Generalizations of the Secretary problem

Simon Grad

May 2024

1 Introduction

The Secretary Problem, also known as the "marriage problem," is a classic problem in decision theory and optimal stopping theory. It involves a scenario where one aims to choose the best candidate (secretary) from a pool of applicants who are interviewed sequentially and in random order. The challenge lies in deciding the optimal point to stop and make a selection, based on the relative ranks of the candidates seen so far. Traditional solutions to this problem provide a clear, probabilistic strategy for selecting the best candidate with a high probability.

In recent years, the Secretary Problem has been generalized and extended to a variety of more complex and realistic scenarios. These generalizations include multiple-choice variants, weighted selections, and adaptations to different distributions and constraints. The applicability of these generalizations spans fields such as online algorithms, resource allocation, and machine learning.

This report delves into several such generalizations of the Secretary Problem. To achieve a comprehensive understanding, we employ experimental methods using Python. By simulating different variants of the problem, we aim to uncover insights into their performance and optimal strategies under various conditions.

By leveraging the power of experimental methods in Python, this report aims to provide an understanding of the Secretary Problem's generalizations.

2 Standard Secretary Problem

In this section, we review the well-known solutions to the standard Secretary Problem. The standard Secretary Problem is defined as follows:

A known number of n items are to be presented one by one in random order, with all $n!$ possible orders being equally likely. The observer can rank the items presented so far in order of desirability. As each item is presented, the observer must either accept it, in which case the process stops, or reject it, in which case the next item in the sequence is presented and the observer faces the same choice as before.

The observer's goal is to maximize the probability that the item chosen is, in fact, the best of the n items available. We shall abbreviate this outcome to the single word 'win'. Since the observer is never able to go back and choose a previously-presented item which, in retrospect, turns out to be the best, he clearly has to balance the danger of stopping too soon and accepting an apparently-desirable item when an even better one might still be to come, against the danger of going on too long and finding that the best item was rejected earlier on.

In the following two subchapters, we explore the standard method and the average method for solving the standard Secretary Problem.

2.1 Standard Method

In this subsection, we examine the traditional approach to solving the Secretary Problem. We start by rejecting the first r candidates and then choose the first one that comes along who is better than everyone we have seen so far. This method also has an elegant analytical solution.

By $P(r)$ we denote the probability that we have chosen the best candidate if we rejected the first $r-1$ candidates. We can calculate;

$$
\begin{aligned}
P(r) &= \sum_{i=r}^{n} P(i \text{ is the best candidate} \cap i \text{ is selected}) \\
&= \sum_{i=r}^{n} P(i \text{ is selected} | i \text{ is the best candidate}) P(i \text{ is the best candidate}) \\
&= \sum_{i=r}^{n} P(\text{the best of } i-1 \text{ candidates is in the first } r-1 \text{ candidates} | i \text{ is the best candidate}) \cdot \frac{1}{n} \\
&= \sum_{i=r}^{n} \frac{r-1}{i-1} \cdot \frac{1}{n} \\
&= \frac{r-1}{n} \sum_{i=r}^{n} \frac{1}{i-1}
\end{aligned}
$$

So we can draw a graph of this function for $n = 100$ (Figure 1).

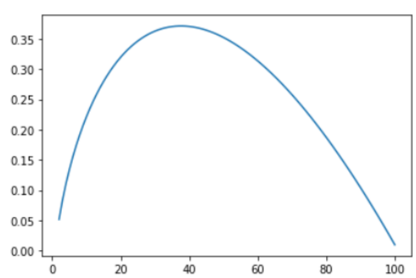


Figure 1: A graph of the probability function.

As we can see in Figure 1, the maximum of the function is achieved at $r = 38$. This means that the best strategy for us is to reject the first 37 candidates and then choose the first one who is better than everyone seen so far. In this case, we will choose the best candidate 37 percent of the time, which is quite good considering our goal is to select the best candidate.

This is inline with the analytical approach we can approximate the probability function by an integral written below.

$$P(x) \approx x \int_x^1 \frac{1}{t} dt = -x \ln x$$

Then we can take the derivative of this function setting it to 0 and solving for x.

$$\begin{aligned} P'(x) &= -lnx - 1 \\ lnx &= -1 \\ x &= e^{-1} \approx 0.36787944 \end{aligned}$$

We now analyze the simulation of the standard method for the standard Secretary Problem with n (number of candidates) being 100. We rated all candidates from 1 to 100 and shuffled them. Then, using the standard method, we selected a candidate. Using basic Monte Carlo simulation, we ran this $m = 100,000$ times to determine the percentage of successful attempts at finding the best candidate.

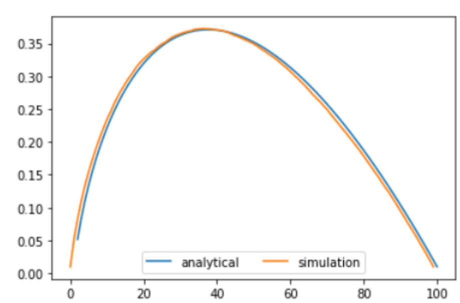
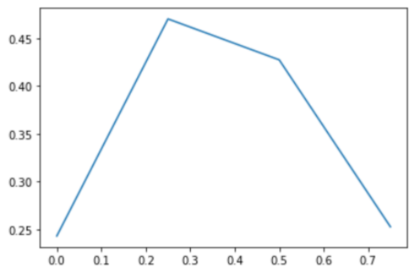


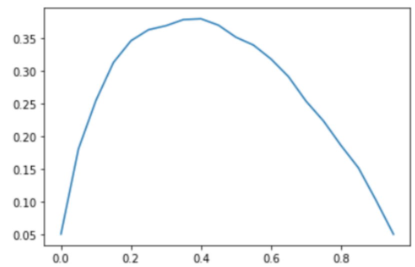
Figure 2: Comparison between analytical solution and simulation results.

The percentage we obtained for $r = n \cdot \frac{1}{e}$ is 0.37132, which is very close to the analytical solution of 0.37104. Then we checked if $\frac{1}{e}$ (or in our case, stopping at the 37th candidate) was truly the optimal solution. The results show that it was. In fact, in Figure 2, we can compare how well we performed with our Monte Carlo simulation. This gives us affirmation that our simulation works as it should and will provide correct results in the future.

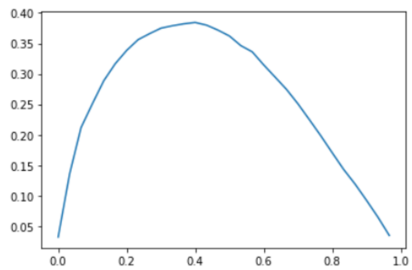
We conclude this subsection by examining graphs with different values of n (number of candidates)[Figure 3]. Specifically, we consider the following values for n: 4, 20, 30, 50, 500, and 100.



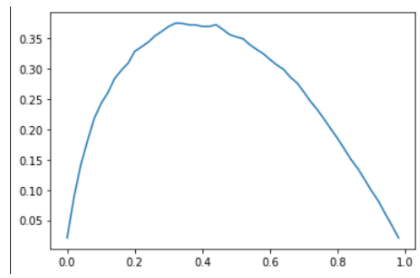
(a) $n = 4$



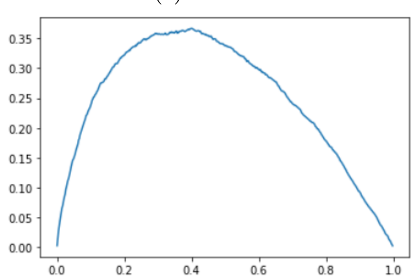
(b) $n = 20$



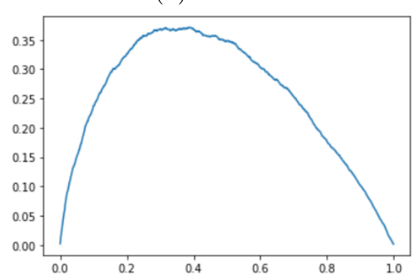
(c) $n = 30$



(d) $n = 50$



(e) $n = 500$



(f) $n = 1000$

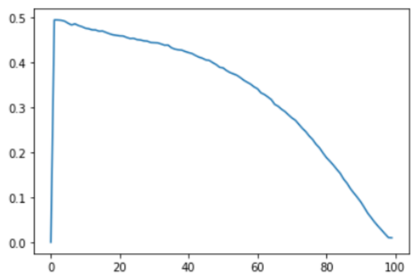
Figure 3: Graphs with different values of n (number of candidates)

2.2 Average Method

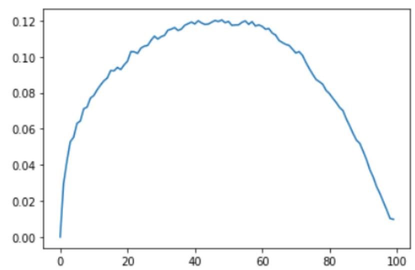
In this section, we explore an alternative approach that uses an average-based strategy. This method involves calculating the average rank of the candidates up until the r-th candidate and then choosing the first candidate that is better than the calculated average multiplied by a parameter α.

With this method, we aim to prioritize selecting the best possible candidate, thereby avoiding scenarios where we might choose the second-best option simply because it appeared before the best one. It seems that our parameter α should be somewhere around 2 since the expected average is $\frac{n}{2}$. By multiplying this average by 2, we get n (or the best possible option).

In Figure 4a, we examine the success rate of selecting a candidate, even if it is not the optimal one. In Figure 4b, we analyze our effectiveness in choosing the best candidate using this method.



(a) Selecting a candidate.

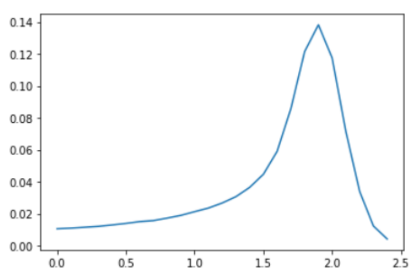


(b) Choosing the best candidate.

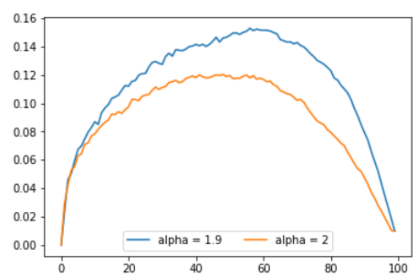
Figure 4: Graphs of Average Method.

We can see that the average method performs poorly. Perhaps the issue lies in the selection of α. We might need to use a smaller α, as we were unable to choose the best candidate more than 50% of the time.

Next, we test to determine the optimal value of α for this strategy. Focusing on the scenario where $r = 37$, we plot a probability function with respect to α (Figure 5a).



(a) Probability function with respect to α.



(b) Average method with $\alpha = 1.9$ and $\alpha = 2$.

Figure 5

Our findings indicate that the best choice is $\alpha = 1.9$. In Figure 5b, we compare the probability functions for successfully selecting the best candidate using the average method with $\alpha = 1.9$ and $\alpha = 2$. The comparison clearly illustrates the difference in performance.

3 Analysis of Equal Suitability

In this section we try to answer the questions;

- Is there a distinction when considering individuals of equal suitability, and does this affect our results?

- Moreover, does our outcome differ depending on whether we strictly prefer a superior candidate or are content with someone ranking equally to candidates we've previously encountered?

An immediate question arises: how will we determine the rating for each candidate?

In the real world, each candidate would be given an individual rating based on her merits. In our model, it makes the most sense to assign each candidate a random rating from 1 to *max_rating*. Of course, this introduces the 'flaw' of reality, where we might end up with only candidates rated as 1.

The second option, though less realistic, is more mathematically consistent. We assume that we have an equal number of candidates with ratings of one, two, three, and so on up to *max_rating*. This ensures that the number of candidates with each rating is the same.

While we have discussed only uniform distributions above, we could similarly use a normal distribution of candidate ratings (i.e. there would be more candidates with a rating of 5 than those with ratings of 1 or 10).

We can further ask how the above distributions will be affected if we decide to change *max_rating* in relation to n (the number of candidates). The possibilities are indeed vast. Therefore, we will focus on just a few options.

3.1 Real world ($\boldsymbol{max\_rating} = n$)

In this subsection, we discuss the option most likely to occur in the real world: we have 100 candidates who are each given a random number between 1 and 100. Because people are very different, we usually do not rank them equally, though there are some cases where this might happen. This is why I find this method the most realistic.

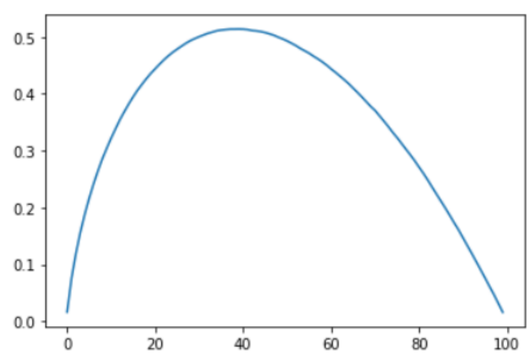


Figure 6

We begin by drawing the probability function (Figure 6) of how successful we are in choosing the best candidate. Note: this is not necessarily the candidate with the ranking of 100, because randomness might prevent such a candidate from existing. Additionally, we might have multiple best candidates if there are several people with the same rank.

From Figure 6, we can immediately observe that the probabilities are significantly higher than in the standard Secretary Problem. Additionally, the maximum is shifted slightly to the right; instead of occurring at 37, it peaks around 39.

In Figure 6, we allowed for choosing a candidate who is equally as good as a previously seen candidate. In Figure 7, we examine the difference between the two probability functions: the blue line represents the scenario where we can choose a candidate with the same ranking, while the orange line represents the scenario where we must choose a strictly better candidate.

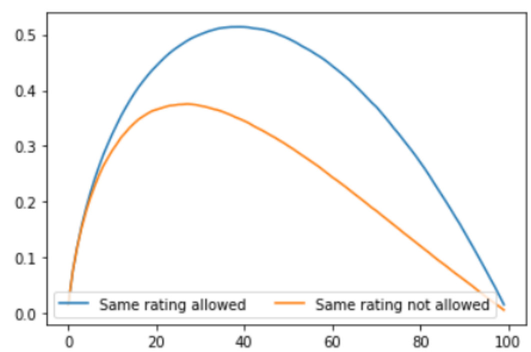


Figure 7: Probability functions with and without selecting the same-ranked candidate.

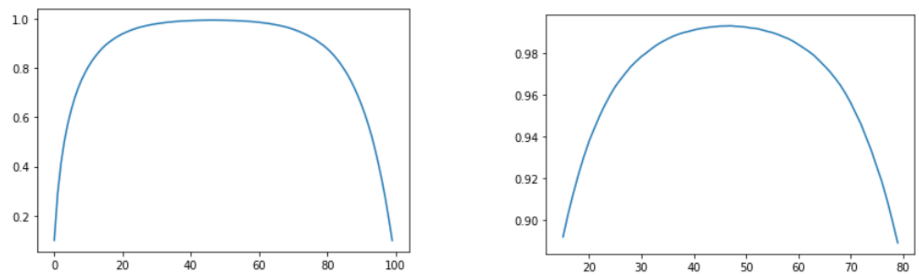
3.2 Rating from 1 to 10

In this subsection, we examine the case where we have 100 people ($n = 100$) and we rate each candidate on a scale of 1 to 10 (with 10 being the best candidates). We are successful in hiring the best candidate if we hire a candidate with the highest grade. In most cases, this will be a candidate with a grade of 10, and almost always there will be more than one candidate with a grade of 10.

In Figure 8a we notice that the probability is indeed significantly higher. Even if we randomly select the first candidate, we will have 10 % chance of selecting the best one. This time, it is worth waiting longer (reviewing and rejecting a larger sample) because we aim to find the person with the highest score, which is almost certain to occur (due to the large sample size of 100 candidates).

In Figure 8b, I provided a close-up of the middle part of the function so that we can better observe that the maximum of the probability function is now achieved by stopping at 47.

If we decide that we are not satisfied with a candidate who ranks the same as a previous one, we get the graph in Figure 9.



(a) (b)

Figure 8: Probability function with grades from 1 to 10.

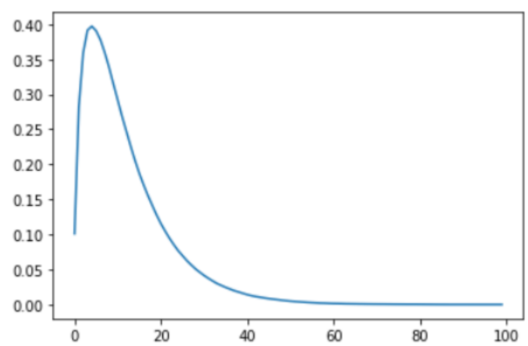


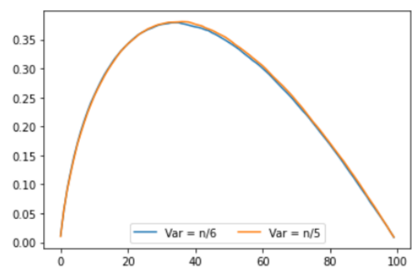
Figure 9: Probability function with grades from 1 to 10 when candidates with the same rank as a previous one are not considered satisfactory.

In this selection process, we need to be cautious. If we encounter a candidate with the highest score too early, we risk not selecting anyone at all. Consequently, our function (Figure 9) demonstrates that the probability of selecting the best candidate approaches zero for high stopping points. Maximum of the function occurs early, specifically when stopping at 4.

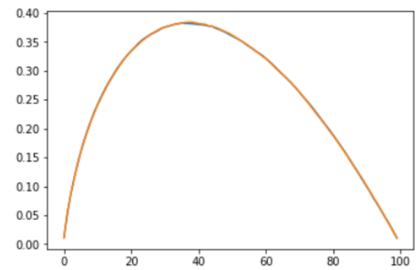
3.3 Normal distribution

In the final subsection of this chapter, we examine what happens if we do not choose a candidate's score uniformly, but instead use a normal distribution. This means that there will likely be more candidates with a score of 50 than with a score of 1 or 100 (for $n = 100$).

We decide that the normal distribution will be $N\left(\frac{n}{2}, \sigma\right)$. We also have to choose the variance (σ), as results can vary for different variances. For now, we will set aside the discussion of various variances and focus on what happens when the variances are $\frac{n}{5}$ and $\frac{n}{6}$. This allows us to approximately cover scores from 1 to 100. Naturally, when assigning scores, we only accept whole numbers to preserve the possibility that two candidates might have the same score.



(a) Probability function of success in finding the best candidate without being satisfied with someone of the same rank as a previous candidate.



(b) Probability function of success in finding the best candidate when we are satisfied with choosing someone of the same rank as a previous candidate.

Figure 10: Case where grades are normally distributed.

In Figure 10a, we observe that there is no significant difference between the two variances. We also notice that the maximum, determined by a normal distribution, is reached slightly earlier than in the standard case, at 34. Additionally, the probability is slightly higher, with a success rate of 38% instead of 37%.

We also examine the case where, instead of using the standard method, we use the average method for finding the best candidate. We set α to 1.9, as it has been identified as the optimal value in previous examples.

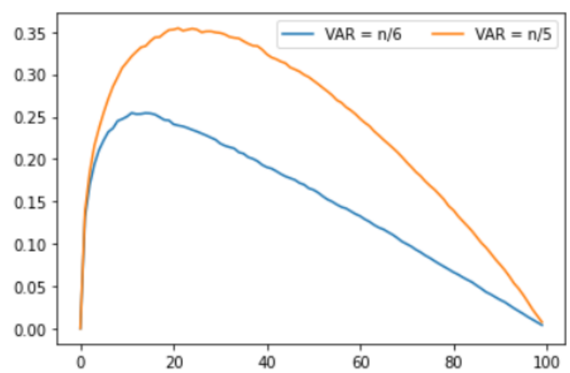


Figure 11: The average method with normally distributed ratings.

In Figure 11, we observe that, unlike in the previous case with normal distribution where the graphs showed virtually no difference, here they differ significantly for different variances. We also note that, with the normal distribution and the average method, we are more successful than with the standard method on normally distributed ratings.

4 Relaxing the Criteria for Success

In previous attempts, we were only satisfied if we ultimately chose the best candidate. With the right model, we successfully selected the best candidate nearly one-third of the time.

While this is an impressive result, if the alternative is that we find no partner in the remaining two-thirds of the cases, it is unlikely anyone would choose this strategy.

In reality, it is not true that we are only satisfied with the best choice while disregarding all others. Therefore, in this chapter, we explore other ways of evaluating the success of the model.

1. Relax the restriction to a top $p\%$. (This means we are satisfied if our chosen candidate is in the top $p\%$ of all candidates.)
2. Satisfaction is exactly proportional to the candidate's rating. (To determine success, we calculate the average rating of all selected candidates.)
3. A combination of the above two. (We are satisfied only with the top $p\%$ candidates and average their ratings.)

4.1 Top $p\%$ Strategy

In this subsection, we explore the first strategy described above.

We start by drawing two probability functions: one for the standard method and one for the average method (with $\alpha = 1.8$). We consider the selection of a candidate a success if they are in the top 10% of candidates. For our model with 100 people, this means that success is achieved if we end up with a candidate with a rating higher than 90.

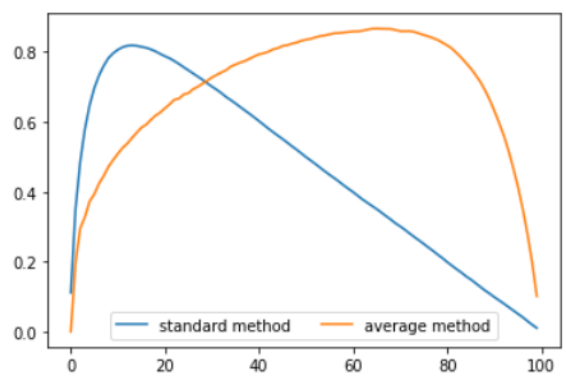


Figure 12: Probability function where success is defined as selecting a candidate in the top 10 %.

As we can see in Figure 12, the best method here is to use the average method and stop around 66. However, if you want to stop rejecting candidates before seeing half of the total candidates, you are better off using the standard method, which peaks after seeing the first 12 candidates. With both methods, if stopping at the optimal time, the success rate is slightly above 80%.

Figure 13 guided us in choosing $\alpha = 1.8$, as it emerged as the optimal value to use.

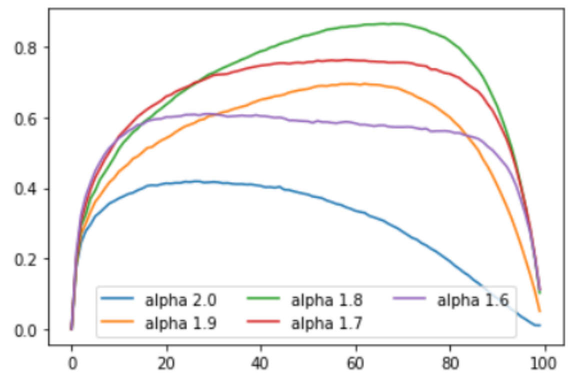
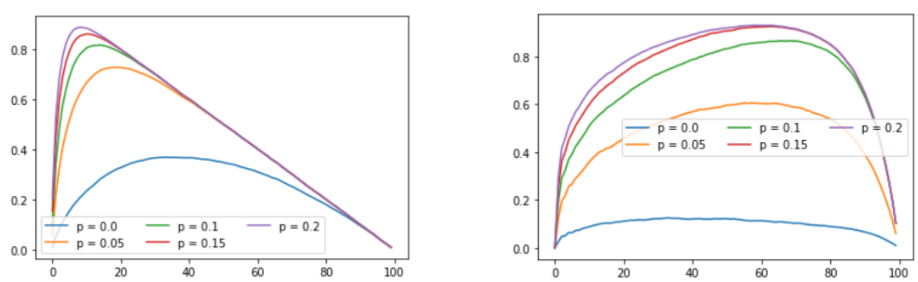


Figure 13: Average method where success is defined as selecting a candidate in the top 10%.

To conclude this subsection, we will examine how the probability function changes for different percentages (p) of success (Figure 14).



(a) Standard method.

(b) Average method.

Figure 14

4.2 Rating Strategy

In this subsection, we explore the second strategy described above. This strategy evaluates success by making satisfaction exactly proportional to the candidate's rating. To determine success, we calculate the average rating of all selected candidates.

We begin with the standard method. Figure 15 is fundamentally different from the previous figures, as the y-axis represents the average rating of the selected candidate (on a scale from 0 to 1) rather than the probability. Additionally, we distinguish between two scenarios: one where, if we have not selected a candidate by the end, we hire the last candidate, and another where, if the search is unsuccessful, we do not hire any candidate.

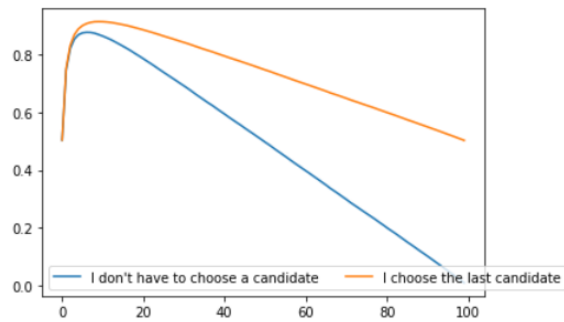
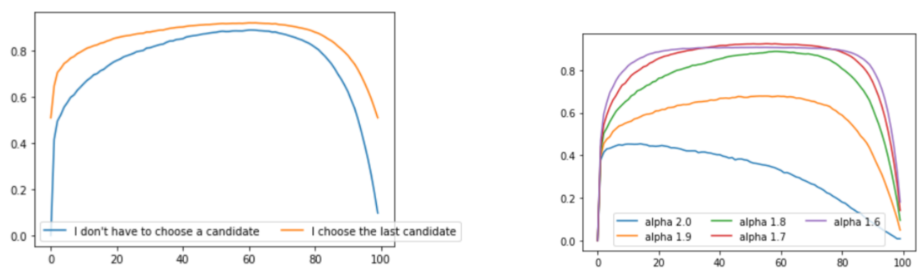


Figure 15: Standard method.

From Figure 15, we observe that if we want to hire the best candidate on average, we should stop at 6. This is drastically different from 37, which is the stopping point if we are only satisfied with the very best candidate.



(a) Average method.

(b) α test.

Figure 16

We continue with the average method for $\alpha = 1.8$ (Figure 16a), but we notice in Figure 16b, where we have graphs for different α values, that the maximum is reached at $\alpha = 1.7$.

5 Optimal Strategies for Hiring Two Candidates

In this section, we answer the following questions:

1. What criteria is best if we want to hire two people and maximize the probability of selecting the best two candidates?

2. What strategy is best if we can choose two people and we are happy if one of them is the best?

5.1 The best two candidates

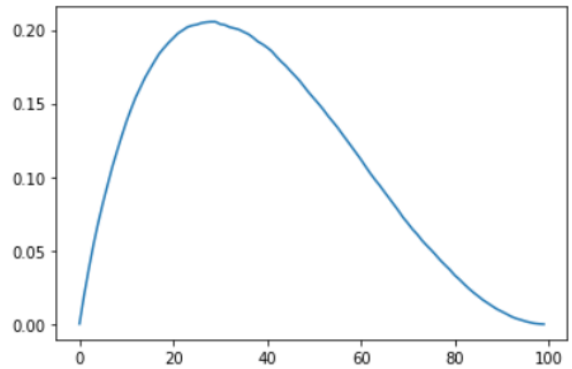


Figure 17: Standard method.

We begin with the first question posed above. In Figure 17, we applied the standard method (where we review all the candidates and reject everyone until a certain stopping point, after which we hire the first two candidates who are better than those we previously rejected). We observe that we are successful (i.e. hiring both the best and the second-best candidates) a little over 20% of the time if we stop at the 28th candidate. This means we need to stop earlier, which makes sense since we need to find two candidates instead of one.

When applying the average method to this problem, the results are less favourable.

5.2 One of them is the best

In this subsection, it is not mandatory to hire two candidates, but we have the option to do so. Our goal is to hire the best candidate.

We hypothesize that the probability of success will be higher than if we did not have the option to hire two candidates. Additionally, we predict that the maximum probability will be reached earlier than at 37, but later than at 28, since we want to hire just the best candidate.

In Figure 18, we see that our hypotheses were correct. The maximum is reached if we stop at 30, and the probability is greater than 50%.

When applying the average method to this problem, the results are also less favourable.

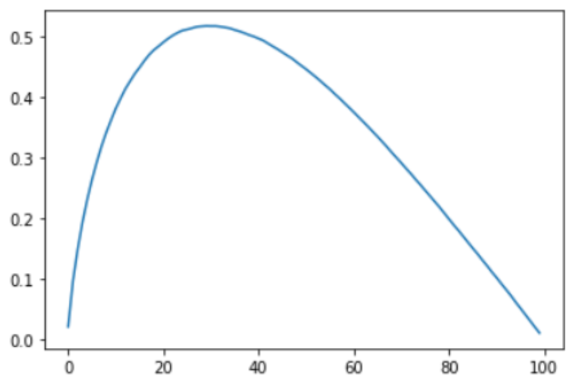


Figure 18: Standard method.

6 Optimal Strategies for Hiring k Candidates

In this section we examine optimal strategies for hiring at most k candidates.

6.1 The average rank

In this chapter, we will address the question: What criteria are best if we want to maximize the average rank of the selected candidates?

Note that the y-axis represents the average rating of the selected candidate (on a scale from 0 to 1) rather than the probability.

First, we try the standard method for different values of k. In Figure 19, we can see that the higher the k, the higher the function. This is logical because hiring 10 people makes it impossible to achieve an average rating of 100, which is theoretically possible with $k = 1$.

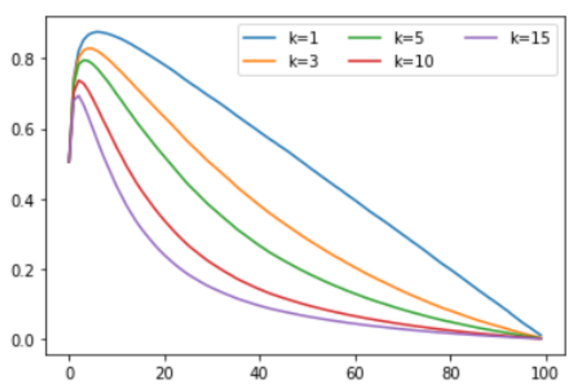
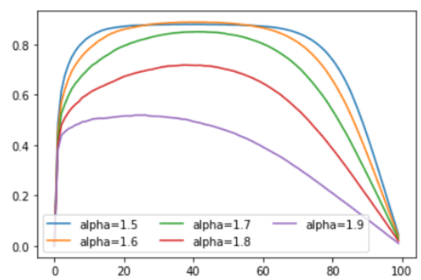


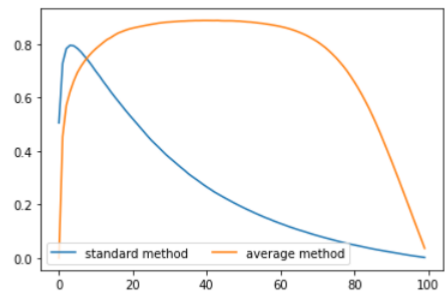
Figure 19: Standard method for different k.

In Figure 19, we see that as k increases, the maximum becomes smaller and shifts to the left. This indicates that we should stop earlier when k grows.

We decide to examine the case for $k = 5$ in more detail. Let's look at the graph for the average method for different alpha values when $k = 5$.



(a) α test.



(b) Standard method and average method for $k = 5$.

Figure 20

In Figure 20a, we determine that the best alpha is 1.6. Therefore, in Figure 20b, we compare both methods together.

6.2 Minimize the rank of the worst one

The question we examine in this chapter is: What criteria is best if we want to hire k people and want to minimize the rank of the worst one that was hired?

Same as before we run the standard method for different k (Figure 21). We notice that there is much less difference for different k.

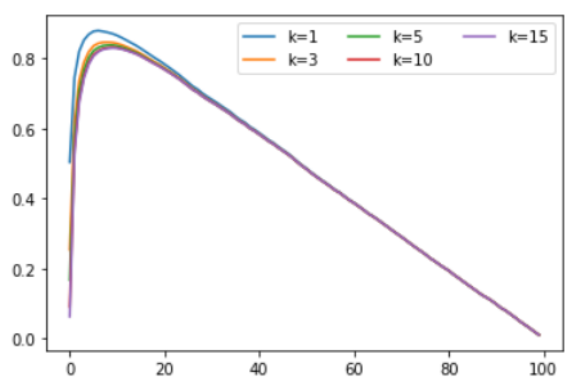
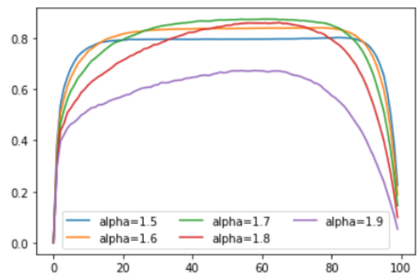
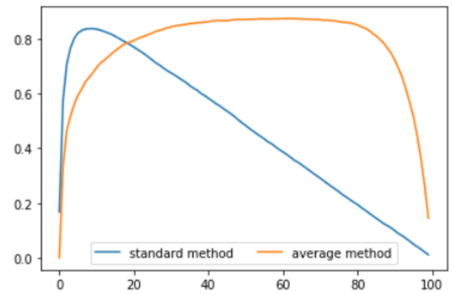


Figure 21: Standard method for different k.

We also conduct an alpha test, determining that the optimal alpha value to use is 1.7.



(a) α test.



(b) Standard method and average method for $k = 5$.

Figure 22

7 Conclusion

In this report, we explored various strategies for solving the Secretary Problem and its variants, focusing on the standard method and the average method. We examined different scenarios, including cases where candidates have the same rank and where ratings are normally distributed.

Key findings include:

1. **Standard Method Efficiency**: The optimal stopping point for the standard method changes depending on the criteria for success. For selecting the top candidate, stopping at the 37th candidate yields the highest success probability. However, when allowing for equally ranked candidates, the stopping point shifts to around the 39th candidate.

2. **Average Method Performance**: Using the average method with an alpha value of 1.9 generally resulted in better outcomes, especially when considering top 10% candidates or normally distributed ratings.

These insights underline the importance of adapting the selection strategy based on the specific criteria and distribution of candidate ratings. Future research could further investigate these strategies and explore additional variations in candidate distribution and selection criteria.

8 Questions and Answers

- *Is there a distinction when considering individuals of equal suitability, and does this affect our results?*

  There is a distinction when considering individuals of equal suitability, and this generally affects our results. (How much it affects the results depends on the system we decide to rank individuals.)

- *Moreover, does our outcome differ depending on whether we strictly prefer a superior candidate or are content with someone ranking equally to candidates we've previously encountered?*

  Yes, our outcome differs. (How much it differs depends on the system we decide to rank individuals.)

- *What criteria is best if we want to hire two people and want to maximize the probability of selecting the best two candidates?*

  If we want to hire two people and maximize the probability of selecting the best two candidates, it is best to use the standard method, stopping at 28% of the candidates.

- *What strategy is best if we can choose two people and we are happy if one of them is the best?*

  If we can choose two people and we are happy if one of them is the best, it is best to use the standard method, stopping at 30% of the candidates.

- *What criteria is best if we want to hire k people and want to minimize the rank of the worst one that was hired?*

  If we want to hire k people and want to minimize the rank of the worst one that was hired, it is best to use the average method with $\alpha = 1.7$ and stopping at 60% of the candidates.