J. Appl. Prob. **27**, 722–733 (1989)
Printed in Israel

EXTREMAL PROCESSES, SECRETARY PROBLEMS AND THE 1/*e* LAW

DIETMAR PFEIFER,* *University of Oldenburg*

Abstract

We consider a class of secretary problems in which the order of arrival of candidates is no longer uniformly distributed. By a suitable embedding in a time-transformed extremal process it is shown that the asymptotic winning probability is again $1/e$ as in the classical situation. Extensions of the problem to more than one choice are also considered.

NON-HOMOGENEOUS POISSON PROCESS; POISSON BINOMIAL DISTRIBUTION; COUPLING; RECORD VALUES

1. Introduction

In recent years many generalizations of the classical secretary problem as described by Gilbert and Mosteller (1966) have been investigated by several authors (see Freeman (1983) for an exhaustive survey and bibliography). The problem is, briefly, to find with a high probability the best out of $n \geqq 1$ candidates (secretaries) by comparisons, when recalling a person examined earlier is not possible. (The same situation occurs if a tourist wants to take a photo of the most beautiful site to be visited, when there is only one shot left and all n places are visited only once.) However, the basic assumption in most of the models considered is the uniform distribution of the order of arrival of the n candidates, i.e., assigning an equal probability of $1/n!$ to each of the possible permutations of the candidates' ranks. On comparing an early approach of Dynkin (1963) with the so-called record problem of Rényi (1962) it is immediately obvious that the original problem can be reformulated in terms of extremal statistics in the following way.

Let $\{X_n\}$ be an i.i.d. sequence of random variables with continuous c.d.f. F and define

$$(1.1)\qquad I_n = \begin{cases} 1, & \text{if } X_n > \max(X_1, \cdots, X_{n-1}) \\ 0, & \text{otherwise} \end{cases} \quad \text{for } n \geqq 2$$

where we use the convention $I_1 = 1$. If $I_n = 1$ for some n, we say that X_n is a *record value* of the sequence. Rényi (1962) has shown that the $\{I_n\}$ form an independent sequence with

Received 16 June 1987; revision received 25 October 1988.
* Postal address: Fachbereich 6 Mathematik, Universität Oldenburg, Postfach 25 03, D-2900 Oldenburg, W. Germany.

(1.2) $$P(I_n = 1) = 1 - P(I_n = 0) = 1/n, \qquad n \geqq 1.$$

Considering the superposition of the Bernoulli point processes δ_{I_n} (where δ_x is the atomic probability measure with all mass at x) as a point process ζ, i.e. $\zeta(A) = \Sigma_{n\in A} I_n$ for Borel sets A, it is easily verified that the corresponding arrival time sequence $\{T_n\}$, say, is a homogeneous Markov chain with transition probabilities given by

(1.3) $$P(T_{n+1} > j \mid T_n = i) = i/j, \qquad 1 \leqq i \leqq j.$$

The $\{T_n\}$ are also called the record times associated with the sequence $\{X_n\}$. (A 'time-reverse' Markov chain to $\{T_n\}$ is considered by Ross (1982), (1983) in connection with a model for the average-case analysis of the simplex algorithm in linear programming). By the i.i.d. assumption for $\{X_n\}$, all possible orderings of values for $X_1, \cdots X_n$ are equally probable, again with value $1/n!$ The original problem is hence equivalent to detecting the maximum in the sequence $X_1, \cdots, X_n$, or in other words detecting the last '1' in the sequence $I_1, \cdots, I_n$ by means of stopping rules (i.e. by making decisions on the basis of observations preceding the stopping moment). A similar problem arising in computer science has recently been treated by Kemp (1984), in connection with searching algorithms (pp. 21–26). By the independence of the I_k it is now easy to derive the optimal stopping rule for the classical problem (cf. also Bruss (1984)), i.e. the stopping rule which maximizes the detection probability for the best candidate, as follows.

For $1 \leqq c \leqq n$, let $S_{n,c} = \Sigma_{k=c}^{n} I_k$ denote the number of records between observation c and observation n. Determine c such that $P(S_{n,c} = 1)$ is maximal. The optimal strategy then is to accept the first candidate from c onwards who is better than the previous ones (i.e., who induces a record value).

Note that here $S_{n,c}$ follows a Poisson binomial distribution with

(1.4) $$\begin{aligned} P(S_{n,c} = 1) &= P(\text{exactly one of } I_c, \cdots, I_n \text{ is } 1) \\ &= \sum_{k=c}^{n} p_k \prod_{\substack{i=c \\ i\neq k}}^{n} (1 - p_i) = \prod_{i=c}^{n} (1 - p_i) \sum_{k=c}^{n} p_k/(1 - p_k) \end{aligned}$$

where $p_k = 1/k$, $k \geqq 1$. Explicit calculation of (1.4) gives, more simply,

(1.5) $$P(S_{n,c} = 1) = \frac{c-1}{n} \sum_{k=c}^{n} \frac{1}{k-1}$$

as is well known. Since in the present case, $\Sigma_{k=1}^{\infty} p_k^2 < \infty$, $\Sigma_{k=1}^{\infty} p_k = \infty$, It is clear that $S_{n,c}$ is asymptotically Poisson distributed with mean $\Sigma_{k=c}^{n} p_k = \log(n/c) + o(1) = \lambda + o(1)$, say (cf. Deheuvels and Pfeifer (1986)), whenever c is dependent on n in such a way that n/c is asymptotically constant. Likewise, when $n/c \to \infty$ the total variation distance between the distribution of $S_{n,c}$ and the corresponding Poisson distribution with mean $\log(n/c)$ tends to zero. Hence we can conclude that the asymptotic winning probability is $\sup_c P(S_{n,c} = 1)$ tending to $\sup\{\lambda e^{-\lambda} \mid \lambda > 0\} = 1/e$, the latter being achieved for $\lambda = 1$, which means $c \sim n/e$ (to be rounded to a suitable integer; cf. Gilbert and Mosteller (1966)).

Another approach to the Poisson setting of the problem can be made by a suitable embedding into an extremal-F process $\{E(t) \mid t > 0\}$. Such embeddings have been successfully applied before, for example by Resnick (1973), (1974), (1975), Resnick and Rubinovitch (1973), Weissman (1975), and most recently by Ballerini and Resnick (1985), (1987a,b) in connection with linear trend models for records involving non-i.i.d. random variables. It is the purpose of this paper to show that similar embedding techniques for the i.i.d. case can be used to give more insight into the classical secretary problem, and an embedding into non-homogeneous extremal processes to be defined later provides a simple tool for generalizing and solving the problem in the case of non-uniformly distributed arrivals of candidates. Returning to the classical situation, some characteristics of $\{E(t)\}$ are $\{\max_{1 \leqq k \leqq n} X_k\} \overset{\mathcal{D}}{=} \{E(n)\}$ where $\overset{\mathcal{D}}{=}$ means equality in distribution, and the jump times $\{\tau_n \mid -\infty < n < \infty\}$ of the extremal process form a non-homogeneous Poisson point process ξ with intensity measure $E\xi(B) = \int_B (1/s)ds$ for Borel sets B contained in $\mathbb{R}^+$. Letting $Y_n = \xi((n-1, n])$, we see that $\{Y_n\}$ forms an independent Poisson-distributed sequence with $E(Y_n) = -\log(1-p_n)$ where again $p_n = 1/n$, $n \geqq 2$ and also $I_n = \min(1, Y_n)$ which gives a maximal (and also individually optimal) coupling in the sense of Serfling (1978). Hence we obtain immediately lower Poisson bounds for the winning probability for an arbitrary choice of c by the fact that the Y_n are stochastically larger than the I_n; in other words, if ζ again denotes the Bernoulli point process $\Sigma_{n=1}^{\infty} \delta_{I_n}$, we have, for any c,

$$
\begin{aligned}
P(S_{n,c} = 1) &= P(\zeta([c, n]) = 1) = P(\xi((c-1, n]) \geqq 1) \geqq P(\xi((c-1, n]) = 1) \\
&= P\left(\sum_{k=c}^{n} Y_k = 1\right) = \mu e^{-\mu} = \frac{c-1}{n} \log \frac{n}{c-1}
\end{aligned} \tag{1.6}
$$

with

$$
\mu = \sum_{k=c}^{n} -\log\left(1 - \frac{1}{k}\right) = \log \frac{n}{c-1} \quad (n, c \geqq 2).
$$

(See also Pfeifer (1986).)

(Note that $\xi((c, d]) = 0$, whenever $\zeta((c, d]) = 0$, for $c < d$.)

The full power of the Poisson approach becomes apparent, however, when the possibility of more than one choice is considered, in the sense described in Gilbert and Mosteller (1966). Although the optimal strategy for this case is rather complicated (cf. also Székely (1986)), we can easily obtain a suboptimal strategy which is nearly as good as the optimal strategy. Suppose that the possibility of K, say, choices is given ($1 < K < n$). Then we could determine a value of c such that $P(1 \leqq S_{n,c} \leqq K)$ is maximal, and choose up to K candidates starting with c with improving ranks (i.e., up to K 'ones' in the sequence $I_c, \cdots, I_n$). The probability of having the best candidate among these can then be estimated in a similar way by

$$
\begin{aligned}
P(1 \leqq S_{n,c} \leqq K) &= P(1 \leqq \zeta([c, n]) \leqq K) \geqq P(1 \leqq \xi((c-1, n] \leqq K) \\
&= e^{-\mu} \sum_{k=1}^{K} \mu^k / k!
\end{aligned} \tag{1.7}
$$

where again μ is as in (1.6). The right-hand side of (1.7) is now maximized by $\mu \sim \sqrt[K]{K!}$, or $c \sim n/\exp(\sqrt[K]{K!})$. The foregoing analysis shows why the Poisson probabilities given in Gilbert and Mosteller (1966) are always lower bounds for the asymptotic optimal winning probability; a fact which has not been proved rigorously before. Numerical computations even show that the bounds obtained by (1.7) are also superior to those previously described in the literature (see Székely (1986)). Table 1 compares such upper and lower bounds for the asymptotically optimal winning probability with the above Poisson probabilities, as well as the exact values from Gilbert and Mosteller (1966).

TABLE 1

K	2	3	4	5	6	7
$1-e^{-K/e}$	0.521	0.668	0.770	0.841	0.890	0.924
$\Sigma_{k=1}^{K} e^{-\mu}\mu^k/k!$	0.587	0.726	0.817	0.877	0.917	0.944
$1-(1-1/e)^K$	0.600	0.747	0.840	0.899	0.936	0.960
Asymptotically optimal winning probabilities	0.591	0.732	0.823	0.883	0.922	0.948

In the following section we shall drop the assumption of the equidistribution of possible orders of arrival of candidates in allowing unequal distributions for the sequence $\{X_n\}$.

2. A generalized secretary problem

Here we shall suppose that the sequence $\{X_n\}$ of 'qualities' of candidates is still independent, but each X_n possesses a c.d.f. of the form F^{α_n} with $\{\alpha_n\}$ such that $\alpha_n > 0$ and, with

$$p_n = \alpha_n \Big/ \sum_{k=1}^{n} \alpha_k, \tag{2.1}$$

$$\sum_{n=1}^{\infty} p_n = \infty, \qquad \sum_{n=1}^{\infty} p_n^2 < \infty. \tag{2.2}$$

Typically, we might have $\alpha_n = n^\beta$ with some real $\beta \geqq -1$ in which case we have

$$p_n \sim \begin{cases} \dfrac{1}{n \log n}, & \beta = -1 \\ (1+\beta)\dfrac{1}{n}, & \beta > -1. \end{cases} \tag{2.3}$$

The case of an increasing sequence $\{\alpha_n\}$ models, for instance, the situation where candidates are preselected according to increasing formal qualifications, whereas the $\{X_n\}$ correspond to some random influence such as personal behaviour, additional skills, etc. In the situation described above we would expect to wait for a longer time period before choosing a candidate since we expect the best candidate to be among the latter ones; similarly, in case that $\{\alpha_n\}$ is decreasing, we would expect to make a decision at an earlier stage which might be advantageous in order to save interviewing costs. At first sight it may seem that the winning probability under the optimal strategy is quite different for the two situations outlined above; actually, under the conditions specified in (2.1), it can be shown that — at least asymptotically — the winning probability will also be close to $1/e$ as in the classical case. This is due to the fact that as long as (2.1) holds we are in the range of Poisson approximation, hence the arguments presented in the introduction can be applied again. In what follows we shall make these arguments more rigorous by embedding the problem into a time-transformed extremal process which shows that we are confronted with the same situation as before except for a change of time scale.

For this purpose, suppose that a real-valued function $A(t)$, $t > 0$, is defined in such a way that

$$A(n) = \sum_{k=1}^{n} \alpha_k, \qquad n \in \mathbb{N}; A(0) = 0 \tag{2.4}$$

$$A(t), \quad t > 0 \text{ is strictly increasing.} \tag{2.5}$$

Note that with our assumptions such a function $A(t)$ always exists. Without loss of generality, we also want to assume that $\alpha_1 = 1$, which merely corresponds to a proper normalization of the distributions considered. Then the stochastic process $\{E^*(t) \mid t > 0\}$ defined by

$$E^*(t) = E(A(t)), \qquad t > 0 \tag{2.6}$$

is a non-homogeneous Markov jump process (called a non-homogeneous extremal-F process) with the property that

$$\{E^*(n)\} \stackrel{\mathscr{D}}{=} \left\{ \max_{1 \leqq k \leqq n} X_k \right\}; \tag{2.7}$$

note that in particular, $P(E^*(t) \leqq x) = F^{A(t)}(x)$ $(t > 0, x \in \mathbb{R})$, hence

$$P(E^*(n) \leqq x) = \prod_{k=1}^{n} F^{\alpha_k}(x) = P\left(\max_{1 \leqq k \leqq n} X_k \leqq x \right).$$

From (2.6) it follows that the corresponding jump times $\{\tau_n^* \mid -\infty < n < \infty\}$ form again a Poisson point process ξ^* with

$$E\xi^*((c, d]) = \log \frac{A(d)}{A(c)} \quad \text{for } 0 < c < d. \tag{2.8}$$

(Such processes have also been considered recently by Zhang (1988).)

If we define $\{I_n^*\}$ and $\{Y_n^*\}$ analogously to the classical case we see again that $I_n^* = \max(1, Y_n^*)$, this time with

(2.9) $$E(Y_n^*) = \log \frac{A(n)}{A(n-1)}, \qquad n \geqq 2,$$

implying

(2.10) $$p_n = P(I_n^* = 1) = 1 - \exp(-E(Y_n^*)) = 1 - \frac{A(n-1)}{A(n)} = \alpha_n \Big/ \sum_{k=1}^{n} \alpha_k$$

as in (2.2) (cf. also Nevzorov (1986)). In particular, $\{I_n^*\}$ again forms an independent Bernoulli sequence, and it is easy to see that for the corresponding record time sequence $\{T_n^*\}$ we again obtain a homogeneous Markov chain with transition probabilities

(2.11) $$P(T_{n+1}^* > j \mid T_n^* = i) = \frac{A(i)}{A(j)}, \qquad 1 \leqq i \leqq j$$

(the time-reversed Markov chain to this Markov chain was also considered by Ross (1982) in connection with a more general model for an average-case analysis for the simplex method).

Note that the divergence part of condition (2.2) guarantees that this Markov chain is non-degenerate (i.e., is almost surely infinite) by the Borel–Cantelli lemma applied to the sequence $\{I_n^*\}$.

By the independence of the $\{I_n^*\}$ we can now conclude, as in the classical case, that for the one-choice problem the best strategy is again the one-step-ahead policy: determine a value c^* such that $P(S_{n,c^*} = 1)$ is maximal, where again $P(S_{n,c^*} = 1)$ is given by relation (1.4), but this time with the p_k as in (2.2). By our assumptions, again

(2.12) $$P(S_{n,c^*} = 1) = \frac{A(c^* - 1)}{A(n)} \sum_{k=c^*}^{n} \frac{\alpha_k}{A(k-1)} \geqq P(\xi^*((c^* - 1, n]) = 1) = \mu^* e^{-\mu^*}$$

with

(2.13) $$\mu^* = \log \frac{A(n)}{A(c-1)},$$

with the right-hand side being the limit of the left-hand side for $n \to \infty$ (see Deheuvels and Pfeifer (1986)). This means that for the asymptotically optimal c^* we must again have $\mu^* \sim 1$, or

(2.14) $$c^* \sim A^{-1}\left(\frac{A(n)}{e}\right),$$

again with an asymptotically optimal winning probability of $1/e$ by (2.12). For example, with the choice $\alpha_n = n^\beta$, $\beta \geqq -1$ we have

(2.15) $$c^* \sim \begin{cases} n \exp(-1/(1+\beta)), & \beta > -1 \\ n^{1/e}, & \beta = -1 \end{cases}$$

which reflects the fact that for increasing sequences $\{\alpha_n\}$ the stopping moment is relatively later than in the classical situation, while for decreasing sequences $\{\alpha_n\}$ the stopping moment is relatively earlier.

Note that by the convergence part of conditions (2.1) there exists an a.s. finite random variable S such that $\xi^*([2, n]) - S_{n,1} \to S$ a.s. for $n \to \infty$ (see Pfeifer (1986) or Resnick (1973), (1975)). From here it follows by the LLN that for the number $S_{n,1}$ of choosable candidates available (i.e., of records) we have

$$\frac{S_{n,1}}{\log A(n)} \to 1 \quad \text{a.s. } (n \to \infty). \tag{2.16}$$

In our examples,

$$\log A(n) \sim \begin{cases} (1+\beta)\log n, & \beta > -1 \\ \log\log n, & \beta = -1; \end{cases} \tag{2.17}$$

similarly, asymptotic normality of $(S_{n,1} - \log A(n))/\sqrt{\log A(n)}$ can be proved, as well as a corresponding LIL, i.e.,

$$\limsup_{n\to\infty} \pm (S_{n,1} - \log A(n))/\sqrt{2\log A(n) \log\log\log A(n)} = 1 \quad \text{a.s.}$$

Similarly, relation (2.11) can be used for a strong Poisson approximation for the times T_n^* when new outstanding candidates (or records) occur. Using Theorem 1 in Pfeifer (1987), we see that we can without loss of generality define a Poisson arrival process $\{\tau_n^*; n \geqq 1\}$ with unit rate and a non-negative random variable Z on the same probability space such that

Z and $\{(\tau_n^* - n)/\sqrt{n}\}$ are asymptotically independent;

$$\log A(T_n^*) = Z + \tau_n^* + o(1) \quad \text{a.s.} \tag{2.18}$$

Actually, Z can be expressed as

$$Z = \sum_{k=1}^{\infty} \log\left(1 + W_k \frac{p_{T_k^*}}{1 - p_{T_k^*}}\right) \tag{2.19}$$

where $\{W_k\}$ is an i.i.d. sequence on the same probability space, uniformly distributed over (0, 1), and independent of $\{T_k^*\}$. From here we immediately obtain dual limit theorems to those for $\{S_{n,1}\}$:

$$\frac{\log A(T_n^*)}{n} \to 1 \quad \text{a.s. } (n \to \infty); \tag{2.20}$$

$$(\log A(T_n^*) - n)/\sqrt{n} \tag{2.21}$$

is asymptotically normally distributed with zero mean and unit variance;

$$\limsup_{n\to\infty} \pm \frac{\log A(T_n^*) - n}{\sqrt{2n \log\log n}} = 1 \quad \text{a.s.} \tag{2.22}$$

This can be proved by observing that

$$Z \leqq \sum_{k=1}^{\infty} -\log(1 - p_{T_k^*}) \leqq \sum_{k=1}^{\infty} \frac{p_{T_k^*}}{1 - p_{T_k^*}} = \sum_{k=1}^{\infty} \frac{p_k I_k}{1 - p_k I_k} \tag{2.23}$$

which is almost surely finite if $\Sigma_{k=1}^{\infty} p_k^2 < \infty$, which holds by (2.1).

Note that in our examples, we also have asymptotically

$$\log A(T_n^*) \sim \begin{cases} (1+\beta)\log T_n^*, & \beta > -1 \\ \log\log T_n^*, & \beta = -1. \end{cases} \tag{2.24}$$

Of course, the same arguments as before also apply to the case of K choices. Analogously to (1.7) we obtain

$$P(1 \leqq S_{n,c^*} \leqq K) \geqq P(1 \leqq \xi^*((c^*-1, n]) \leqq K) = e^{-\mu^*} \sum_{k=1}^{K} \mu^{*k}/k! \tag{2.25}$$

with μ^* again as in (2.13) which shows that asymptotically the values given in Table 1 are again lower bounds for the optimal winning probability in the general K-choice problem. Actually, following the arguments in Gilbert and Mosteller (1966), it can be shown that for the optimal strategy also we obtain the same values as before in the classical situation, due to the fact that we only have a change in the time scale. Note that for the suboptimal solution, we have, analogously to (2.14),

$$c^* \sim A^{-1}(A(n)\exp(-\sqrt[K]{K!})). \tag{2.26}$$

Finally, in what follows we shall investigate the relationship between the distributional assumptions for the sequence $\{X_n\}$ and the distribution of possible orders of arrival of candidates. The following result will be the key for such an investigation.

Theorem. For any permutation σ of $\{1, \cdots, n\}$ we have

$$P(X_{\sigma(1)} < \cdots < X_{\sigma(n)}) = \prod_{k=1}^{n} \alpha_k \Big/ \prod_{k=1}^{n} \sum_{i=1}^{k} \alpha_{\sigma(i)}, \tag{2.27}$$

if σ and σ^* are permutations such that $\sigma^*(j) = \sigma(i) < \sigma(j) = \sigma^*(i)$, and $\sigma(k) = \sigma^*(k)$ for all $k \neq i, j$, then also

$$P(X_{\sigma(1)} < \cdots < X_{\sigma(n)}) > P(X_{\sigma^*(1)} < \cdots < X_{\sigma^*(n)}) \tag{2.28}$$

whenever $\alpha_1 < \cdots < \alpha_n$, and the converse relation holding in (2.28) whenever $\alpha_1 > \cdots > \alpha_n$.

Proof. Let $\alpha_k^* = \alpha_{\sigma(k)}$, $1 \leqq k \leqq n$. Define correspondingly $X_k^* = X_{\sigma(k)}$, and I_k^*, p_k^* as in (1.1) and (2.2), respectively. Then again $X_1^*, \cdots, X_n^*$ and hence also $I_1^*, \cdots, I_n^*$ (via the embedding) are independent. Now

$$P(X_{\sigma(1)} < \cdots < X_{\sigma(n)}) = P(I_1^* = \cdots = I_n^* = 1) = \prod_{k=1}^{n} p_k^* = \prod_{k=1}^{n} \frac{\alpha_k^*}{\sum_{i=1}^{k} \alpha_i^*}$$

whence (2.27) follows. For the second statement it suffices to assume $1 < j = i + 1 \leqq n$. Then

$$(2.29)\quad \frac{P(X_{\sigma(1)}<\cdots<X_{\sigma(n)})}{P(X_{\sigma^*(1)}<\cdots<X_{\sigma^*(n)})}=\prod_{k=1}^{n}\frac{\sum_{m=1}^{k}\alpha_{\sigma^*(m)}}{\sum_{m=1}^{k}\alpha_{\sigma(m)}}=\frac{\alpha_{\sigma(1)}+\cdots+\alpha_{\sigma(i-1)}+\alpha_{\sigma(i+1)}}{\alpha_{\sigma(1)}+\cdots+\alpha_{\sigma(i-1)}+\alpha_{\sigma(i)}}>1;$$

correspondingly for the second case.

Let Σ denote the order of candidates with respect to increasing ranks, and suppose that Σ has a distribution of the form

$$P(\Sigma=\sigma)=\prod_{k=1}^{n}\alpha_k\Big/\prod_{k=1}^{n}\sum_{i=1}^{k}\alpha_{\sigma(i)}.$$

Then the 'qualities' $\{X_n\}$ of the candidates can be considered as being independent random variables with distributions specified by the $\{\alpha_n\}$ as above. Note that if N denotes the position of the maximum in the sequence $X_1,\cdots,X_n$, then we have

$$(2.30)\quad P(N=k)=P(I_k=1,I_{k+1}=\cdots=I_n=0)=p_k\prod_{j=k+1}^{n}(1-p_j)=\frac{\alpha_k}{A(n)}$$

with the p_j as in (2.2).

Example 1. Suppose that $\alpha_k=k$, $1\leqq k\leqq n$. Then X_k corresponds to the maximum of k i.i.d. random variables with c.d.f. F. In this case,

$$(2.31)\qquad p_k=2/(k+1),\qquad 1\leqq k\leqq n,$$

and

$$(2.32)\qquad P(N=k)=2k/(n(n+1)),$$

which means that N follows a triangular distribution. Further,

$$(2.33)\qquad P(S_{n,c}=1)=\frac{(c-1)c}{n(n+1)}\sum_{k=c}^{n}\frac{2}{k-1},\qquad c\geqq 2.$$

Explicit calculation for $n=4$ gives optimality for $c=3$ with a corresponding winning probability of 0.5 (in the classical case, we would have $c=2$ with winning probability of 0.4583). Table 2 contains the distribution of the corresponding Σ.

TABLE 2

σ	(1234)	(1243)	(1324)	(1342)	(1423)	(1432)	(2134)	(2143)
$P(\Sigma=\sigma)$	0.1333	0.1143	0.1000	0.0750	0.0686	0.0600	0.0667	0.0571
σ	(2314)	(2341)	(2413)	(2431)	(3124)	(3142)	(3214)	(3241)
$P(\Sigma=\sigma)$	0.0400	0.0267	0.0286	0.0222	0.0333	0.0250	0.0267	0.0178
σ	(3412)	(3421)	(4123)	(4132)	(4213)	(4231)	(4312)	(4321)
$P(\Sigma=\sigma)$	0.0143	0.0127	0.0171	0.0150	0.0143	0.0111	0.0107	0.0095

Note that the corresponding distribution in the classical case would yield a constant probability of 0.0417 for each of the 24 cases.

Example 2. Suppose that $\alpha_k = 1/k$, $1 \leqq k \leqq n$. In this case, nice simple formulas for the expressions involved are not available; however,

$$P(N=k) \sim \frac{1}{k \log n} \tag{2.34}$$

and

$$P(S_{n,c}=1) \sim \frac{\log(c-1)}{\log n} \sum_{k=c}^{n} \frac{1}{k \log k}. \tag{2.35}$$

For $n = 4$, we have more precisely

k	1	2	3	4
$P(N=k)$	0.48	0.24	0.16	0.12

where the optimal strategy requires $c = 1$, i.e., choosing the first candidate without inspecting the other ones. Here the winning probability is $P(N=1) = 0.48$. The distribution of the corresponding Σ is given in Table 3.

TABLE 3

σ	(1234)	(1243)	(1324)	(1342)	(1423)	(1432)	(2134)	(2143)
$P(\Sigma=\sigma)$	0.0073	0.0076	0.0082	0.0095	0.0091	0.0101	0.0145	0.0152
σ	(2314)	(2341)	(2413)	(2431)	(3124)	(3142)	(3214)	(3241)
$P(\Sigma=\sigma)$	0.0262	0.0443	0.0305	0.0492	0.0245	0.0284	0.0393	0.0665
σ	(3412)	(3421)	(4123)	(4132)	(4213)	(4231)	(4312)	(4321)
$P(\Sigma=\sigma)$	0.0650	0.0949	0.0366	0.0404	0.0610	0.0985	0.0866	0.1266

The foregoing examples show that when N has a non-uniform distribution we usually have a gain in information in that the optimal winning probability is larger than in the classical case where N has a uniform distribution. However, for n large, and under our assumptions, the gain of information tends to zero since the asymptotic winning probability is precisely what we would obtain in the classical case.

Note also that by (2.30), to each possible distribution of N with support $\{1, \cdots, n\}$ there corresponds a class of models of the above form which are equivalent in that they differ only in the (fixed) underlying c.d.f. F (i.e., a fixed multiplying constant for the α_k.)

The foregoing analysis shows that as long as we are in the range of Poisson approximation we cannot expect to obtain asymptotically a better result than in the classical case. The following example shows that an improvement is indeed possible outside this approximation range.

Example 3. Suppose that we are given a sequence $\{\alpha_n\}$ such that for the p_n defined in (2.2) we have $p_n \to p$ for some $0 < p < 1$ as $n \to \infty$. (Such a situation occurs e.g. when we have $\alpha_n = \gamma^{n-1}$ for some $\gamma > 1$, with $p = (\gamma - 1)/\gamma$.) In this case, the asymptotically optimal strategy yields $c \sim n - m$ where m is the integer maximizing the expression $(m+1)p(1-p)^m$, i.e., $m \simeq \{-1/\log(1-p)\} - 1$. (This follows from relation (1.4), for example). Whenever this last expression is integer we obtain a winning probability of $-p/(e(1-p)\log(1-p))$ which is strictly larger than $1/e$, and in the general case this expression is always an upper bound for the optimal winning probability which by some numerical analysis can also be shown to remain larger than $1/e$ in general. For instance, if p is larger than 0.5, i.e., $\gamma > 2$, then we always have $m = 0$ which means that we choose the last candidate throughout without inspecting the other ones, with a winning probability of p. If, for instance, $p = 0.2$, i.e., $\gamma = 1.25$, then $m = 3$ or $m = 4$ is asymptotically optimal, with a winning probability of 0.4096. Note that the (asymptotically) optimal m depends only on p which means that the (asymptotically) optimal strategy considers only a fixed number of candidates, independent of the total number of candidates.

Acknowlegements

The author would like to express his gratitude to the Deutsche Forschungsgemeinschaft for financial support from a Heisenberg research grant, and to J. Gani for his kind hospitality during the author's stay at the University of California at Santa Barbara, where the paper was eventually finished. He would also like to thank F. T. Bruss for some inspiring discussions, and J. Preater, University of Keele, for pointing out an error in Table 1.

References

BALLERINI, R. AND RESNICK, S. I. (1985) Records from improving populations. *J. Appl. Prob.* **22**, 487–502.

BALLERINI, R. AND RESNICK, S. I. (1987a) Records in the presence of a linear trend. *Adv. Appl. Prob.* **19**, 801–828.

BALLERINI, R. AND RESNICK, S. I. (1987b) Embedding sequences of successive maxima in extremal processes, with applications. *J. Appl. Prob.* **24**, 827–837.

BRUSS, F. T. (1984) Patterns of relative maxima in random permutations. *Ann. Soc. Sci. Bruxelles* **98** (I), 19–28.

DEHEUVELS, P. AND PFEIFER, D. (1986) A semigroup approach to Poisson approximation. *Ann. Prob.* **14**, 663–676.

DYNKIN, E. B. (1963) The optimal choice of the stopping moment for a Markov process. *Dokl. Akad. Nauk. SSSR* **150**, 238–240.

FREEMAN, P. R. (1983) The secretary problem and its extensions: A review. *Internat. Statist. Rev.* **51**, 189–206.

GILBERT, J. AND MOSTELLER, F. (1966) Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61**, 35–73.

KEMP, R. (1984) *Fundamentals of the Average Case Analysis of Particular Algorithms*. Wiley, New York; Teubner, Stuttgart (Chapter 3).

NEVZOROV, V. B. (1986) Two characterizations using records. In *Stability Problems for Stochastic Models*. Lecture Notes in Mathematics **1233**, Springer-Verlag, 79–85.

PFEIFER, D. (1986) Extremal processes, record times and strong approximation. *Pub. Inst. Stat. Univ. Paris* **XXXI**, 47–65.

PFEIFER, D. (1987) On a joint strong approximation theorem for record and inter-record times. *Prob. Theory Rel. Fields* **75**, 213–221.

RÉNYI, A. (1962) Théorie des éléments saillants d'une suite d'observations. *Colloquium on Combinatorial Methods in Probability Theory*, 104–115, Mathematisk Institut, Aarhus Universitet, Denmark.

RESNICK, S. I. (1973) Extremal processes and record value times. *J. Appl. Prob.* **10**, 863–868.

RESNICK, S. I. (1974) Inverses of extremal processes. *Adv. Appl. Prob.* **6**, 392–406.

RESNICK, S. I. (1975) Weak convergence to extremal processes. *Ann. Prob.* **3**, 951–960.

RESNICK, S. I. AND RUBINOVITCH, M. (1973) The structure of extremal processes. *Adv. Appl. Prob.* **5**, 287–307.

ROSS, S. M. (1982) A simple heuristic approach to simplex efficiency. *European J. Operat. Res.* **9**, 344–346.

ROSS, S. M. (1983) *Stochastic Processes*. Wiley, New York (Chapter 4).

SERFLING, R. J. (1978) Some elementary results on Poisson approximation in a sequence of Bernoulli trials. *SIAM Rev.* **20**, 567–579.

SZÉKELY, G. J. (1986) *Paradoxes in Probability Theory and Mathematical Statistics*. D. Reidel, Dordrecht (Chapter 3).

WEISSMAN, I. (1975) Extremal processes generated by non-identically distributed random variables. *Ann. Prob.* **3**, 172–177.

ZHANG, Y. S. (1988) Strong Approximations in Extremal Statistics (in German). Ph. D. Thesis, RWTH Aachen.