i.e., $P(N=0)$ arbitrary; $(N-1)|\{N>0\}$ geometric]. □

We can also show that, if the posterior distribution of N is Pascal, i.e., if (5) reduces to

$$\frac{\binom{n}{k}(1-t)^{n-k}t^kP(N=n)}{\sum_{r=k}^{\infty}\binom{r}{k}(1-t)^{r-k}t^kP(N=r)} = \binom{n}{k}\varphi_t^{k+1}(1-\varphi_t)^{n-k} \tag{17}$$

for each $t \in (0,1)$ and for $k = 0, 1, \ldots, n$, then the prior distribution must be geometric. Indeed (17) is equivalent to

$$\sum_{r=k}^{\infty}\binom{r}{k}(1-t)^{r-k}t^kP(N=r) = \left(\frac{t}{\varphi_t}\right)^k\left(\frac{1-t}{1-\varphi_t}\right)^{n-k}\frac{1}{\varphi_t}P(N=n). \tag{18}$$

For fixed t, the right side of (18) is manifestly a constant function of n for $n = k, k+1, \ldots,$ so

$$P(N=k+m) = \left(\frac{1-\varphi_t}{1-t}\right)^m P(N=k) \tag{19}$$

for each $k \geq 0$ and $m \geq 1$. But the prior distribution of N does not depend on t, so $(1-\varphi_t)/(1-t)$ must be a constant and N has a geometric distribution. Specifically, $(1-\varphi_t)/(1-t) = 1 - P(N=0)$, so, if $P(N=0) = \theta/(\theta+1)$, then $\varphi_t = (\theta+t)/(\theta+1)$.

2.5. *An open problem.* We have seen that quasi-stationarity occurs only for geometric priors and that quasi-stationarity does not imply stationarity. What about the converse? We would be quite surprised if stationarity did not imply quasi-stationarity, but we have no proof.

It is our feeling that the lack of stationarity which we have demonstrated here serves to underscore the importance of the minimax results in Bruss (1984, Table 1) and Bruss and Samuels (1987).

Acknowledgments. This research was supported in part by grant ST2J-0277-C from the European Communities for the Study of Automatic Decision Strategies. Part of the research for this paper was carried out while the first author was Visiting Associate Professor at the University of California at Santa Barbara and part at the University of Arizona and at Purdue University, while holding a research fellowship from the Alexander von-Humboldt Foundation. He would like to thank the foundation for its support and these institutions for their hospitality.

REFERENCES

BRUSS, F. T. (1984). A unified approach to a class of best choice problems with an unknown number of options. *Ann. Probab.* **12** 882–889.

BRUSS, F. T. (1988). Invariant record processes and applications to best choice modelling. *Stochastic Process. Appl.* **30** 303–316.

BRUSS, F. T. and SAMUELS, S. M. (1987). A unified approach to a class of optimal selection problems with an unknown number of options. *Ann. Probab.* **15** 824–830.

FRANK, A. Q. and SAMUELS, S. M. (1980). On an optimal stopping problem of Gusein–Zade. *Stochastic Process. Appl.* **10** 299–311.

GIANINI, J. (1977). The infinite secretary problem as the limit of the finite problem. *Ann. Probab.* **5** 636–644.

GIANINI, J. and SAMUELS, S. M. (1976). The infinite secretary problem. *Ann. Probab.* **4** 418–432.

STEWART, T. J. (1981). The secretary problem with an unknown number of options. *Oper. Res.* **29** 130–145.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024-1555

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907