The Annals of Probability
1977, Vol. 5, No. 4, 627-635

THE FINITE-MEMORY SECRETARY PROBLEM[1]

BY H. RUBIN AND S. M. SAMUELS

Purdue University

The expected rank of the individual selected in the secretary problem can be kept bounded as the number of individuals becomes infinite *even if we are not permitted to remember more than one individual at a time.* The least upper bound of these expected ranks is derived: it is approximately 7.4.

1. Introduction. The central question, when dealing with any of the so-called secretary problems, is: "Does the minimal risk remain bounded—or the success probability remain bounded away from zero—as the number of arriving individuals becomes infinite?" And the answer is yes, surprisingly often (the most notable exception, discovered by Gianini [4], occurs when the total number of arriving individuals is unknown, uniformly distributed on 1 to n, as n becomes infinite), especially considering the severity of the constraints:

(1) no information from the data may be used other than the sequence of relative ranks of successive arrivals;

(2) an individual may be selected only at the time it arrives; later recall is forbidden.

(Recent papers of Yang [9] and Smith and Deely [8] allow recall but with the penalty that the individual may no longer be available for selection. The latter paper, despite the similarity of its title to ours, uses the term "finite memory" in an entirely different sense.)

Perhaps the most surprising "yes" so far is the one we shall give in this paper, because not only do we exhibit sequences of stopping rules whose bounded risk is more transparently clear than previously thought possible; but these rules obey yet another very severe constraint:

(3) only one individual at a time can be "remembered."

To properly introduce our rules we must begin by recalling briefly the two "standard" secretary problems: the best choice problem and the rank problem.

1.1. *The best choice problem.* The object is to maximize the probability of selecting the best (rank 1) of n rankable individuals arriving in random order. (See any of many references, e.g., [2], [3, page 87] or [6].) This goal is not nearly as difficult to achieve as it may seem initially. For example, the rule

Received August 27, 1976.

[1] Research supported by National Science Foundation Grants MCS 76-08316 (Rubin) and MCS 76-07254 (Samuels) at Purdue University.

AMS 1970 *subject classification.* 60G40.

Key words and phrases. Secretary problem, optimal stopping rules, relative ranks, loss functions, dynamic programming.

"Let half the arrivals go by, then select the first one better than all previous ones (if any)," has success probability greater than $\frac{1}{4}$ for *all* n. The optimal rules are of this form except that we let s_n individuals go by, where s_n/n is about three-eighths. These rules obey the third constraint since at any given time we only need to remember the best of those individuals who have arrived so far.

1.2. *The rank problem.* Now the object is to *minimize* the *expected rank* of the individual chosen. (For this to make sense, we must demand that some individual be chosen: the last one to arrive if no earlier one was selected.) Chow et al. [1] derived the limiting minimal risk which is approximately 3.87, and the limiting form of the optimal rules. The optimal rules are of the form: "if we haven't stopped by 'time' $s_n(k)$, select the first arrival in $[s_n(k), s_n(k+1))$ which is one of the k best so far (if any)," where $s_n(k)/n$, for each k, tends to a limit which is *strictly* between 0 and 1.

Indeed there is no way to keep the risk bounded with rules which never accept an individual of relative rank higher than some preassigned value. Hence it has been supposed that it can never be as easy as in the best choice problem to see that the minimal risk remains bounded, and that bounded risk must surely be incompatible with the third constraint.

What we shall show in Section 3.3 is that both of these suppositions are false. That demonstration will be given explicitly for an "infinite secretary problem," as introduced by Rubin [7] and by Gianini and Samuels [5]. But implicitly, as explained in Section 3.2, it holds for the ordinary finite problem, which is the subject of Section 2.

While simplifying an old problem, we are led to a much more complicated new one: What are the best "finite memory" rules? Some of the difficulties are described in Sections 4—6.

2. Memory-length-one rules for the finite problem.

2.1. *The finite problem.* n individuals ranked one (= best), two (= second best), and so on, up to n (= worst), arrive in random order. As each one arrives, we have three choices: accept it (and stop), ignore it, or remember it. Only one arrival at a time can be remembered; so, if we choose to remember the current arrival, we must discard the previously remembered one. The one and only thing we can observe about the current arrival is whether it is better (lower rank) or worse (higher rank) than the currently remembered individual. Hence the only available strategies are all possible "action" strings:

$$\{W_i/B_i : i = 2, 3, \cdots, n-1\}$$

where each W and B is either accept, ignore, or remember, and "W_i/B_i" means if we haven't selected one of the first $i-1$ arrivals, do W_i if the ith arrival is *worse* than the remembered one, and do B_i if it is *better*.

(We have omitted $i = 1$ and n because every strategy must accept the nth arrival, if it has not previously stopped. And, for every criterion we consider in this paper, it is better to remember the first arrival, rather than to accept it.)

2.2. *The best choice problem.* If the object is to maximize the probability of selecting the best individual, we have already seen that this memory constraint does not prevent us from using the optimal rule for the classical problem, namely

$$\begin{aligned} W_i/B_i &= \text{ignore/remember} \qquad \text{for} \quad i < s_n \\ &= \text{ignore/accept} \qquad \text{for} \quad i \geqq s_n . \end{aligned}$$

2.3. *The best* 3-*action rule for the rank problem.* In addition to the two actions ignore/remember and ignore/accept, let us also allow just one more, namely remember/remember. The logic behind the latter action is that if we have been remembering an individual for a long time without encountering a better arrival, the remembered individual may be *too* good. Better to discard it and, in effect, start over again, rather than risk being forced to accept the last arrival.

The sense in which we "start over again" when we use remember/remember is this: If W_k/B_k = remember/remember then the conditional expected rank of the individual selected, given that we stop after time k, is just $(n+1)/(n-k+1)$ times the expected rank when the strategy

$$W_i'/B_i' = W_{k+i}/B_{k+i} \qquad 1 < i < n-k$$

is used in the $n-k$ individual problem. This follows easily from the elementary fact that the expected rank of the jth best among the last $n-k$ arrivals is $j(n+1)/(n-k+1)$.

We now give an algorithm for finding the best 3-action rule and its risk for $n \geqq 2$. Let

$$\begin{aligned} d_n(k) &= \text{minimal expected rank among all 3-action rules with} \\ &\qquad W_i/B_i = \text{ignore/remember for} \quad i \leqq k . \\ d_n(1) &= \text{minimal expected rank among all 3-action rules.} \end{aligned}$$

Then

$$\begin{aligned} d_n(n-1) &= (n+1)/2 \\ d_n(k) &= \min\left\{d_n(k+1),\ \frac{1}{k+1}\cdot\frac{n+1}{k+2} + \frac{k}{k+1}\,d_n(k+1),\right. \\ &\qquad \left.\frac{n+1}{n-k+1}\,d_{n-k}(1)\right\} \qquad \text{for} \quad 1 \leqq k < n-1 . \end{aligned}$$

The three terms in the bracket represent the minimal risks for W_{k+1}/B_{k+1} = ignore/remember, ignore/accept and remember/remember, respectively. The middle term follows from the fact that we are remembering the best of the first k arrivals, so the $(k+1)$st arrival has probability $1/(k+1)$ of being better, and, if better, has expected rank $(n+1)/(k+2)$.

The second term becomes smaller than the first at

$$a_n = \min\{k : (k+2)d_n(k+1) > n+1\} .$$

The third term becomes minimal at, say, r_n, which must be bigger than a_n since otherwise a best rule would be to, in effect, totally ignore the first $r_n - 1$ arrivals, which is impossible. (In fact for $n \leqq 10$, the third term is never minimal.)

A best 3-action rule is then

$$\begin{aligned} W_i/B_i &= \text{ignore/remember} && \text{for } i < a_n \\ &= \text{ignore/accept} && \text{for } a_n \leqq i < r_n \\ &= \text{remember/remember} && \text{for } i = r_n \\ &= W'_{i-r_n}/B'_{i-r_n} && \text{for } i > r_n, \end{aligned}$$

where

$$\{W_j'/B_j' : 1 \leqq j \leqq n - r_n\}$$

is best for the $n - r_n + 1$ individual problem.

Table 1 gives some values of a_n, r_n and $d_n(1)$, the minimal risk.

TABLE 1

Best rules and risks using only the three actions: Ignore/Remember, Ignore/Accept and Remember/Remember

$n =$ Total number of individuals	$d_n(1) =$ Minimal risk	$a_n =$ Earliest ignore/accept	$r_n =$ Earliest remember/remember
3	1.667	2	—
4	1.875	2	—
5	2.100	2	—
6	2.333	3	—
7	2.476	3	—
8	2.625	3	—
9	2.778	3	—
10	2.933	3	—
11	3.071	3	8
20	3.869	5	14
21	3.933	5	14
25	4.185	6	18
47	5.033	9	28
50	5.114	9	30
100	5.885	17	54
250	6.599	38	126
500	6.932	72	241
1000	7.138	140	468

From this table it can be seen that the best rule for $n = 50$ uses remember/remember with the 30th and 43rd arrivals and ignore/accept with the 9th through 29th, 34th through 42nd, and 45th through 49th arrivals. The reader can also work out the best 3-action rule for $n = 100$.

It does appear from this table that the risks are remaining bounded as n becomes infinite, but the evidence will not become conclusive until we look at the infinite problem in the next section.

3. Memory-length-one rules for the infinite rank problem.

3.1. *The infinite problem.* As in Gianini and Samuels [5], the arrival times of the best, the second best, etc., individuals are an infinite sequence of independent random variables, each uniformly distributed on (0, 1). As we shall see later, it is not at all clear how best to further define the problem by specifying precisely what are the allowable strategies. Let us instead merely look at the analogues of the 3-action rules studied for the finite problem. These are defined by choosing a sequence

$$0 = R_0 < A_1 < R_1 < A_2 < \cdots < A_k < R_k < \cdots < 1 .$$

and stipulating that in each (R_{i-1}, A_i) we remember the best arrival in that subinterval only; then in (A_i, R_i) we accept the first arrival (if any) better than the best in (R_{i-1}, A_i).

3.2. *Infinite problem risk > finite problem risk.* We shall shortly demonstrate by a very simple argument that there are rules for which the expected rank of the individual chosen is finite and then we shall compute the minimal risk and the rule which achieves it. But first let us underline the significance of that result: namely that this minimal risk is an *upper bound* for the minimal finite rank problem risks. (In fact it is actually the limit of the minimal 3-action risks, $\{d_n(1)\}$, as a modest amount of analysis would show.)

The argument is quite simple. Suppose we modify the rules by, in effect, allowing knowledge of when an arrival is one of the n best—by accepting the first arrival (if any) in (R_{i-1}, A_i) which is *both* better than the best in (R_{i-1}, A_i) *and* one of the n best overall, and by accepting the last of the n best to arrive, if we have not stopped sooner. Obviously the risk is reduced. But the resulting rule is now simply a *randomized* 3-action rule for the n individual problem.

3.3. *The minimal infinite problem risk is finite.* Choose the A_i's and R_i's so that for all $i = 0, 1, \cdots$

$$(R_{i+1} - R_i) = R_1(1 - R_1)^i \tag{1}$$

and

$$(A_{i+1} - R_i)/(R_{i+1} - R_i) \equiv p . \tag{2}$$

Let T be the stopping time and X be the rank of the individual selected. Clearly $P(T > R_i) = p^i$ so $T < 1$ a.s. Moreover, since the kth best arrival in $(R_1, 1)$ has expected rank $k/(1 - R_1)$, we easily conclude that

$$EX = EXI_{\{T<R_1\}} + P(T > R_1)EX/(1 - R_1) , \tag{3}$$

which is finite if and only if

$$P(T > R_1) = p < 1 - R_1 .$$

To see immediately that the minimal risk is less than 12, note that $E(X \mid T) = T^{-1}$, so

$$EXI_{\{T<R_1\}} < P(T < R_1)/A_1 = (1 - p)/A_1 = (1 - p)/pR_1 ;$$

hence, from (3),

$$EX < [(1 - p)/pR_1]/[1 - p/(1 - R_1)] .$$

Now take $p = R_1 = \frac{1}{3}$.

3.4. *Minimal risk and best rule.* First we assert that an optimal choice of the A's and the R's must satisfy (1) and (2) so that our task will then simply be to minimize EX with respect to the two parameters p and R_1; and the values of these for which the risk is minimized will completely specify the optimal rule. The basic idea is the same here as in the finite problem but slightly simpler to describe:

Since the arrival times of the best, second best, and so on, of those arriving after some time t_0 are independent, each uniform on $(t_0, 1)$; and since the expected rank of the kth best of these is $k/(1 - t_0)$ it follows that, for any m,

$$E(X \mid T > R_m) = E^*X/(1 - R_m)$$

where E^*X is the risk when we use the rule defined by the sequence

$$R_i^* = (R_{m+i} - R_m)/(1 - R_m)$$
$$A_i^* = (A_{m+i} - R_m)/(1 - R_m)\,;$$

the rest of the argument is straightforward. (Notice that as soon as we verify that there is a best choice of A_1 and R_1, it follows that the minimal risk is actually attained.)

Now, for given R_1 and p,

$$\begin{aligned} EXI_{\{T<R_1\}} &= \int_{a_1}^{R_1} E(X \mid T = t)\, dF_T(t) \\ &= \int_{pR_1}^{R_1} t^{-1}(pR_1 t^{-2})\, dt \\ &= (p^{-1} - p)/2R_1\,; \end{aligned}$$

hence

$$EX = (p^{-1} - p)(1 - R_1)/2R_1(1 - R_1 - p)\,. \tag{4}$$

For fixed p, the right side of (4) is minimized at $(1 - R_1) = p^{\frac{1}{2}}$. Substituting and minimizing with respect to p, the result is that $(1 - R_1)$ is the positive root of $x^3 + x^2 + x - 1 = 0$. Thus

$$\begin{aligned} R_1 &= .456311 \\ p &= .295598 \\ A_1 &= pR_1 = .13488 \end{aligned}$$

and the minimal risk is $1/A_1$, as we would expect from dynamic programming considerations:

$$EX' = 7.41375\,.$$

These values of R_1, A_1, and EX are the limits of r_n/n, a_n/n and $d_n(1)$, respectively, from the finite 3-action problem.

3.5. *Even for polynomial loss, the minimal risk is finite.* The same simple argument we have used when the loss is the rank (X) of the individual selected can be applied to any loss of the form $X(X + 1) \cdots (X + m)$. This is because the distribution of the rank (Z_k) of the kth best to arrive after some time, t_0, is

Pascal $(k, 1 - t_0)$. Hence

$$EZ_k(Z_k + 1) \cdots (Z_k + m) = k(k + 1) \cdots (k + m)/(1 - t_0)^m ;$$

so, as before, we conclude that any rule satisfying (1) and (2) has finite risk if and only if $p < (1 - R_1)^m$.

Of course every polynomial is dominated by one of these factorial forms. Thus we have an elementary demonstration that, whenever the loss is a polynomial function of the ranks, the minimal risk remains bounded as n becomes infinite, even when we are constrained to use memory-length-one rules.

4. A finite memory "counter" rule. Here is a competitor to the rules in Section 3. Whether or not to call it a memory-length-one rule is a matter of taste.

Having chosen A's and R's as in Section 3, choose a B_i between each A_i and R_{i+1}. In each (A_i, B_i) accept not the *first* arrival better than the best in (R_i, A_i), but the *second* one. In (B_i, R_i) accept the *first* better arrival just as before.

Can we improve on the rules of Section 3 this way? The answer is *no* as can be shown by direct calculation. Here is a streamlined argument:

First we only need to consider A's and R's satisfying (1) and (2) and B's such that

$$(B_{i+1} - R_i)/(R_{i+1} - R_i) \equiv q > p . \tag{4}$$

Now we compare this modified rule with the corresponding unmodified rule without the B's.

Let T and T' be the stopping times for the unmodified and modified rules respectively, and let X_T and $X_{T'}$ be the ranks of the individuals selected. Let D be the event that there is exactly one arrival in (A_1, R_1) better than the best in $(0, A_1)$ and its arrival time is in (A_1, B_1). It is easy to see that

$$\begin{aligned} E(X_{T'} \mid T' < R_1) &= E(X_{T'} \mid T' < R_1, D^c) \\ &= E(X_T \mid T' < R_1, D^c) \\ &= E(X_T \mid T < R_1, D^c) ; \\ E(X_T \mid D) &= E(X_T \mid T < R_1, D) \\ &= 1/R_1 \\ &< E(X_T \mid T < R_1) \\ &< E(X_T \mid T < R_1, D^c) . \end{aligned}$$

Thus

$$E(X_T \mid T < R_1) < E(X_{T'} \mid T' < R_1) .$$

Similarly, for all i,

$$E(X_T \mid T \in (R_i, R_{i+1})) < E(X_{T'} \mid T \in (R_i, R_{i+1})) .$$

Now both sides of this inequality are increasing in i, in fact

$$E(X_T \mid T \in (R_i, R_{i+1})) = E(X_T \mid T \in (0, R_1))/(1 - R_1)^i ,$$

and $T' > T$. The result follows immediately.

5. Optimal memory-length-one rules.

5.1. *The finite rank problem.* There are only nine possible actions, of which four—accept/accept, accept/ignore, accept/remember, and ignore/ignore—we can show should never be used. That leaves us with 5^{n-2} possible strategies for the n-arrival problem and, embarrassingly, no efficient way to analyze them. The standard method of backward induction is no help here, unless we rule out the actions remember/ignore and remember/accept. We are happy not to use the former, although unable to prove it should never be used, but sorry not to use the latter, which we know can improve the 3-action rules of Section 2 (though for large n the improvement is small).

For any *given*, *prescribed* strategy, we can compute its risk by forward induction using the following facts: let

T = stopping time,

R_i = relative rank of remembered individual after i arrivals (undefined on $T \leqq i$),

Y_i = relative rank of ith arrival.

Here is the distribution of R_{i+1} in terms of the distribution of R_i and the action we take with the $(i+1)$st arrival (all probabilities and expectations are conditioned on $\{T > i\}$, or $\{T > i+1\}$, whichever is appropriate):

$(i+1)$st *Action*	$P(R_{i+1} = j)$
ignore/remember	$\frac{i+1-j}{i+1} P(R_i = j) + \frac{1}{i+1} P(R_i \geqq j)$
ignore/accept	$(i+1-j)P(R_i = j)/E(i+1-R_i)$
remember/remember	$1/(i+1)$
remember/accept	$P(R_i < j)/E(i+1-R_i)$
remember/ignore	$\frac{1}{i+1} P(R_i < j) + \frac{j-1}{i+1} P(R_i = j-1)$

The two other formulas we need to compute the risk are:

$$P(i+1\text{st arrival better than remembered one}) \equiv P(Y_{i+1} \leqq R_i) = ER_i/(i+1)$$

$$E(\text{rank of } i+1\text{st arrival} \mid Y_{i+1} \leqq R_i) \equiv \frac{n+1}{i+2} E(Y_{i+1} \mid Y_{i+1} \leqq R_i)$$

$$= \frac{n+1}{i+2} ER_i(R_i+1)/2ER_i .$$

The above formulas help to illustrate the difficulty which arises when we try to do backward induction: in order to know what to do after time i, we need to know the distribution of R_i, which in turn depends on what we have done up to time i. The net result is that backward induction becomes just a way of directly evaluating *all* possible strategies.

6. Memory length greater than one. As bad as the situation is for memory-length-one rules, it is much worse when we are allowed to remember, say, m previous arrivals. Now with each arrival we have $m + 2$ choices: accept it, ignore it, or remember it instead of the kth best of the currently remembered ones, $k = 1, 2, \cdots, m$.

We have no results at all, not even in the infinite case, which for $m = 1$ is much more tractable than the finite case. We do venture the guess that the best strategy involves from time to time discarding the *best* of the remembered individuals.

REFERENCES

[1] CHOW, Y. S., MORIGUTI, S., ROBBINS, H. and SAMUELS, S. M. (1964). Optimal selection based on relative rank (The "Secretary Problem"). *Israel J. Math.* **2** 81–90.

[2] CHOW, Y. S., ROBBINS, H. and SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin, Boston.

[3] DYNKIN, E. B. and YUSHKEVICH, A. A. (1969). *Markov Processes, Theorems and Problems.* (In English.) Plenum Press, New York.

[4] GIANINI, J. (1976). The secretary problem with random number of individuals. In progress.

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

[6] GILBERT, J. P. and MOSTELLER, F. (1966). Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61** 35–73.

[7] RUBIN, H. (1966). The "secretary" problem. *Ann. Math. Statist.* **37** 544.

[8] SMITH, M. H. and DEELY, J. J. (1975). A secretary problem with finite memory. *J. Amer. Statist. Assoc.* **70** 357–361.

[9] YANG, W. C. K. (1974). Recognizing the maximum of a random sequence based on relative rank with backward solicitation. *J. Appl. Probability* **11** 504–512.

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
MATHEMATICAL SCIENCES BUILDING
WEST LAFAYETTE, INDIANA 47907