Choosing the Best: Decision-Making Under Uncertainty

Allan J. Rossman, Department of Statistics, Cal Poly – San Luis Obispo
arossman@calpoly.edu https://askgoodquestions.blog

presented for Math and Stat Club, CSU – Monterey Bay
October 7, 2019

Problem:
Your supervisor has assigned you the task of hiring a new employee subject to these rules:
1. A total of n candidates have applied for the position.
2. The candidates arrive in random order.
3. You interview candidates one at a time.
4. You can rank the candidates that you have interviewed from best to worst, but you have no *a priori* knowledge about the quality of the candidates.
5. Once you have interviewed a candidate, you must decide *immediately* whether to hire him/her. If you decide to hire, the process ends. If you opt not to hire, the process continues but that candidate can no longer be considered.
6. Your supervisor will be satisfied only if you hire the *best* candidate. Hiring the second best candidate is no better than hiring the very worst.

Predictions:
Before we start, take a guess concerning the optimal probability of success (hiring the best candidate) for the numbers of applicants listed in the table:

Number of applicants	4	12	24	500	7,601,872,508
Optimal probability of success (guess)					

Strategy:
We will start by considering possible strategies in cases with just a few applicants.

1. For n = 1 candidate, there is only one possible interview order: A: 1

The probability of choosing the best is:

2. For n = 2 candidates, there are two possible interview orders: A: 12 B: 21

The only possible strategies are:

The probability of choosing the best is:

3. For $n = 3$ candidates, there are six possible interview orders:

A: 123	B: 132	C: 213
D: 231	E: 312	F: 321

The optimal strategy is:

The probability of choosing the best with the optimal strategy is:

4. For $n = 4$ candidates, there are twenty-four possible interview orders:

A: 1234	B: 1243	C: 1324	D: 1342	E: 1423	F: 1432
G: 2134	H: 2143	I: 2314	J: 2341	K: 2413	L: 2431
M: 3124	N: 3142	O: 3214	P: 3241	Q: 3412	R: 3421
S: 4123	T: 4132	U: 4213	V: 4231	W: 4312	X: 4321

We will analyze two possible strategies. The optimal strategy turns out to be:

The probability of choosing the best with the optimal strategy is:

These relatively simple cases suggest a general form for the optimal strategy with n candidates:

Analysis:

Let n be the number of applicants and r be the position of first "contender." In other words, the strategy is to let the first $(r - 1)$ candidates go by. Because the best candidate is equally likely to be in any of the n possible positions, the probability of choosing the best is simply the average of the conditional probabilities of choosing the best given each possible position of the best candidate:

- If the best candidate is in one of the first $r - 1$ positions, then the probability of success (choosing the best) is:

- If the best candidate is in position r, then the probability of success is:

- If the best candidate is past position r and in position i, the probability of success is:

Thus, the overall probability of choosing the best can be expressed as:

For a given number of candidates n, this probability can be calculated for all possible values of r. Then the value of r can be selected to maximize this probability.

n	3	4	12	24	50	100	500	1000	10,000
r^*									
prob									

Notice that this probability is

Practice:

We can apply the optimal strategy on the following 25 simulated orderings of 12 applicants:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	4	8	7	4	7	4	2	9	12	1	5	3	12	1	5	3	1	12	3	6	3	6	8	5	7
2	12	1	5	3	12	1	5	3	11	11	10	7	8	5	11	11	10	7	1	12	4	8	7	4	3
3	1	12	3	6	3	6	8	5	7	4	2	9	6	10	12	1	9	2	4	8	10	7	1	12	10
4	5	3	9	2	9	2	4	8	1	12	3	6	3	6	8	5	2	9	6	10	2	9	6	10	4
5	6	10	12	1	1	12	3	6	8	5	11	11	5	3	9	2	12	1	5	3	12	1	5	3	11
6	7	4	2	9	6	10	12	1	5	3	9	2	9	2	4	8	6	10	12	1	1	12	3	6	8
7	11	11	10	7	8	5	11	11	6	10	12	1	1	12	3	6	4	8	7	4	7	4	2	9	12
8	3	6	8	5	11	11	10	7	2	9	6	10	2	9	6	10	3	6	8	5	11	11	10	7	1
9	8	5	11	11	5	3	9	2	9	2	4	8	10	7	1	12	7	4	2	9	6	10	12	1	5
10	10	7	1	12	4	8	7	4	3	6	8	5	11	11	10	7	8	5	11	11	5	3	9	2	9
11	2	9	6	10	2	9	6	10	4	8	7	4	7	4	2	9	11	11	10	7	8	5	11	11	6
12	9	2	4	8	10	7	1	12	10	7	1	12	4	8	7	4	5	3	9	2	9	2	4	8	2

The proportion of these orderings for which the optimal strategy succeeds in choosing the best is:

Approximate analysis:
For large values of *n*, we can approximate:

The probability of choosing the best can therefore be approximated by:

We can investigate how well this approximation performs for various values of *n*.

Asymptotic analysis:
For large values of *n*, we can determine the value of *r* that maximizes this probability by setting the derivative of the probability function with respect to *r* equal to 0 and solving for *r*:

For large values of *n*, the optimal probability is then:

Thus, for large values of *n*, the optimal procedure is:

With this strategy, even if every person on the planet applies for this position, the probability of *choosing the best* is approximately equal to:

Extensions:
1. Number game
2. Finding a soul-mate

References:
DeGroot, M.H. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.
DeGroot, M.H. (1986), *Probability and Statistics* (2nd ed.), Addison-Wesley, Reading MA.
Ferguson, T.S. (1989), "Who Solved the Secretary Problem?", *Statistical Science*, 4, 282-289.