**УДК 519.21**

*D. Melnyk, N. Shchestyuk, Y. Zakhariichenko*

DOI: 10.18523/2617-70808202540-45

# THE STOCHASTIC EXPERIMENT FOR SOME GENERALIZATIONS OF THE SECRETARY PROBLEM

*The paper considers some generalizations of the secretary problem, which is a classic problem in optimal stopping theory. We assume that the manager is somewhat more flexible and changes his goal to hire one of the top two best candidates. Another generalization is the searching the candidate of the top $\epsilon$ percent. It means that we agree to choose the candidate who differs from the absolute leader by no more than a specified amount ($\epsilon$ percent). Starting with classical secretary problem, we discuss in detail optimal solution for the secretary problem with the two best, following results in various sources. We review some approaches to this problem, which give the same optimal solution. After that we present our results of the stochastic experiments for both generalizations. By simulating numerous iterations of the candidate selection process, we estimate the probability of successfully selecting the best candidate. We demonstrate that with increasing $\epsilon$, the probability (rate) of success increases, and the number of candidates that were previously rejected decreases. Moreover, when we generate a list of candidates with random quality scores we use a random number generator to assign scores from different kind of distribution that reflects the quality of candidates.*

*We conclude that stochastic experiment based on Monte Carlo method is a powerful statistical technique that can be employed to analyze the different generalizations od Secretary Problem*

*Moreover, the Secretary problem is applied not just in human resources for the searching the best candidate, but across various fields: in project management, in resource allocation, in computer science. Thanks to the proposed approaches, the manager or other scientist gets a tools, which allows him to use a strategy that maximizes the chance of stopping with the two or more best candidate and take into account the different kind of distribution that reflects the quality of candidates.*

**Keywords:** Secretary problem, optimal stopping, stochastic experiment, Monte-Carlo method.

## Introduction

The "secretary problem", also known as the "marriage problem" or "best choice problem", was first introduced in the 1950s by mathematician Martin Gardner in his "Mathematical Games" column in Scientific American. Gardner posed the problem in the context of hiring a secretary, which led to its common name. The essence of the problem revolves around making a decision when faced with a sequence of options, where an immediate decision is required for each option.

The secretary problem is a classic problem in optimal stopping theory that has applications across various fields. There are a lot of fields where the secretary problem is applied. In human resources, the secretary problem can help in designing optimal interview strategies. Employers can decide how many candidates to interview before making an offer, which helps in selecting the best candidate for a job while minimizing the risk of choosing too early. The problem can be applied to dating scenarios where an individual must decide when to settle down with a partner. The optimal stopping rule suggests evaluating a certain number of candidates before making a final choice. The secretary problem can help in investment decisions. Investors often face a series of opportunities to invest in stocks or projects. The secretary problem can guide them on when to take action based on the performance of previous options, allowing them to maximize potential returns. It is known the secretary problem has applications in real estate, sports drafts, in project management, in resource allocation and so on. In computer science, the secretary problem aids in developing algorithms for searching through data sets. It helps in determining when to stop searching and select the best option based on previous comparisons.

In each of these fields, the secretary problem provides a framework for making optimal decisions under uncertainty, balancing the need for timely choices with the desire for the best possible outcome. Today, the secretary problem remains an active area of research. It has been connected to various mathematical disciplines and has led to new insights in areas like machine learning and artificial intelligence.

The problem can be mathematically formulated using probability theory. The optimal strategy involves rejecting a predefined number of candidates (often around 37 percents of the total pool) before starting to accept the next best candidate you encounter. This strategy maximizes the probability of selecting the best candidate. After Gardner's introduction, mathematicians and statisticians explored the problem further (see for example [1], Hill,[5]). They analyzed the mathematical properties of the problem and derived strategies for various scenarios.

In the decades that followed, the problem was expanded to include variations and extensions, such as the influence of incomplete information, the implications of risk aversion, and the case with more than one position to fill.

In our paper we consider some generalizations of this problem. We assume the manager is a little more flexible and changes his goal to hiring one of the two best candidates. Another generalization is in determining the probability that we end with a candidate in the top $\epsilon$ percent. It means that we agree to choose the candidate who differs from the absolute leader by no more than a specified amount ($\epsilon$ percent) We would like to figure out how many people would we interview in this case and what would be our probability of success.

We discuss in detail one in the section 2. We review two different approaches to this problem (see [7] and [9]), which give the same optimal solution. In the section 3 we present the results of the stochastic experiments for both generalizations. By simulating numerous iterations of the candidate selection process, we estimate the probability of successfully selecting the best candidate. Moreower, when we generate a list of candidates with random quality scores we use a random number generator to assign scores from different kind of distribution that reflects the quality of candidates.

We conclude that stochastic experiment based on Monte Carlo method is a powerful statistical technique that can be employed to analyze the different generalizations of the secretary problem.

## Optimal solution for the secretary problem with the two best

***Classical secretary problem.*** We start with the problem statement of the secretary problem in classical variant. The basic problem can be stated as follows (see for example [1], [5]):

1. There is a single position to fill.
2. There is a fixed and known number n of applicants for a single position.
3. The applicants (candidates, secretaries) can be ranked from best to worst unambiguously.
4. The applicants are interviewed sequentially in random order, with each order being equally likely.
5. Immediately after an interview, the interviewed applicant is either accepted or rejected, and the decision is irrevocable.
6. The decision to accept or reject an applicant can be based only on the relative ranks of the applicants interviewed so far.

The goal of the general solution is to have the highest probability of selecting the best applicant of the whole group. This is the same as maximizing the expected payoff, with payoff defined to be one for the best applicant and zero otherwise.

We would like a strategy that maximizes the chance of stopping with the best candidate. The following strategy $S_k$ is proposed for example in [1],[5]. Under it, the manager rejects the first $k$ applicants and defines the best applicant among these $k$ applicants. Then he selects the first subsequent applicant that is better than applicant from first $k$. Suppose the best applicant is at position $m$. Then our strategy $S_k$ results in our selecting the best applicant if and only if there is no one among people $k+1, ..., m-1$ who is better than the best person in the first $k$. Thus, if the best applicant is at position $m$ then we select the best person precisely when the best person among the first $m-1$ is in the first $k$ people. The probability the best of the first $m-1$ is in the first $k$ is just $\frac{k}{m-1}$. Therefor for an arbitrary rejecting first $k$ candidates, probability strategy $S_k$ wins is

$$Prob(S_k\, wins) = \sum_{m=k+1}^{n} Prob(win|\, best\, at\, m) *$$

$$* Prob(best\, at\, m) = \sum_{m=k+1}^{n} \frac{k}{m-1}\frac{1}{n} =$$

$$= \frac{k}{n}\sum_{m=k+1}^{n} \frac{1}{m-1} = \frac{k}{n}(\sum_{m=1}^{n-1}\frac{1}{m} - \sum_{m=1}^{k-1}\frac{1}{m}).$$

If we take into account harmonic number $l-th$ is approximately $\ln l$ for $l$ large, then

$$Prob(S_k\, wins) = \frac{k}{n} ln(\frac{n-1}{k-1}) = \frac{ln(\frac{n-1}{k-1})}{\frac{n}{k}}.$$

For $n$ and $k$ large, we may replace $n-1$ with $n$ and $k-1$ with k. Thus we are trying to optimize $g(x) = \frac{lnx}{x}$, where $1 \le x = \frac{n}{k} \le n$.

To find where a function is largest, we check the critical and endpoints. Letting $g(x) = \frac{lnx}{x}$, we see the endpoints give $g(n) = \frac{lnn}{n}$. As

$$g'(x) = \frac{1-lnx}{x^2},$$

$g'(x) = 0$ implies $lnx = 1$ or $x = e$. Thus the optimal $k$ is about $\frac{n}{e}$ and the probability we end up with the best is approximately

$$\frac{lne}{e} = \frac{1}{e} \approx 36,8\%.$$

***Secretary problem with the two best.*** The next question would be: what if the manager is a little more flexible and changes his goal to hiring one of the two best candidates? In this case what strategy gives the largest probability that we end up with either the best or second best candidate? The answer turns out to be over 50 percent! We'll denote the location of the best and second best candidates as $m_1$ and $m_2$. We assume again we have a simple strategy of inter- viewing the first $k$ candidates, and afterwards discuss some variants (see [7], p.36-37 in more detail).

If both $m_1 \leq k$ and $m_2 \leq k$ we always lose, because we skip the first k candidate.

If the best is in the first $k$ and the second is not, we lose unless the second best happens to be in the final position. Thus the probability we win in this case is $\frac{k}{n}\frac{1}{n}$.

If the second best candidate is in the first $k$ and the best is not, we automatically win with this strategy! The probability of this happening is $\frac{k}{n}\frac{n-k}{n}$. If $k$ is of the same order of magnitude as $n$, then this will be a significant probability of success.

After that we analyze the case when the top two candidates are not in the first $k$. Denote $A =$ "$S_k\, wins$" for our case, $H_{m_1,m_2} =$ = "$best\, at\, m_1,\, second\, at\, m_2$" The probability of success $A$ in this case is

$$\sum_{m_1=k+1}^{n-1} \sum_{m_2=m_1+1}^{n} Prob(A|H_{m_1,m_2})Prob(H_{m_1,m_2})$$

$$= \sum_{m_1=k+1}^{n-1} \sum_{m_2=m_1+1}^{n} \frac{k}{m_1-1}\frac{2}{n(n-1)} =$$

$$= \frac{2k}{n(n-1)} \sum_{m_1=k+1}^{n-1} \frac{1}{m_1-1}(n-m_1) =$$

$$= \frac{2k}{n}(ln(\frac{n-2}{k-1}) - 1 + \frac{k}{n}).$$

In the last expression to evaluate the sum − we wrote $n-m_1$ as $n-1-(m_1-1)$ and took into account harmonic number. Combining all the different probabilities, we see the probability of winning is

$$Prob(S_k\, wins) = \frac{k}{n^2} + \frac{k}{n}(1-\frac{k}{n}) + \frac{2k}{n}(ln(\frac{n}{k}) - 1 + \frac{k}{n}).$$

As $k$ will be of the same size as $n$, the $k/n^2$ term is negligible and if we let $x = n/k$ we just need to optimize the function

$$g(x) = \frac{1}{x}(1-\frac{1}{x}) + \frac{2}{x}(lnx - 1 + \frac{1}{x})$$

To simplify calculus we let $y = 1/x = k/n$ and get

$$h(y) = y(1-y) + 2y(-lny - 1 + y)$$

Then the derivatives are

$$h'(y) = -3 + 2y - 2lny,$$

$$h''(y) = 2(1 - \frac{1}{y}).$$

Numerically solving $h'(y) = 0$ implies $y \approx 0,3071$ and we can check this is a maximum. Substituting this into our formula, we find the probability of winning with this strategy is about 0.51239. It means we have greater than a 50% chance of getting one of the top two candidates!

***Remark.*** We would like to notice that there is an similar exploration to solve the secretary problem with two best, which the reader can find in [9].

## Stochastic experiment for the secretary problem

In this section we present the results of the stochastic experiments for the secretary problem obtained by Monte Carlo approach. Monte Carlo method is a powerful statistical technique that can be employed to various problems: option pricing [3], [10], diffusion modeling[2], portfolio optimization [11] and other.

By simulating numerous iterations of the candidate selection process, you can estimate the probability of successfully selecting the best candidate and evaluate the effectiveness of different strategies. For Monte Carlo Simulation we use the next steps [8], [4]:

1. Define the parameters: a number of candidates (n) and a number of simulations (N).

2. Simulate candidate quality. For each iteration, generate a list of candidates with random quality scores. We have done it using a random number generator to assign scores from a uniform distribution and then some other distributions that reflects the quality of candidates.

3. Implement the Selection Strategy. First we choose parameter k. Then suppose the manager's strategy is (like in the original version of the secretary problem) to reject the first k candidates and then hire the next candidate who is better than all candidates seen thus far.

4. Run the simulation. For each of the N iterations we generate the candidate quality scores (n) and apply the stopping rule to decide whether to accept or reject each candidate. Then for the first generalization we track whether the selected candidate is the best or second best candidate among

all those presented. For the second generalization we track whether the selected candidate is in the top $\epsilon$ percent.

5. Collect Results. Record the number of times the selected candidate is the best or second best (for one generalization) or is in the top $\epsilon$ percent (for the second generalization) across all iterations. This will allow us to calculate the success rate of the strategy.

6. Analyze the outcomes. Calculate the proportion of successful selections (number of successful selections / total iterations). Compare the success rates of different strategies, including the classic optimal stopping rule and any alternative methods we have tested. Visualize the results using graphs, to illustrate the distribution of outcomes.

First, we apply the Monte Carlo simulation for Secretary problem with the two best candidates. Using Python, we visualize the success rate of the strategy and the theoretical probability of the optimal stopping for different k.

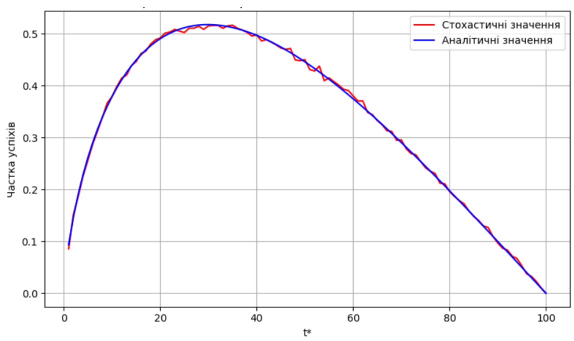


**Figure 1.** The compare theoretical probability and the success rate of the optimal strategy for the two best.

Then we visualize the Monte Carlo simulation for Secretary problem with the top $\epsilon$ percent (for the second generalization). On the picture we can see the dependence of the success rate from $\epsilon$ for uniform distribution.

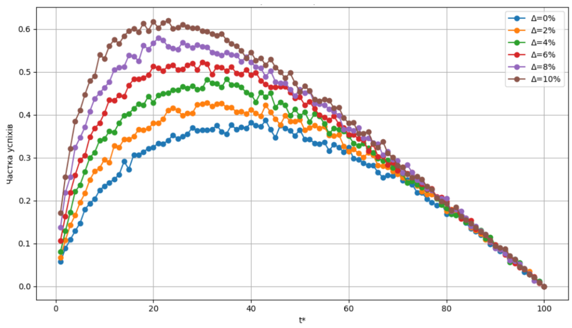


**Figure 2.** Graphical representation for the dependence of the success rate from $\epsilon$ for uniform distribution.

From the graph in Fig. 2 we clearly see that with increasing $\epsilon$, the probability (rate) of success increases, and the number of candidates k that were previously rejected decreases. For example, for $\epsilon = 2\%$ the maximum value of the success rate is $45,77$ and $k = 31$, for $\epsilon = 4\%$ the maximum is $52,42$ and $k = 25$, for $\epsilon = 10\%$ these values are $68,55$ and $k = 19$ respectively.

The next step is to learn the behavior of the rate of successful experiments for a normal distribution. Compared to a uniform distribution, we have a less rapid growth with increasing degree. However, the probability of success still increases, and the number of candidates that need to be reviewed and rejected decreases.

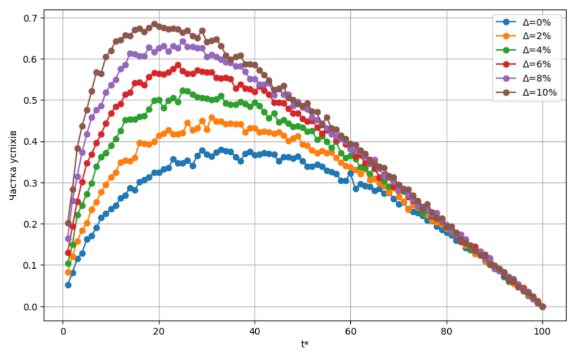


**Figure 3.** Graphical representation for the dependence of the success rate from $\epsilon$ for normal distribution

Similarly, let us examine the behavior of the probability of success for the exponential distribution.

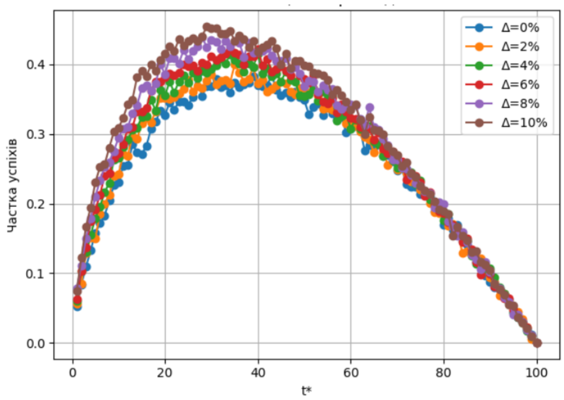


**Figure 4.** Graphical representation for the dependence of the success rate from $\epsilon$ for exponential distribution

Compared to the previous graphs, we immediately notice a very weak increase in the proportion of successful experiments with increasing $\epsilon$.

## Conclusion

In the paper some generalizations of the classical secretary problem were considered to be closer to the conditions of the real world. The analytical analysis and the results of the stochastic experiments showed that if the selection criteria are made softer, choosing not only the best candidate, but also others beside him, the probability of a successful choice will be significantly increased. Thus, if a manager is some flexible and ready to make a concession of up to 10%, the probability of success increases to 68.55%, which is almost twice as much, compared to the value of 36.8% for the classical option. In this case the number of candidates that need to be rejected, also decreases from 37 to 19 (for 100 candidates). At the same time, the modification, which allows choosing not only the best, but also the second best candidate, also gives a significant increase in success. With a smaller number of candidates considered, namely 30%, the probability of success is 51.24%. The results are quite applicable, and can be used for real-world situations in which it is necessary to make irreversible, after-the-fact decisions.

Moreover, in stochastic experiment of the secretary problem we apply a random number generator to assign scores to candidates not only from a uniform distribution, but from some other distributions that reflect the quality of candidates. We assume that the normal distribution is naturally more suitable for describing the distribution of candidate scores.

### *References*

1. J. Bearden and R. Murphy, Uncertainty and Risk. **41**, 187–205 (2007).
2. K. Boluh and N. Shchestyk, Mohyla Mathematical Journal. **3**, 25-30 (2020).
3. V. Pauk, O. Petrenko and N. Shchestyuk, Mohyla Mathematical Journal. **5**, 38–45 (2022).
4. L. Fainzilberg, Control Systems and Computers. **1**, 11–21 (2019).
5. T. Ferguson, Statistical Science. **4**, 282–289 (2003).
6. T. Hill, American Scientist. **97** (2), 126–133 (2009).
7. S. Miller, Williams College, **34–38** (2009).
8. R. Rubinstein and D. Kroese, New York, **432** (2016).
9. M. Spivey, *A Narrow Margin.* (2017). https://mikespivey.wordpress.com/2017/07/19/toptwo/(2017).
10. N. Shchestyuk and S. Tyshchenko, Bulletin of the Taras Shevchenko National University of Kyiv. Physics and Mathematics. 2, 85–92 (2021).
11. G. Solomanchuk and N. Shchestyuk, Mohyla Mathematical Journal. 4, 28–33 (2021).

*Мельник Д. Л., Щестюк Н. Ю., Захарійченко Ю. О.*

# СТОХАСТИЧНИЙ ЕКСПЕРИМЕНТ ДЛЯ ДЕЯКИХ УЗАГАЛЬНЕНЬ ПРОБЛЕМИ СЕКРЕТАРЯ

У статті розглядаються деякі узагальнення задачі секретаря, яка є класичною задачею в теорії оптимальної зупинки. Ми припускаємо, що менеджер дещо гнучкіший і змінює свою мету так, щоб найняти одного з двох найкращих кандидатів. Іншим узагальненням є пошук кандидата серед претендентів з верхнього $\epsilon$ квантиля. Це означає, що ми погоджуємося вибрати кандидата, який відрізняється від абсолютного лідера не більше ніж на задану величину ($\epsilon$ відсотків). Починаючи з класичної задачі секретаря, ми детально розглядаємо оптимальне рішення задачі пошуку секретаря з двома найкращими кандидатами, дотримуючись результатів, наведених у різних джерелах. Ми наводимо пояснення з різних джерел, які дають однакове оптимальне рішення. Після цього ми презентуємо результати наших стохастичних експериментів для обох узагальнень. Симулюючи чисельні ітерації процесу відбору кандидатів, ми оцінюємо ймовірність успішного вибору найкращого кандидата. Ми демонструємо, що зі збільшенням $\epsilon$ ймовірність (коефіцієнт) успіху зростає, а кількість кандидатів, які раніше були відхилені, зменшується. Більше того, коли ми генеруємо список кандидатів із випадковими оцінками якості, ми використовуємо генератор випадкових чисел для присвоєння оцінок з різних типів розподілу, що відображає якість кандидатів.

Ми робимо висновок, що стохастичний експеримент, заснований на методі Монте-Карло, є потужним статистичним методом, який можна використовувати для аналізу різних узагальнень задачі секретаря.

Необхідно зауважити, що задача секретаря застосовується не лише в управлінні персоналом для пошуку найкращого кандидата, а й у різних галузях: в управлінні проєктами, у розподілі ресурсів, в інформатиці. Завдяки запропонованим підходам, менеджер або інший фахівець отримує інструменти, які дозволяють йому використовувати стратегію, що максимізує ймовірність зупинитися на найкращому кандидаті у випадку, якщо його задовольнить вибір одного з двох або більше топ-кандидатів, і враховує різні типи розподілу, що відображає якість кандидатів.

**Ключові слова:** проблема секретаря, оптимальна зупинка, стохастичний експеримент, метод Монте-Карло.

*Матеріал надійшов 01.08.2025*