GILBERT, J. and MOSTELLER, F. (1966). Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61** 35–73.
GNEDIN, A. V. (1983). Effective stopping on a Pareto-optimal option. *Avtomat. i Telemekh.* 44 (3) 166–170. (In Russian).
GUTTMAN, I. (1960). On a problem of L. Moser. *Canad. Math. Bull.* **3** 35–39.
KARLIN, S. (1962). Stochastic models and optimal policy for selling an asset. In *Studies in Applied Probability and Management Science* (K. J. Arrow, S. Karlin and H. Scarf, eds.) 148–158. Stanford Univ. Press, Stanford, Calif.
KOESTLER, A. (1960). *The Watershed. A Biography of Johannes Kepler*. Anchor Books, Garden City, NY.
LINDLEY, D. V. (1961). Dynamic programming and decision theory. *Appl. Statist.* **10** 39–51.
LORENZEN, T. J. (1979). Generalizing the secretary problem. *Adv. in Appl. Probab.* **11** 384–396.
LORENZEN, T. J. (1981). Optimal stopping with sampling cost. *Ann. Probab.* **9** 167–172.
MACQUEEN, J. and MILLER, R. P., JR. (1960). Optimal persistence policies. *Oper. Res.* **8** 362–380.
MOSER, L. (1956). On a problem of Cayley. *Scripta Math.* **22** 289–292.
MUCCI, A. G. (1973). On a class of secretary problems. *Ann. Probab.* **1** 417–427.
PETRUCCELLI, J. D. (1980). On a best choice problem with partial information. *Ann. Statist.* **8** 1171–1174.
PETRUCCELLI, J. D. (1981). Best-choice problems involving uncertainty of selection and recall of observation. *J. Appl. Probab.* **18** 415–425.
PETRUCCELLI, J. D. (1984). Best-choice problems involving recall with uncertainty of selection when the number of observations is random. *Adv. in Appl. Probab.* **16** 111–130.
PRESMAN, E. L. and SONIN, I. M. (1972). The best-choice problem for a random number of objects. *Theory Probab. Appl.* **17** 657–668.
PRESMAN, E. L. and SONIN, I. M. (1975). Equilibrium points in a game related to the best choice problem. *Theory Probab. Appl.* **20** 770–781.
RASMUSSEN, W. T. and PLISKA, S. R. (1976). Choosing the maximum from a sequence with a discount function. *Appl. Math. Optim.* **2** 279–289.
ROSE, J. (1984). Optimal sequential selection based on relative ranks with renewable call options. *J. Amer. Statist. Assoc.* **79** 430–435.
SAKAGUCHI, M. (1961). Dynamic programming of some sequential sampling design. *J. Math. Anal. Appl.* **2** 446–466.
SAKAGUCHI, M. (1976). Optimal stopping problems for randomly arriving offers. *Math. Japon.* **21** 201–217.
SAKAGUCHI, M. (1986). Best choice problems for randomly arriving offers during a random lifetime. *Math. Japon.* **31** 107–117.
SAMUELS, S. M. (1981). Minimax stopping rules when the underlying distribution is uniform. *J. Amer. Statist. Assoc.* **76** 188–197.
SAMUELS, S. M. (1985). A best-choice problem with linear travel cost. *J. Amer. Statist. Assoc.* **80** 461–464.
SAMUELS, S. M. and CHOTLOS, B. (1986). A multiple criteria optimal selection problem. In *Adaptive Statistical Procedures and Related Topics* (J. Van Ryzin, ed.) 62–78. IMS, Hayward, Calif.
STADJE, W. (1980). Efficient stopping of a random series of partially ordered points. *Multiple Criteria Decision Making. Theory and Application Lecture Notes in Econ. Math. Syst.* **177** 430–447. Springer, New York.
STEWART, T. J. (1978). Optimal selection from a random sequence with learning of the underlying distribution. *J. Amer. Statist. Assoc.* **73** 775–780.
TAMAKI, M. (1986). A full-information best-choice problem with finite memory. *J. Appl. Probab.* **23** 718–735.
YANG, M. C. K. (1974). Recognizing the maximum of a sequence with backward solicitation. *J. Appl. Probab.* **11** 504–512.

Comment: Who Will Solve the Secretary Problem?

Stephen M. Samuels

Just like Johannes Kepler, who threw a new curve at the solar system, Tom Ferguson has given a different slant to the Secretary Problem. To its many practitioners who ritually begin by saying "all that we can observe are the relative ranks," Ferguson (citing historical precedent), in effect, responds "let's not take that assumption for granted." The heart of his paper, as I see it, is the following *Ferguson Secretary Problem*:

Stephen M. Samuels is Professor of Statistics and Mathematics at Purdue University. His mailing address is: Department of Statistics, Mathematical Sciences Building, Purdue University, West Lafayette, Indiana 47907.

Given n, either find an *exchangeable* sequence of continuous random variables, $X_1, X_2, \cdots, X_n$, for which, among all stopping rules, τ, based on the X's,

$$\sup_{\tau} P\{X_\tau = \max(X_1, X_2, \cdots, X_n)\}$$

is achieved by a rule based only on the *relative ranks* of the X's—or prove that no such sequence exists.

Ferguson has come within *epsilon* of solving this problem. He has exhibited exchangeable sequences, for each n and $\varepsilon > 0$, such that the best rule based only on relative ranks has success probability within ε of the supremum. But he has left open the question of whether this supremum can actually be attained.

For $n = 2$, the answer is easy; there is no such sequence. The following elementary argument, which

I heard from Tom Cover several years ago, neatly settles the matter. Pick any number, x, said Cover, and choose X_1 if $X_1 > x$; otherwise choose X_2. If both X_1 and X_2 turn out to be bigger than x, or if both are smaller than x, then (by exchangeability) this rule selects the larger of the two with probability ½, while, if one random variable is larger than x while the other is smaller, the larger one is sure to be chosen. Thus,

$$P\{X_\tau = \max(X_1, X_2)\} = \tfrac{1}{2}[1 + P\{\min(X_1, X_2) < x < \max(X_1, X_2)\}],$$

which is strictly greater than ½ as long as x is between the ess inf and the ess sup of the X's. This beats rules based only on relative ranks, which, for $n = 2$, are necessarily constants, so have probability ½ of success.

Cover's argument cannot be extended beyond $n = 2$, but here is a quite general negative result due to Hill (1968, Section 6): For no $m + 1 \geq 2$ is there an exchangeable sequence satisfying

$$P\{X_{(i)} < X_{m+1} < X_{(i+1)} \mid X_1, \cdots, X_m\} \equiv (m+1)^{-1} \quad \text{a.s.}$$

for each $i = 0, 1, \cdots, m$, where $X_{(1)}, \cdots, X_{(m)}$ are the order statistics of $(X_1, X_2, \cdots, X_m)$, $X_{(0)} = -\infty$, and $X_{(m+1)} = \infty$.

Hill proved this by first showing it for $m = 1$, using an argument similar to Cover's, and then showing that, if it were true for a sequence of length $m + 1$, it would also be true for the first m elements of the same sequence. Thus, Hill has shown that there is no exchangeable sequence such that observing the sequence itself is exactly like observing the relative ranks of the sequence.

But Hill's result by no means solves the *Ferguson Secretary Problem.* For example, for $n = 3$, the best rule based on relative ranks selects X_2 if $X_2 > X_1$, and X_3 otherwise; its success probability is ½. In order for this rule to be optimal among all rules based on the X's, it is necessary and sufficient that both of the following hold:

(a) $P\{X_1 = \max(X_1, X_2, X_3) \mid X_1\} \leq \tfrac{1}{2}$ a.s.
(b) $P\{\max(X_1, X_2) = \max(X_1, X_2, X_3) \mid X_1, X_2\} \geq \tfrac{1}{2}$ a.s.

Hill's result merely establishes the impossibility of these two conditional probabilities having, almost surely, the values ⅓ and ⅔, respectively.

For iid continuous X's, these random conditional probabilities take on all values in the unit interval; but, for other exchangeable sequences, the question is: what other possibilities are there? (Hill's result rules out constants.) Let us examine, for example, the conjugate prior mixtures of uniforms which Ferguson considered.

Let

$$U_j = \max(X_1, \cdots, X_j),$$
$$L_j = \min(X_1, \cdots, X_j).$$

When the X's are conditionally iid, uniform on $[0, \theta]$, and θ has the one-sided Pareto density given by Ferguson's (8.2), then, extending his (8.6),

$$\begin{aligned} P\{U_j = U_n \mid X_1, \cdots, X_j\} &= \int_{\max(m_0, U_j)}^{\infty} \left(\frac{U_j}{\theta}\right)^{n-j} (\alpha + j) \frac{[\max(m_0, U_j)]^{\alpha+j}}{\theta^{\alpha+j+1}} \, d\theta \\ &= \frac{\alpha + j}{\alpha + n} \left[\frac{U_j}{\max(m_0, U_j)}\right]^{n-j} \\ &\qquad \text{for all } U_j > 0 \text{ and any } m_0 > 0 \text{ and } \alpha > 0. \end{aligned}$$

For $n = 3$, $j = 1$, condition (a) is satisfied if $\alpha \leq 1$. But, since U_j can be arbitrarily close to zero, so can the above probability; hence condition (b) cannot be satisfied.

In this respect, the two-sided Pareto priors are clearly more promising. Their densities are given by Ferguson's (7.2). Then, from Ferguson's (7.3), we have

$$\begin{aligned} P\{U_j = U_n \mid X_1, \cdots, X_j\} &= \int_{\max(u_0, U_j)}^{\infty} \int_{-\infty}^{\min(l_0, L_j)} \left(\frac{U_j - \alpha}{\beta - \alpha}\right)^{n-j} \frac{(k+j)(k+j+1)}{(\beta - \alpha)^{k+j+2}} \\ &\qquad \cdot [\max(u_0, U_j) - \min(l_0, L_j)]^{k+j} \, d\alpha \, d\beta \\ &\qquad \text{for all } L_j \leq U_j \text{ and any } l_0 < u_0 \text{ and } k > 0. \end{aligned}$$

This integral has maximal value $(k + j + 1)/(k + n + 1)$ identically on $\{U_j \geq u_0\}$ [Ferguson's (7.4)], and minimal value (after a change of variable)

$$\begin{aligned} &\int_0^1 \frac{(k + j)(k + j + 1)}{(n + k + 1)} \pi^{k+j-1}(1 - \pi)^{n-j} \, d\pi \\ &\qquad = (n - j)!/(k + n + 1)(k + n) \cdots (k + j + 2) \end{aligned}$$

identically on $\{U_j = L_j \leq l_0\}$. This lower bound can also be derived from the upper bound because, on $\{U_j = L_j \leq l_0\}$,

$$\begin{aligned} P\{U_j = U_n \mid X_1, \cdots, X_n\} \\ = (n - j)! \, P\{L_j > X_{j+1} > \cdots > X_n \mid X_1, \cdots, X_n\}. \end{aligned}$$

Now, taking $n = 3$, $j = 2$, we have

$$P\{U_2 = U_3 \mid X_1, X_2\} \geq 1/(k + 4),$$

which, being strictly positive, is a qualitative improvement over the lower bound for the one-sided Pareto priors. But the parameter k has to be positive, so this bound is less than ¼, which is not good enough to meet condition (b). Moreover, the upper bound for $n = 3, j + 1$ is $(2 + k)/(4 + k)$ which is bigger than ½, so condition (a) cannot be satisfied either.

Here is one more result for Pareto priors: If we reflect the one-sided priors (i.e., look at $-X_1, \cdots, -X_n$), then the lower bound calculations are virtually the same as in the two-sided case, and the result is a slight improvement, to $(n - j)!/(\alpha + n)(\alpha + n - 1) \cdots (\alpha + j + 1)$ for any $\alpha > 0$. For $n = 3, j = 2$, this is an improvement from $1/(k + 4)$ to $1/(\alpha + 3)$—still not good enough to meet condition (b).

Thus the *Ferguson Secretary Problem* remains unsolved. Indeed, from these Pareto prior examples, it is not at all clear what the solution is: do the required exchangeable sequences exist or don't they? This quest for sufficiently "non-informative priors" should interest some Bayesians, too.

ADDITIONAL REFERENCE

HILL, B. M. (1968). Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *J. Amer. Statist. Assoc.* **63** 677–691.

Comment

Herbert Robbins

I am confused by Tom's attempt to clear up the confusion among various versions of the secretary problem. In Section 2 he defines the simplest form of the problem, in Section 4 he distinguishes secretary problems from Cayley's problem, etc. in which one observes numerical values of some possibly continuous random variable rather than just relative ranks, and in Section 5 he defines the 'general' secretary problem to be "a sequential observation and selection problem in which the payoff depends on the observations only through their relative ranks and not otherwise on their actual values." So far, so good. Then in Section 6 he introduces into the discussion the two-person googol game, which is not a secretary problem, and in Section 7 and Section 8 says that nobody has solved "the" secretary problem, possibly because no one realized that there was a game-theoretical problem to be solved. I can't agree with that.

Consider two cases of the secretary problem: (I) the payoff is 1 if we choose the best of the the n applicants, 0 otherwise, and we want to maximize the expected payoff, and (II) the loss is the absolute rank of the person selected (1 for the best, $\cdots$, n for the worst), and we want to minimize the expected loss. When all $n!$ orders of the applicants are equally likely the solutions of (I) and (II) have been known and published for some time. And when the probabilities of the various permutations are controlled by an antagonist, so that (I) and (II) become game-theoretical (minimax) problems, their solutions are *also* in the literature: See problem 7 on page 60 of Chow, Robbins and Siegmund (1971), and page 89 of Chow, Moriguti, Robbins and Samuels (1964).

In the latter reference it is also shown that when the $n!$ permutations *are* equally likely, the minimal expected loss for (II) with n applicants tends as $n \to \infty$ to the finite limit

$$A_1 = \prod_{j=1}^{\infty} \left(1 + \frac{2}{j}\right)^{1/(1+j)} \cong 3.8695.$$

This surprising result can be obtained by a heuristic argument involving a sequence of differential equations, but the argument is hard to make rigorous. The same heuristic argument yields a more general result: if the loss is taken to be $x(x + 1) \cdots (x + k - 1)$, where x is the absolute rank of the person selected and k is a fixed positive integer, then the minimal expected loss as $n \to \infty$ tends to

$$A_k = k!\left\{\prod_{j=1}^{\infty} \left(1 + \frac{k+1}{j}\right)^{1/(k+j)}\right\}^k.$$

(As $k \to \infty$ the quantity in braces tends to $e^{\pi^2/6} \cong 5.1807$.) But when the loss is x^2, rather than x or $x(x + 1)$, the limit as $n \to \infty$ of the minimal expected loss has not been exhibited explicitly by any formula such as this (it is, of course, less than A_2), nor has the minimax game-theoretical probability distribution of permutations been obtained for this case. Down with googol and up with problems like these!

Herbert Robbins is Professor of Statistics at Rutgers University. His mailing address is: Hill Center for the Mathematical Sciences, Busch Campus, Rutgers University, New Brunswick, New Jersey 08903.

ADDITIONAL REFERENCE

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