

# A VARIANT OF THE SECRETARY PROBLEM: THE BEST OR THE WORST

L. BAYÓN, J. GRAU, A. M. OLLER-MARCÉN, M. RUIZ, AND P.M. SUÁREZ

ABSTRACT. We consider a variant of the secretary problem in which the candidates state their expected salary at the interview, which we assume is in accordance with their qualifications. The goal is for the employer to hire the best or the worst (cheapest), indifferent between the two cases. We focus on the complete information variant as well as on the cases when the number of applicants is a random variable with a uniform distribution  $U[1, n]$  or with a Poisson distribution of parameter  $\lambda$ . Moreover, we also study two variants of the original problem in which we consider payoffs depending on the number of conducted interviews.

Keywords: Secretary problem, Combinatorial Optimization  
AMS 2010 Mathematics Subject Classification 60G40, 62L15

## 1. PRELIMINARIES

The optimal stopping problem, known as the classical *secretary problem*, can be stated as follows: an employer is willing to hire the best secretary out of  $n$  rankable candidates for a position. The candidates are interviewed one by one in a random order. A decision about each particular candidate is to be made immediately after the interview. Once rejected, a candidate cannot be recalled. During the interview, the administrator can rank the applicant among all applicants interviewed so far, but is unaware of the quality of yet unseen applicants. The question is about the optimal strategy (stopping rule) to maximize the probability of selecting the best candidate. The secretary problem is one of many names for a famous problem of optimal stopping theory. It is also known as the marriage problem, the sultan's dowry problem, the fussy suitor problem and others. This problem has been extensively studied in the fields of applied probability, statistics, and decision theory and has been considered by many authors (see [6], [7] and [12] for an extensive bibliography). It can also be posed as a decision-taking problem in a game with the following rules:

- (1) You want to choose one object.
- (2) The number of objects is known.
- (3) The objects appear sequentially in random order.
- (4) The objects are rankable.
- (5) Each object is accepted or rejected before the next object appears.
- (6) The decision depends only on the relative ranks.
- (7) Rejected objects cannot be recalled.
- (8) Payoff: You only win if the best object is chosen.

This problem has a very elegant solution. Dynkin [5] and Lindley [10] independently proved that the best strategy consists in observing roughly  $n/e$  of the candidates and then choosing the first one that is better than all those observed

so far. This strategy returns the best candidate with a probability of at least  $1/e$ , this being its approximate value for large values of  $n$ . This well-known solution was refined by Gilbert and Mosteller in [8], where they show that  $[(n - \frac{1}{2})e^{-1} + \frac{1}{2}]$  is a better approximation than  $[n/e]$ , although the difference is never greater than 1. There are other variants of the secretary problem that also have simple, elegant solutions. In the “postdoc” problem, for instance, the desire to pick the best is replaced by the desire to pick the second best (because, according to Vanderbei [13], the “best” will go to Harvard). For this problem, the probability of success for an even number of applicants is exactly  $\frac{n}{4(n-1)}$ . This probability tends to  $1/4$  as  $n$  tends to infinity, illustrating the fact that it is easier to pick the best than the second best.

If the number of candidates is unknown, the observer faces an additional risk. If he rejects any candidate, he may then discover that it was the last one, in which case he receives nothing at all. In [11] the case in which the number of candidates is a random variable with a discrete uniform distribution  $U[1, n]$  was studied. In this case, the cutoff value for large  $N$  is approximately  $Ne^{-1/2}$  and the probability of success is  $2e^{-2}$ . This same paper also tackles the problem assuming that the number of candidates is Poisson distributed with parameter  $\lambda$ .

Another interesting variant of the problem was introduced by Bearden [1], considering a positive payoff that increases with the number interviews. In this case, the optimal cutoff value is not proportional to the number of candidates and it is the square root of such number.

In the present paper, we consider a variant of the secretary problem in which the candidates state their expected salary at the interview, which we assume is in accordance with their qualifications. The goal is for the employer to hire the best or the worst (cheapest), indifferent between the two cases. In other words, it is the classical problem in which rule 8 above is replaced by: “Payoff: You only win if the best or the worst is chosen”. The problem is studied in Section 2 in its version with full information regarding the number of candidates. In Section 3, we consider the problem if the number of candidates is a random variable with a uniform distribution  $U[1, n]$ . In Section 4, the number of candidates is a random variable with a Poisson distribution of parameter  $\lambda$ . In Section 5, we consider payoffs other than the classical binary payoff that are dependent on the number of interviews carried out. In one case, we consider payoff proportional to the number of interviews and in another, payoff proportional to the number of interviews not carried out.

The techniques used throughout the paper are suitable adaptations of those employed in the study of the solution to the classical secretary problem (limit of Riemann integrals, expansions at infinity, etc.) to obtain asymptotic formulas of the type  $\alpha \cdot n + \beta + o(1)$  for the optimal cutoff value. In some proofs related to the calculation of some limits, details have been omitted and their correctness has been tested with the powerful symbolic computation tool Mathematica. Several of the results are formulated in terms of  $W_0$  and  $W_{-1}$  functions, the lower real branches of the Lambert  $W$  function, respectively; see Corless et al. [3] for further details.

## 2. A VARIATION OF THE SECRETARY PROBLEM: THE BEST OR THE WORST

In this section, we study a variation of the classical secretary problem considering that the success of the player resides in choosing the worst or the best object,

indifferent between the two cases. In this case, as in the secretary problem, the strategies to consider consist in rejecting a certain number of objects (cutoff value) and then taking the first object that is better or worse than all those previously rejected. After rejecting  $r$  objects, the probability of successfully choosing the  $k$ -th inspected object,  $\mathcal{P}(k)$ , is that of this object being the best or the worst and that of not having interrupted the inspections by previously choosing another object. Given that both are independent events, the probability of success will be the product of the probability of both events. The probability that the  $k$ -th object to be seen is the best or worst of all the objects is  $2/n$  while the probability that the best and the worst among those seen prior to the  $k$ -th is among the first  $r$  objects (i.e. the inspections have not ceased before seeing the  $k$ -th object) is

$$\frac{\binom{r}{2}}{\binom{k-1}{2}} = \frac{r(r-1)}{(k-1)(k-2)}.$$

Thus, provided  $1 < r < k \leq n$ , we have that

$$\mathcal{P}(k) = \frac{2}{n} \frac{\binom{r}{2}}{\binom{k-1}{2}}$$

and the probability of success, rejecting the first  $r$  objects, and accepting afterwards the first one which is either better than all the previous ones or worse than all the previous ones, is

$$P_n(r) := \sum_{k=r+1}^n \mathcal{P}(k) = \frac{2\binom{r}{2}}{n} \sum_{k=r+1}^n \frac{1}{\binom{k-1}{2}}.$$

Note that for  $n > r \in \{0, 1\}$ , it is straightforward to see that the probability of success is

$$P_n(0) = P_n(1) = \frac{2}{n}.$$

**Theorem 1.** *Given a positive integer  $n > 2$ , consider the function*

$$P_n(r) := \frac{2\binom{r}{2}}{n} \sum_{k=r+1}^n \frac{1}{\binom{k-1}{2}}$$

*defined for every integer  $r \in [2, n-1]$ . Let us denote by  $\mathcal{M}(n)$  the value for which the function  $P_n$  reaches its maximum. Then,*

- i)  $P_n(r) = \frac{2r(n-r)}{n(n-1)}$ .
- ii)  $\mathcal{M}(n) = \lfloor n/2 \rfloor$ .
- iii) *The maximum of  $P_n$  is:*

$$\mathfrak{p}_n := P_n(\mathcal{M}(n)) = \frac{\lfloor \frac{1+n}{2} \rfloor}{2\lfloor \frac{1+n}{2} \rfloor - 1} = \begin{cases} \frac{n}{2(n-1)}, & \text{if } n \text{ is even;} \\ \frac{n+1}{2n}, & \text{if } n \text{ is odd.} \end{cases}$$

*Proof.*

- i) First of all, observe that

$$P_n(r) = \frac{2\binom{r}{2}}{n} \sum_{k=r+1}^n \frac{1}{\binom{k-1}{2}} = \frac{r(r-1)}{n} \sum_{k=r+1}^n \frac{2}{(k-1)(k-2)}.$$

Now, using telescopic sums, we have that

$$\begin{aligned} \sum_{k=r+1}^n \frac{2}{(k-1)(k-2)} &= 2 \sum_{k=r+1}^n \left( \frac{1}{k-2} - \frac{1}{k-1} \right) = 2 \left( \frac{1}{r-1} - \frac{1}{n-1} \right) = \\ &= \frac{2(n-r)}{(r-1)(n-1)} \end{aligned}$$

and the result follows.

- ii) Since  $P_n(r) = -\frac{2}{n(n-1)}r^2 + \frac{2}{(n-1)}r$  is the equation of a parabola in the variable  $r$ , it is clear that

$$\mathcal{M}(n) = \min \{r \in [2, n-1] : P_n(r) \geq P_n(r+1)\}.$$

Now,

$$P_n(r+1) - P_n(r) = \frac{2}{n(n-1)}(n-2r-1)$$

so it follows that

$$P_n(r+1) - P_n(r) \leq 0 \Leftrightarrow (n-2r-1) \leq 0 \Leftrightarrow r \geq \frac{n-1}{2}.$$

Consequently,

$$\mathcal{M}(n) = \min \left\{ r \in [2, n-1] : r \geq \frac{n-1}{2} \right\} = \lfloor n/2 \rfloor$$

as claimed.

- iii) It is enough to apply the previous result.

If  $n$  is even, then  $n = 2N$  and

$$P_n(\mathcal{M}(n)) = P_n(N) = \frac{2N(n-N)}{n(n-1)} = \frac{2N^2}{2N(2N-1)} = \frac{N}{2N-1}$$

Moreover, in this case

$$\left\lfloor \frac{1+n}{2} \right\rfloor = \left\lfloor \frac{1+2N}{2} \right\rfloor = N$$

so it follows that

$$P_n(\mathcal{M}(n)) = \frac{N}{2N-1} = \frac{N}{2 \left\lfloor \frac{1+n}{2} \right\rfloor - 1}$$

as claimed.

Otherwise, if  $n$  is odd, then  $n = 2N+1$  and

$$P_n(\mathcal{M}(n)) = P_n(N) = \frac{2N(n-N)}{n(n-1)} = \frac{2N(2N+1-N)}{(2N+1)2N} = \frac{N+1}{2N+1}.$$

In this case

$$\left\lfloor \frac{1+n}{2} \right\rfloor = \left\lfloor \frac{1+2N+1}{2} \right\rfloor = N+1$$

so we also have that

$$P_n(\mathcal{M}(n)) = \frac{N+1}{2N+1} = \frac{N+1}{2 \left\lfloor \frac{1+n}{2} \right\rfloor - 1}$$

and the proof is complete.  $\square$

It can be deduced from this result that, for  $n > 2$ , the strategy that maximizes the probability of success consists in rejecting the first  $\lfloor \frac{n}{2} \rfloor$  (optimal cutoff value) objects and then choosing the first to appear that is better or worse than all the preceding ones. The probability of success following this strategy is

$$\frac{\lfloor \frac{1+n}{2} \rfloor}{2\lfloor \frac{1+n}{2} \rfloor - 1}$$

In the cases  $n \in \{1, 2\}$ , it is evident that an optimal cutoff value is  $r = 0$ , i.e. to accept the first object that we are shown; the value of the probability of success being 1 in both cases:  $\mathbf{p}_1 = \mathbf{p}_2 = 1$

### 3. UNKNOWN NUMBER OF CANDIDATES WITH UNIFORM DISTRIBUTION $U[1, n]$

In this section, we analyze the case in which the number of objects is unknown, but it is known that it can be anywhere between 1 and  $n$  with equal probability. As we saw in Theorem 1 i), the probability of success with  $i > 1$  objects while rejecting the first  $r$  observed objects is  $P_i(r) = \frac{2r(i-r)}{i(i-1)}$ . Hence, the probability of success in this case will be the mean probability of the  $n$  equally probable possible scenarios with probability  $1/n$ . That is, the probability of success when rejecting the  $r$  first objects before comparing with the previous ones and with an unknown number of candidates that follows a uniform distribution  $U[1, n]$  will be:

$$P(r, n) := \sum_{i=r+1}^n \frac{P_i(r)}{n} = \sum_{i=r+1}^n \frac{2r(i-r)}{i(i-1)n}$$

We denote by  $\mathbf{m}(n)$  the value for which  $P(\cdot, n)$  reaches its maximum value in  $[1, n]$  and by  $\mathbf{P}(n)$ , the maximum value it reaches, i.e.  $\mathbf{P}(n) := P(\mathbf{m}(n), n)$ . We shall first prove that the probability of success in the game,  $\mathbf{P}(n)$ , is strictly decreasing for  $n > 1$ , where clearly  $\mathbf{P}(1) = \mathbf{P}(2) = 1$  and  $\mathbf{m}(1) = \mathbf{m}(2) = 0$ . The technique used in the proof is an adaptation of the strategy-stealing argument: starting from any given strategy in the game with a number of objects following a uniform distribution  $[1, n + 1]$ , two strategies are developed for the game with a number of objects following a uniform distribution  $[1, n]$ , one or other of which necessarily improves the initial probability of success. More specifically, we show that if  $r$  is the optimal cutoff value for the problem associated with a uniform distribution  $[1, n + 1]$ , then a higher probability of success is achieved in the problem associated with a uniform distribution  $[1, n]$  with either  $r$  or  $r - 1$  as the cutoff value.

**Lemma 1.** *Consider the function*

$$P(r, n) := \sum_{i=r+1}^n \frac{2(i-r)r}{i(i-1)n}$$

*defined for every pair of integers  $(r, n)$  with  $0 < r < n$ . Then, one of the following inequalities holds:*

$$\begin{aligned} P(r, n + 1) &< P(r - 1, n), \\ P(r, n + 1) &< P(r, n). \end{aligned}$$

*Proof.* Let us denote  $H(n, r) := \sum_{i=r}^{n-1} \frac{1}{i}$ . Then we have that

$$\begin{aligned} P(r, n) &= \sum_{i=r+1}^n \frac{2(i-r)r}{i(i-1)n} = \frac{2r}{n} \sum_{i=r+1}^n \left( \frac{r}{i} + \frac{1}{i-1} + \frac{r}{i-1} \right) = \\ &= \frac{2r}{n} \left[ \left( \frac{r}{n} - 1 \right) + \sum_{i=r+1}^n \frac{1}{i-1} \right] = \frac{2r}{n} \left[ \left( \frac{r}{n} - 1 \right) + H(n, r) \right]. \end{aligned}$$

In the same way we find that

$$\begin{aligned} P(r, n+1) &= \frac{2r}{n+1} \left[ \left( \frac{r}{n+1} - 1 + \frac{1}{n} \right) + H(n, r) \right], \\ P(r-1, n) &= \frac{2(r-1)}{n} \left[ \left( \frac{r-1}{n} - 1 + \frac{1}{r-1} \right) + H(n, r) \right] \end{aligned}$$

As a consequence, it follows that

$$\begin{aligned} P(r, n+1) - P(r-1, n) &= \frac{2(1+n-r)}{n(n+1)} \left[ -\frac{(1+2n)(1+n-r)}{n(n+1)} + H(n, r) \right], \\ P(r, n+1) - P(r, n) &= \frac{2r}{n(n+1)} \left[ \frac{2n(1+n-r) - r}{n(n+1)} - H(n, r) \right] \end{aligned}$$

Thus, if we assume that both

$$P(r, n+1) - P(r-1, n) \geq 0 \text{ and } P(r, n+1) - P(r, n) \geq 0,$$

it follows that

$$\begin{aligned} A &:= H(n, r)n(n+1) - (1+2n)(1+n-r) \geq 0, \\ B &:= -(H(n, r)n(n+1)) + 2n(1+n-r) - r \geq 0. \end{aligned}$$

But in this case  $0 \leq A + B = -1 - n < 0$ , a contradiction, and hence the result.  $\square$

**Corollary 1.** *With the previous notation,  $\mathbf{P}(n+1) < \mathbf{P}(n)$  for every  $n > 1$ .*

*Proof.* If  $n > 1$ ,

$$\begin{aligned} \mathbf{P}(n+1) &= P(\mathbf{m}(n+1), n+1) < \max\{P(\mathbf{m}(n+1), n), P(\mathbf{m}(n+1) - 1, n)\} \leq \\ &\leq P(\mathbf{m}(n), n) = \mathbf{P}(n), \end{aligned}$$

where the first inequality follows from Lemma 1 and the second holds by the definition of  $\mathbf{m}(n)$ .  $\square$

**Lemma 2.** *The function  $g(x) = -2x \log x - 2x(1-x)$  reaches its absolute maximum in the interval  $[0, 1]$  at the point*

$$\vartheta := -\frac{1}{2}W\left(-\frac{2}{e^2}\right) = 0.20318786\dots,$$

where  $W$  denotes Lambert  $W$ -function. Moreover, the value of this maximum is:

$$g(\vartheta) = 2(\vartheta - \vartheta^2) = 0.32380511\dots$$

*Proof.* This is an elementary Calculus exercise. The derivative  $g'(x) = -4 + 4x - 2 \log x = 0$  in  $[0, 1]$  if and only if  $x = 1$  or  $x = -\frac{1}{2}W\left(-\frac{2}{e^2}\right)$ . It is easily checked that  $x = 1$  is a minimum and that  $x = -\frac{1}{2}W\left(-\frac{2}{e^2}\right)$  is a maximum and some easy computations conclude the proof.  $\square$

**Theorem 2.** *With all the previous notation, the following hold:*

i) *For every positive integer  $n$ ,*

$$1 = \mathbf{P}(1) = \mathbf{P}(2) > \mathbf{P}(3) > \cdots > \mathbf{P}(n) > \mathbf{P}(n+1) > \cdots > 2(\vartheta - \vartheta^2).$$

ii)  $\lim_{n \rightarrow \infty} \frac{\mathbf{m}(n)}{n} = \vartheta$ ; *i.e.*,  $\mathbf{m}(n) \sim \vartheta n$ .

iii)  $\lim_{n \rightarrow \infty} P(\lfloor \vartheta n \rfloor, n) = \lim_{n \rightarrow \infty} \mathbf{P}(n) = 2(\vartheta - \vartheta^2)$ .

*Proof.*

i) By Corollary 1, we know that  $\mathbf{P}(2) = 1$  and that  $\mathbf{P}(n)$  is decreasing. Now, if we recall Lemma 1, we can put

$$P(r, n) = \frac{2r}{n} \left[ \left( \frac{r}{n} - 1 \right) + \sum_{i=r}^{n-1} \frac{1}{i} \right] = \frac{2r}{n} \left( \frac{r}{n} - 1 \right) + \frac{2r}{n} \sum_{i=r}^{n-1} \frac{1}{i},$$

where the latter sum is a Riemann sum. Consequently,

$$\lim_{n \rightarrow \infty} P(r, n) \simeq 2x(x-1) + 2x \int_x^1 \frac{1}{t} dt = 2x(x-1) - 2x \log x = g(x)$$

with  $x = \lim_{n \rightarrow \infty} r/n$ .

Finally, since  $\mathbf{P}(n)$  is decreasing in  $n$  and  $2(\vartheta - \vartheta^2)$  is the maximum value of  $g(x)$  due to Lemma 2 it follows that  $\mathbf{P}(n) > 2(\vartheta - \vartheta^2)$  for every integer  $n > 1$ , as claimed.

ii) By Lemma 2,  $g(x)$  reaches its maximum value at  $\vartheta$ . Since  $P(\cdot, n)$  reaches its maximum at  $\mathbf{m}(n)$  and  $x = \lim_{n \rightarrow \infty} r/n$ , it follows immediately that

$$\lim_{n \rightarrow \infty} \frac{\mathbf{m}(n)}{n} = \vartheta.$$

iii) From the previous work and Lemma 2, it is clear that

$$\lim_{n \rightarrow \infty} P(\lfloor \vartheta n \rfloor, n) = \lim_{n \rightarrow \infty} \mathbf{P}(n) = g(\vartheta) = 2(\vartheta - \vartheta^2)$$

because  $\lim_{n \rightarrow \infty} \lfloor \vartheta n \rfloor / n = \vartheta$ .

□

From the above theorem, it can be deduced that, for any value of  $n$ , the optimal strategy achieves success with a probability greater than  $2(\vartheta - \vartheta^2) = 0.3238\dots$ . Moreover, this is the approximate value of the probability of success for large values of  $n$ , just by following the simple strategy of rejecting the nearest integer to  $-\frac{n}{2}W(-\frac{2}{e^2}) = n \cdot 0.2031\dots$ . Thus,  $\lfloor \vartheta \cdot n \rfloor$  constitutes a practical estimation of the optimal value of initial rejections, just like  $\lfloor n/e \rfloor$  does in the classical secretary problem. This is, in fact, true:  $\lfloor n\vartheta \rfloor$  is within 1 of the correct answer for all  $n$ , although many errors (20%) are made with this estimation:

3, 8, 13, 18, 23, 32, 37, 42, 47, 52, 57, 62, 67, 72, 77, 82, 96, 101, 106, 111, 116, 121,  $\dots$

However, it is also true that the error is negligible if compared with the probability of the optimal strategy for large values of  $n$ .

Let us now look at a result that provides a better estimate for  $\mathbf{m}(n)$  than the one above. However, let us first consider the following auxiliary result where  $\psi$  stands for the digamma function.

**Lemma 3.** *Let us consider the function  $f(r, n) := -\frac{2r}{n} + \frac{2r^2}{n^2} + \frac{2r}{n}\psi(n) - \frac{2r \log(r)}{n} + \frac{1}{n}$  and let  $\alpha(n)$  be the value for which the function  $f(\cdot, n)$  reaches its maximum in  $[1, n]$ . Then,  $\alpha(n) \approx \mathbf{m}(n)$ .*

*Proof.* If we recall the definition of the digamma function  $\psi$ , we can write

$$P(r, n) = -\frac{2r}{n} + \frac{2r^2}{n^2} + \frac{2r}{n}\psi(n) - \frac{2r}{n}\psi(r).$$

Now, we know that for any integer  $r$

$$\psi(r) = \log r - \frac{1}{2r} + \epsilon(r),$$

with  $\epsilon(r) = \mathcal{O}(1/2r)$  if  $r \rightarrow \infty$ . Thus,

$$P(r, n) = -\frac{2r}{n} + \frac{2r^2}{n^2} + \frac{2r}{n}\psi(n) - \frac{2r \log(r)}{n} + \frac{1}{n} - \frac{2r\epsilon(r)}{n} = f(r, n) - \frac{2r\epsilon(r)}{n}.$$

On the other hand, since  $r \geq 1$ , it follows that there exists  $k$  such that  $|2r\epsilon(r)| \leq k$  for every  $r$  and hence:

$$f(r, n) - \frac{k}{n} \leq P(r, n) \leq f(r, n) + \frac{k}{n}$$

Obviously both functions  $g_1(r, n) := f(r, n) - \frac{k}{n}$  and  $g_2(r, n) := f(r, n) + \frac{k}{n}$  reach their maximum at  $\alpha(n)$ . Now, let  $\beta_1(n)$  and  $\beta_2(n)$  points such that  $g_2(\beta_1(n), n) = g_2(\beta_2(n), n) = g_1(\alpha(n), n)$  and  $\alpha(n) \in [\beta_1(n), \beta_2(n)]$ . The inequality above implies that

$$\beta_1(n) \leq \mathbf{m}(n) \leq \beta_2(n)$$

Finally, for every  $r$  we have that  $g_2(r, n) - g_1(r, n) = 2k/n$  so it follows that  $|\beta_2(n) - \beta_1(n)| \xrightarrow[n \rightarrow \infty]{} 0$  and, consequently also

$$|\alpha(n) - \mathbf{m}(n)| \xrightarrow[n \rightarrow \infty]{} 0.$$

□

**Theorem 3.** *With all the previous notation, the following hold:*

- i)  $\mathbf{m}(n) \approx -\frac{n}{2}W\left(-\frac{2e^{-2+\psi(n)}}{n}\right)$ , where  $\psi$  denotes the digamma function.
- ii)  $\mathbf{m}(n) \approx n\vartheta + \frac{1}{4 - 2e^{2-2\vartheta}}$ .

*Proof.* i) Since  $f(\cdot, n)$  (see Lemma 3) reaches its maximum at  $\alpha(n)$ , it follows that

$$0 = \frac{\partial f}{\partial r}(\alpha(n), n) = -\frac{4}{n} + \frac{4\alpha(n)}{n^2} + \frac{2}{n}\psi(n) - \frac{2 \log(\alpha(n))}{n}.$$

From this, and taking into account the definition of Lambert  $W$  function it follows that

$$\alpha(n) = -\frac{n}{2}W\left(-\frac{2e^{-2+\psi(n)}}{n}\right)$$

so it is enough to apply Lemma 3.



ii) Using i), we have that

$$\lim_{n \rightarrow \infty} (\mathbf{m}(n) - n\vartheta) = \lim_{n \rightarrow \infty} \left( -\frac{n}{2} W\left(-\frac{2e^{-2+\psi(n)}}{n}\right) - n\vartheta \right) = \frac{1}{4 - 2e^{2-2\vartheta}}.$$

Thus,

$$\mathbf{m}(n) \approx n\vartheta + \frac{1}{4 - 2e^{2-2\vartheta}}$$

as claimed. □

Using the formula  $\left[ n\vartheta + \frac{1}{4 - 2e^{2-2\vartheta}} \right]$  as far as our computational capacity let us, we find that the optimal cutoff value produces errors only for very few values of  $n$ . Specifically, only four are known: 2, 3, 23 and 2971. We now show another asymptotic result for the solution  $\mathbf{m}(n) \approx h(n)$  that improves the result of the previous theorem in the sense that there is no known value of  $n > 4$  for which  $[h(n)]$  returns a wrong result.

**Proposition 1.**

$$\mathbf{m}(n) \approx h(n) := \frac{1}{2} + \frac{3W\left(\frac{-2}{e^2}\right)}{4} + \frac{W\left(\frac{-2}{e^2}\right)}{4\left(1 + W\left(\frac{-2}{e^2}\right)\right)} - \frac{1}{2W\left(\frac{1}{e^{\frac{3}{2n}} n W\left(\frac{-2}{e^2}\right)}\right)}.$$

*Proof.* It suffices to verify that  $h(n)$  has the following straight line as an asymptote:

$$n\vartheta + \frac{1}{4 - 2e^{2-2\vartheta}}.$$

□

Although no exceptions to the equality  $\mathbf{m}(n) = [h(n)]$  have been found for  $n > 4$ , there is no absolute guarantee that this occurs *ad infinitum*. This is what happens with the following asymptotic formula for the optimal cutoff value in the classical secretary problem

$$\left[ -\frac{1}{2W\left(-\frac{e^{-(1+1/(2n))}}{2n}\right)} \right]$$

which coincides with the solution for all  $n > 3$  with no known exceptions.

**Remark.** The constant  $\vartheta = -\frac{1}{2}W\left(-\frac{2}{e^2}\right) = 0.20318786\dots$  (A106533 in OEIS) appears in [7] in the context of the secretary problem considering a payoff of  $(n - k + 1)/n$ . In fact, it appears erroneously approximated as 0.20388, but is actually the constant which appears, as in this study, as the solution of the equation  $-\log(x) - 2 + 2x = 0$ . Furthermore, as a noteworthy curiosity, it should be pointed out that this constant has appeared in a completely different context from the one addressed here as the solution to the above equation (the Daley-Kendall model) and is known as the *rumour's constant* [4, 9].

**3.1. What is the value of the information regarding the number of objects?** When the number of available objects is unknown, but it is known that it is a uniform random variable  $U[0, N]$ , we have established that the asymptotic value of the probability of success is  $2(\vartheta - \vartheta^2) = 0.3238\dots$ . It is natural to ask how valuable exact knowledge of the number of objects that we shall be able to observe may be. That is, if we win 1 Euro for making a successful choice, how much

might we pay to know the value of the random variable  $U[0, N]$  that represents the number of available objects.

If, prior to the inspections, we are given the information regarding the number of objects,  $n \in [1, N]$ , in each case we will adopt the optimal strategy set out in the first section of this paper that gives us a probability of success

$$\mathfrak{p}_n := \begin{cases} \frac{n}{2(n-1)}, & \text{if } n > 2 \text{ is even;} \\ \frac{n+1}{2n}, & \text{if } n > 2 \text{ is odd} \\ 1 & \text{if } n \in \{1, 2\} \end{cases}$$

and the overall probability of success will be  $\mathfrak{P}_N := N^{-1} \sum_{n=1}^N \mathfrak{p}_n$ . Thus, it is straightforward to see that the probability of success with a number of objects resulting from a uniform random variable  $[1, N]$ , whose value is revealed before the inspections start, when  $N$  tends to infinity, is

$$\lim_{N \rightarrow \infty} \mathfrak{P}_N = \frac{1}{2}$$

Therefore, the value of the information is  $\mathfrak{P}_N - \mathbf{P}(N)$ , which, when  $N$  tends to infinity, is

$$\lim_{N \rightarrow \infty} (\mathfrak{P}_N - \mathbf{P}(N)) = \frac{1}{2} - 2(\vartheta - \vartheta^2) = 0.1761 \dots$$

#### 4. UNKNOWN NUMBER OF CANDIDATES WITH A $\lambda$ -POISSON DISTRIBUTION

For the classical secretary problem, Presman and Sonin, [11] showed that, if the number of candidates is Poisson distributed with parameter  $\lambda$ , then the optimal stopping limit relation is  $r^*(\lambda)/\lambda \rightarrow e^{-1}$  and that this is also the asymptotic value of the probability of success. In the variant we have introduced in this paper, an analogous conclusion is reached by replacing  $1/e$  by  $1/2$ , as seen in Section 2.

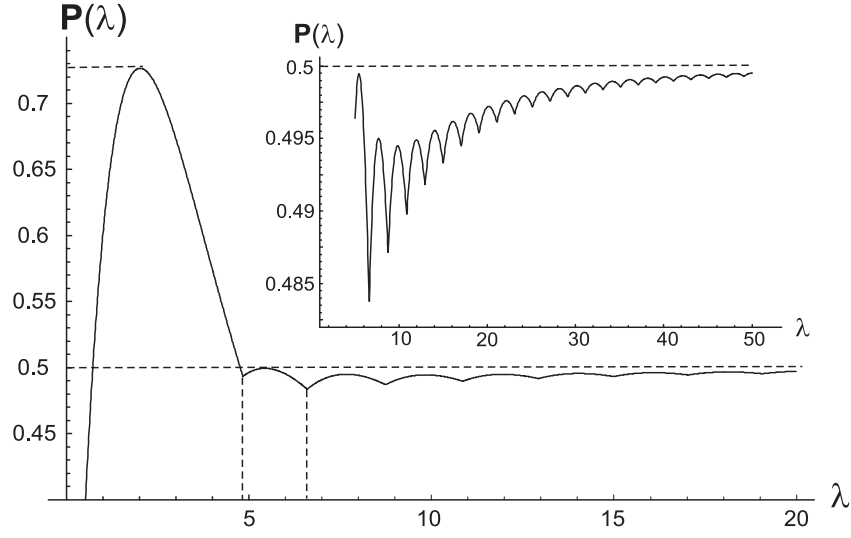
If the number of objects in the best or worst variant is Poisson distributed with parameter  $\lambda$ , the probability of success after rejecting the first  $r$  candidates will be

$$\begin{aligned} P(0, \lambda) &:= \frac{\lambda}{e^\lambda} + \frac{\lambda^2}{2e^\lambda} + \sum_{i=3}^{\infty} \frac{2e^{-\lambda}\lambda^i}{i \cdot i!}, \\ P(1, \lambda) &:= \frac{\lambda^2}{2e^\lambda} + \sum_{i=3}^{\infty} \frac{2e^{-\lambda}\lambda^i}{i \cdot i!}, \\ P(r, \lambda) &:= \sum_{i=r+1}^{\infty} \frac{2(i-r)r e^{-\lambda}\lambda^i}{i(i-1) \cdot i!} \end{aligned}$$

Following the same notation as in Section 3, we denote by  $\mathbf{m}(\lambda)$  the value for which  $P(\cdot, \lambda)$  reaches its maximum value and by  $\mathbf{P}(\lambda)$ , the maximum value it reaches; i.e.  $\mathbf{P}(\lambda) := P(\mathbf{m}(\lambda), \lambda)$ . Unlike what occurred in the previous case, the probability of success in the game,  $\mathbf{P}(\lambda)$ , is not monotone with respect to  $\lambda$ . Next, we show the graph of  $\mathbf{P}(\lambda)$  (Figure 1), which is seen to comprise concave arcs in each interval  $[\lambda_i, \lambda_{i+1}]$ , where  $\lambda_0 = 0$ ,  $\lambda_1$  is such that  $P(0, \lambda_1) = P(2, \lambda_1)$  and  $\lambda_i$  for  $i > 1$  is such that  $P(i, \lambda_i) = P(i+1, \lambda_i)$ .

The maximum probability of success is achieved in  $\lambda = \tilde{\lambda} := 2.01771 \dots$ , the solution to the equation

$$0 = -2 + 2e^x - x + 2\gamma x + x^2 + 2x\Gamma(-x) + 2x \log(-x),$$

FIGURE 1. The probability  $P(\lambda)$ .

where  $\mathbf{m}(\tilde{\lambda}) = 0$  and  $\mathbf{P}(\tilde{\lambda}) = 0.726470\dots$

We will prove that, as Figure 1 suggests,  $\mathbf{P}(\lambda)$  tends to  $1/2$  when  $\lambda$  grows to infinity and that this propability can be asymptotically reached using  $\frac{\lambda}{2}$  as cutoff value. But first we need some technical results.

**Lemma 4.** *Let us define the function*

$$f(r, \lambda) := \sum_{i=2}^r \frac{2(i-r)r e^{-\lambda} \lambda^i}{i(i-1) i!}.$$

Then,  $\lim_{\lambda \rightarrow \infty} f(\frac{\lambda}{2}, \lambda) = 0$ .

*Proof.* First of all, note that

$$\left| \sum_{i=2}^{\lambda/2} \frac{2(i-\lambda/2)\lambda/2 e^{-\lambda} \lambda^i}{i(i-1) i!} \right| < \lambda^2 e^{-\lambda} \sum_{i=2}^{\lambda/2} \frac{\lambda^i}{i!} < \frac{1}{\lambda}$$

so, if we define

$$a_n := 2^2 n^2 e^{-2n} \sum_{i=2}^n \frac{2^i n^i}{i!},$$

we just need to prove that  $\lim_{n \rightarrow \infty} a_n = 0$ .

Now,

$$a_{n+1} = 2^2 (n+1)^2 e^{-2(n+1)} \sum_{i=2}^{n+1} \frac{2^i (n+1)^i}{i!}$$

and since

$$\begin{aligned} \sum_{i=2}^{n+1} \frac{2^i (n+1)^i}{i!} &= \sum_{i=2}^n \frac{2^i n^i}{i!} + \frac{2^2}{2!} (2n+1) + \frac{2^3}{3!} (3n^2 + 3n + 1) + \dots + \\ &+ \frac{2^n}{n!} \left( n^n + \frac{n}{1!} n^{n-1} + \frac{n(n-1)}{2!} n^{n-2} + \dots + \frac{n(n-1)\dots 2 \cdot 1}{n!} \right) + \frac{2^{n+1}}{(n+1)!} (n+1)^{n+1} \end{aligned}$$

we have that

$$\begin{aligned} a_{n+1} &= 2^2(n+1)^2 e^{-2(n+1)} \sum_{i=2}^n \frac{2^i n^i}{i!} + 2^2(n+1)^2 e^{-2(n+1)} \frac{2^2}{2!} (2n+1) + \dots + \\ &\quad + 2^2(n+1)^2 e^{-2(n+1)} \frac{2^n}{n!} \left( n^n + \frac{n}{1!} n^{n-1} + \frac{n(n-1)}{2!} n^{n-2} + \dots + \frac{n(n-1) \cdots 2 \cdot 1}{n!} \right) \\ &\quad + 2^2(n+1)^2 e^{-2(n+1)} \frac{2^{n+1}}{(n+1)!} (n+1)^{n+1} \end{aligned}$$

On the other hand,

$$\sum_{i=2}^n \frac{2^i n^i}{i!} = \frac{a_n}{2^2 n^2 e^{-2n}}.$$

Hence

$$\begin{aligned} a_{n+1} &= \left( \frac{n+1}{n} \right)^2 \frac{a_n}{e^2} + 2^2(n+1)^2 e^{-2(n+1)} \frac{2^2}{2!} (2n+1) + \dots + \\ &\quad + 2^2(n+1)^2 e^{-2(n+1)} \frac{2^n}{n!} \left( n^n + \frac{n}{1!} n^{n-1} + \frac{n(n-1)}{2!} n^{n-2} + \dots + \frac{n(n-1) \cdots 2 \cdot 1}{n!} \right) \\ &\quad + 2^2(n+1)^2 e^{-2(n+1)} \frac{2^{n+1}}{(n+1)!} (n+1)^{n+1}, \end{aligned}$$

which clearly implies that

$$\lim_{n \rightarrow \infty} \left[ a_{n+1} - \left( \frac{n+1}{n} \right)^2 \frac{a_n}{e^2} \right] = 0.$$

Finally,  $a_n$  being decreasing and non-negative, we have that  $\lim_{n \rightarrow \infty} a_{n+1} = \lim_{n \rightarrow \infty} a_n = l$ . Thus,

$$l = \frac{l}{e^2} \Rightarrow l \left( 1 - \frac{1}{e^2} \right) = 0 \Rightarrow l = 0$$

and

$$l = \lim_{n \rightarrow \infty} a_n = 0$$

as claimed.  $\square$

**Theorem 4.** Consider the function  $P^*$  defined over  $\mathbb{N} \times \mathbb{R}^+$  and given by

$$P^*(r, \lambda) := \sum_{i=2}^{\infty} \frac{2(i-r)r e^{-\lambda} \lambda^i}{i(i-1) i!}.$$

For a fixed  $\lambda$ , let us denote by  $\mathcal{M}^*(\lambda)$  the value for which  $P^*(\cdot, \lambda)$  reaches its maximum and put  $\mathcal{P}^*(\lambda) := P^*(\mathcal{M}^*(\lambda), \lambda)$ . Then, the following hold:

- i)  $\lim_{\lambda \rightarrow \infty} \frac{\mathcal{M}^*(\lambda)}{\lambda} = \frac{1}{2}$ .
- ii)  $\lim_{\lambda \rightarrow \infty} P^*\left(\frac{\lambda}{2}, \lambda\right) = \frac{1}{2}$ .
- iii)  $\mathcal{M}^*(\lambda) \approx \frac{\lambda}{2} - 1$ .

*Proof.* First of all, we are going to compute the value of  $\mathcal{M}^*(\lambda)$ . To do so, let us consider functions

$$\begin{aligned} S_1(x) &= \sum_{i=2}^{\infty} \frac{x^i}{(i-1)!}, \\ S_2(x) &= \sum_{i=2}^{\infty} \frac{x^i}{i(i-1)!} = - \sum_{i=2}^{\infty} \frac{x^i}{i!} + S_1(x), \\ S_3(x) &= \sum_{i=2}^{\infty} \frac{x^i}{i!}. \end{aligned}$$

Then, we have that

$$S_1''(x) = \frac{e^x - 1}{x}, \quad S_3'(x) = \frac{e^x - 1}{x} - 1.$$

If we integrate these expressions we obtain that

$$\begin{aligned} S_1(x) &= 1 - e^x + x - \gamma x + x \mathbf{E}(x) - x \log x \\ S_3(x) &= -\gamma + \mathbf{E}(x) - \log x - x \end{aligned}$$

where

$$\mathbf{E}(\lambda) = \gamma + \log \lambda + \int_0^\lambda \frac{e^x - 1}{x} dx.$$

Consequently,

$$S_2(x) = 1 - e^x + 2x + (\gamma - \mathbf{E}(x) + \log x)(1 - x).$$

Finally, note that

$$P^*(r, \lambda) = 2re^{-\lambda} \left[ \sum_{i=2}^{\infty} \frac{\lambda^i}{(i-1)!} - r \sum_{i=2}^{\infty} \frac{\lambda^i}{i(i-1)!} \right] = 2re^{-\lambda} S_1(\lambda) - 2r^2 e^{-\lambda} S_2(\lambda)$$

Hence,

$$\frac{\partial P^*(r, \lambda)}{\partial r} = 2e^{-\lambda} S_1(\lambda) - 4re^{-\lambda} S_2(\lambda)$$

and, consequently, if we define

$$r(\lambda) := \frac{1 - e^\lambda + \lambda - \gamma\lambda + \lambda \mathbf{E}(\lambda) - \lambda \log \lambda}{2 [1 - e^\lambda + 2\lambda + (\gamma - \mathbf{E}(\lambda) + \log \lambda)(1 - \lambda)]}$$

it is clear that  $0 = \frac{\partial P^*(r, \lambda)}{\partial r} \Big|_{r=r(\lambda)}$ . Thus, we have just obtained that  $\mathcal{M}^*(\lambda) = r(\lambda)$ .

Once we have computed the value of  $\mathcal{M}^*(\lambda)$ , we are in the condition to prove the three statements of the theorem

i) We have

$$\lim_{\lambda \rightarrow \infty} \frac{\mathcal{M}^*(\lambda)}{\lambda} = \frac{1}{2} \lim_{\lambda \rightarrow \infty} \frac{1 - e^\lambda + \lambda - \gamma\lambda + \lambda \mathbf{E}(\lambda) - \lambda \log \lambda}{\lambda [1 - e^\lambda + 2\lambda + (\gamma - \mathbf{E}(\lambda) + \log \lambda)(1 - \lambda)]}$$

so, if we recall the definition of  $\mathbf{E}(\lambda)$  we have that

$$\lim_{\lambda \rightarrow \infty} \frac{\mathcal{M}^*(\lambda)}{\lambda} = \frac{1}{2} \lim_{\lambda \rightarrow \infty} \frac{1 - e^\lambda + \lambda + \lambda \int_0^\lambda \frac{e^x - 1}{x} dx}{\lambda [1 - e^\lambda + 2\lambda + (\lambda - 1) \int_0^\lambda \frac{e^x - 1}{x} dx]}$$

and, applying L'Hôpital's rule repeatedly we obtain:

$$\lim_{\lambda \rightarrow \infty} \frac{\mathcal{M}^*(\lambda)}{\lambda} = \frac{1}{2} \lim_{\lambda \rightarrow \infty} \frac{1 + \lambda}{\lambda + 3} = \frac{1}{2},$$

as claimed.

ii) Now, we have that

$$\lim_{\lambda \rightarrow \infty} P^*\left(\frac{\lambda}{2}, \lambda\right) = \lim_{\lambda \rightarrow \infty} \frac{\lambda}{2e^\lambda} \left[ (-\lambda^2 + 3\lambda) \int_0^\lambda \frac{e^x - 1}{x} dx + (\lambda - 2)e^\lambda - 2\lambda^2 + \lambda + 1 \right]$$

so, again by L'Hôpital's rule, we obtain:

$$\lim_{\lambda \rightarrow \infty} P^*\left(\frac{\lambda}{2}, \lambda\right) = \frac{1}{2} \lim_{\lambda \rightarrow \infty} \frac{6 + 2\lambda}{2\lambda + 5} = \frac{1}{2}.$$

iii) Since

$$\lim_{\lambda \rightarrow \infty} (\mathcal{M}^*(\lambda) - \lambda/2) = \frac{1}{2} \lim_{\lambda \rightarrow \infty} \frac{1 - e^\lambda + \lambda e^\lambda - 2\lambda^2 + (2\lambda - \lambda^2) \int_0^\lambda \frac{e^x - 1}{x} dx}{1 - e^\lambda + 2\lambda + (\lambda - 1) \int_0^\lambda \frac{e^x - 1}{x} dx},$$

L'Hôpital's rule leads to

$$\lim_{\lambda \rightarrow \infty} (\mathcal{M}^*(\lambda) - \lambda/2) = \frac{1}{2} \lim_{\lambda \rightarrow \infty} \frac{-2\lambda - 4}{\lambda} = -1.$$

Thus,

$$\mathcal{M}^*(\lambda) \approx \frac{\lambda}{2} - 1,$$

as claimed. □

**Lemma 5.** *Assume that the number of objects  $n$  is Poisson distributed with parameter  $\lambda$ . If the (random) value of  $n$  is revealed before the inspections start, then the probability of success tends to  $1/2$  when  $\lambda$  tends to infinity. In other words:*

$$\lim_{\lambda \rightarrow \infty} \sum_{n=1}^{\infty} \mathbf{p}_n \frac{e^{-\lambda} \lambda^n}{n!} = \frac{1}{2}.$$

*Proof.* We reason just like in Subsection 3.1. If, prior to the inspections, we are given the information regarding the number of objects,  $n \in [0, \infty)$ , in each case we will adopt the optimal strategy set out in the first section of this paper that gives us a probability of success  $\mathbf{p}_n$  and the overall probability of success will be

$$\mathfrak{P}_\lambda := \sum_{n=1}^{\infty} \mathbf{p}_n \frac{e^{-\lambda} \lambda^n}{n!} = 2 \sum_{n=1}^{\infty} \frac{n}{-1 + 2n} \frac{e^{-\lambda} \lambda^n}{n!} = \frac{\lambda S(\lambda)}{e^\lambda},$$

where

$$S(\lambda) = \int_0^\lambda \frac{\sinh(x)}{x} dx.$$

so we have that

$$\lim_{\lambda \rightarrow \infty} \mathfrak{P}_\lambda = \frac{1}{2}. \quad \square$$

**Theorem 5.** *In the previous setting we have that*

$$\lim_{\lambda \rightarrow \infty} P\left(\frac{\lambda}{2}, \lambda\right) = \lim_{\lambda \rightarrow \infty} P(\mathbf{m}(\lambda), \lambda) = \lim_{\lambda \rightarrow \infty} \mathbf{P}(\lambda) = \frac{1}{2}.$$

*Proof.* Since  $\mathfrak{p}_n$  (see Lemma 5) is the maximum value of the probability of succeeding in the problem with complete information with  $n$  objects, it follows for all  $n > 1$  that

$$\mathfrak{p}_n \geq P_n(\mathfrak{m}(\lambda)) = \frac{2(n - \mathfrak{m}(\lambda))\mathfrak{m}(\lambda)}{n(n - 1)}$$

and hence,

$$\sum_{n=1}^{\infty} \mathfrak{p}_n \frac{e^{-\lambda} \lambda^n}{n!} \geq \sum_{n=\mathfrak{m}(\lambda)+1}^{\infty} \frac{2(n - \mathfrak{m}(\lambda))\mathfrak{m}(\lambda)}{n(n - 1)} \frac{e^{-\lambda} \lambda^n}{n!} = \mathbf{P}(\lambda)$$

So, if we take upper limits it is clear that  $\overline{\lim}_{\lambda \rightarrow \infty} P(\mathfrak{m}(\lambda), \lambda) \leq \frac{1}{2}$ .

Now, recalling the function  $f$  from Lemma 4 we have that

$$\mathbf{P}(\lambda) = P(\mathfrak{m}(\lambda), \lambda) \geq P(\lambda/2, \lambda) = P^*(\lambda/2, \lambda) - f(\lambda/2, \lambda).$$

But, due to Lemma 4 and Theorem 4 we obtain that

$$\lim_{\lambda \rightarrow \infty} P^*(\lambda/2, \lambda) - f(\lambda/2, \lambda) = \frac{1}{2}$$

and taking lower limits:

$$\underline{\lim}_{\lambda \rightarrow \infty} \mathbf{P}(\lambda) \geq \frac{1}{2}.$$

In conclusion,

$$\frac{1}{2} \leq \underline{\lim}_{\lambda \rightarrow \infty} \mathbf{P}(\lambda) \leq \overline{\lim}_{\lambda \rightarrow \infty} \mathbf{P}(\lambda) \leq \frac{1}{2}$$

and the proof is complete.  $\square$

This theorem satisfactorily solves the problem, since it implies that  $\lambda/2$  is the best possible estimation for the cutoff value because it guarantees an asymptotical succeeding probability of  $\frac{1}{2}$  (exactly the same that the exact value of  $\mathfrak{m}(\lambda)$  would provide). Thus, this result proves that if the number of objects is unknown with a  $\lambda$ -Poisson distribution, the strategy that we must follow in order to maximize the succeeding probability is to select the first object which is better or worst than the previous ones just after rejecting the first  $\lambda/2$  objects. With this strategy we will succeed approximately one half of the times.

Unlike what happened in the previous section (see 3.1.), when facing with an unknown number of objects resulting from a Poisson distribution with parameter  $\lambda$ , the value of the information of the number of objects is asymptotically null. For large values of  $\lambda$ , the difference between the probability of success with and without this information is negligible.

In addition, as far as our computing capabilities let us check,  $\lfloor \lambda/2 - 1 \rfloor$  coincides with the exact value  $\mathfrak{m}(\lambda)$  for every integral value of  $\lambda$  greater than 1. Hence, the following conjecture.

**Conjecture 1.**  $\mathfrak{m}(\lambda) \approx \mathcal{M}^*(\lambda) \approx \frac{\lambda}{2} - 1$ .

## 5. PAYOFF PROPORTIONAL TO THE NUMBER OF PERFORMED OR NON-PERFORMED INTERVIEWS

We now consider the same underlying game as in the previous sections, but introducing different payoffs based on the number of observed objects. We consider that, in the case of success (choosing the best or the worst object), we shall receive a payoff based on the number of interviews carried out. The aim is therefore to

maximize the payoff. Although the goal is still to choose the best or worst object, there are incentives (or penalties) for carrying out more (or fewer) inspections of the objects than those that maximize the probability of success found in Section 2, as higher payoffs will be obtained. We first consider the case in which the payoff received is proportional to the number of observed objects and then the case in which the payoff received is proportional to the number of non-observed objects.

If we consider that the payoff for success in the game with  $n$  available objects after having carried out  $k$  inspections is  $g(k, n)$ , the expected payoff following the strategy of rejecting  $r$  objects and choosing the first observed object that is better or worse than those seen so far will thus be:

$$\mathbf{G}(r, n) := \frac{2\binom{r}{2}}{n} \sum_{k=r+1}^n \frac{g(k, n)}{\binom{k-1}{2}}$$

We consider two cases: the first, where  $g(k, n) = \frac{k}{n}$ , and the second, where  $g(k, n) = \frac{n-k}{n}$ . In both cases, we obtain asymptotic formulas for the optimal cutoff value and the expected payoff following the optimal strategy.

### 5.1. Payoff proportional to the number of performed interviews.

**Theorem 6.** *Consider the function*

$$\mathbf{G}(r, n) := \frac{2\binom{r}{2}}{n^2} \sum_{k=r+1}^n \frac{k}{\binom{k-1}{2}}$$

defined for every pair of integers  $(r, n)$  such that  $1 < r < n$ . Let us denote by  $\mathcal{M}(n)$  the value for which the function  $\mathbf{G}(\cdot, n)$  reaches its maximum. Then,

- i)  $\lim_{n \rightarrow \infty} \frac{\mathcal{M}(n)}{n} = \frac{1}{\sqrt{e}}$ ; i.e.,  $\mathcal{M}(n) \sim \frac{n}{\sqrt{e}}$ .
- ii)  $\lim_{n \rightarrow \infty} \mathbf{G}\left(\left[\frac{n}{\sqrt{e}}\right], n\right) = \frac{1}{e}$ .

*Proof.* First of all, observe that

$$\mathbf{G}(r, n) = \frac{2\binom{r}{2}}{n^2} \sum_{k=r+1}^n \frac{k}{\binom{k-1}{2}} = \frac{r(r-1)}{n^2} \sum_{k=r+1}^n \frac{2k}{(k-1)(k-2)}$$

Now, using telescopic sums, we have that

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

Thus,

$$\begin{aligned} \mathbf{G}(r, n) &= \frac{2r(r-1)}{n^2} \left[ \frac{2}{r-1} - \frac{2}{n-1} + \sum_{i=r}^{n-1} \frac{1}{i} \right] = \frac{4r(n-r)}{n^2(n-1)} + \frac{2r(r-1)}{n^2} \sum_{i=r}^{n-1} \frac{1}{i} = \\ &= \frac{4}{n-1} \frac{r}{n} \left( 1 - \frac{r}{n} \right) + 2 \frac{r}{n} \left( \frac{r}{n} - \frac{1}{n} \right) \sum_{i=r}^{n-1} \frac{1}{i} \frac{1}{n} \end{aligned}$$

where the latter sum is a Riemann sum. Consequently,

$$\lim_{n \rightarrow \infty} \mathbf{G}(r, n) \simeq 2x^2 \int_x^1 \frac{1}{t} dt = -2x^2 \log x = g(x)$$



with  $x = \lim_{n \rightarrow \infty} r/n$ .

Now, it is an elementary Calculus exercise to see that the function  $g(x) = -2x^2 \log x$  reaches its absolute maximum at the point

$$\frac{1}{\sqrt{e}} = 0.606531 \dots$$

Moreover, the value of this maximum is:

$$g\left(\frac{1}{\sqrt{e}}\right) = \frac{1}{e} = 0.367879 \dots$$

After this previous work we can prove the statements of the theorem.

- i) Since  $\mathbf{G}(\cdot, n)$  reaches its maximum at  $\mathcal{M}(n)$  and  $x = \lim_{n \rightarrow \infty} r/n$ , it follows immediately that

$$\lim_{n \rightarrow \infty} \frac{\mathcal{M}(n)}{n} = \frac{1}{\sqrt{e}}$$

- ii) It is clear that

$$\lim_{n \rightarrow \infty} \mathbf{G}\left(\left[\frac{n}{\sqrt{e}}\right], n\right) = g\left(\frac{1}{\sqrt{e}}\right) = \frac{1}{e}$$

because  $\lim_{n \rightarrow \infty} \left[\frac{n}{\sqrt{e}}\right]/n = \frac{1}{\sqrt{e}}$ .

□

Recall that in Theorem 2 iv), an asymptotic formula was obtained for the solution  $\mathbf{m}(n) = n\vartheta + \mathcal{O}(1)$  taking a series expansion. This was possible because equations of type  $0 = A + Bx + C \log(x)$  are solvable in terms of Lambert  $W$  Function. In this case, the equation that appears is of the type  $A + Bx + C \log(x) + Dx \log(x) = 0$ , for which there is no known solution that can be expressed symbolically. Nonetheless, the simplification of this equation achieved by replacing  $r/n$  by its asymptotic value,  $1/\sqrt{e}$ , gives rise to a formula that improves the estimation  $[n/\sqrt{e}]$  and, in this case, for  $\mathcal{M}(n)$ , we conjecture

**Conjecture 2.**

$$\mathcal{M}(n) \approx \frac{n}{\sqrt{e}} + \mu,$$

where  $\mu := 2 - \frac{5}{2\sqrt{e}} = 0.483673350 \dots$

In fact,  $[\frac{n}{\sqrt{e}} + \mu]$  is the value of  $\mathcal{M}(n)$  for all  $n$  up to 10000, with the only exceptions of  $n \in \{33, 94\}$ .

## 5.2. Payoff proportional to the number of non-performed interviews.

**Theorem 7.** Consider the function

$$G(r, n) := \frac{2\binom{r}{2}}{n^2} \sum_{k=r+1}^n \frac{n-k}{\binom{k-1}{2}}$$

defined for every pair of integers  $(r, n)$  such that  $1 < r < n$  and let

$$\theta := -\frac{1}{2W_{-1}\left(-\frac{1}{2\sqrt{e}}\right)} = e^{\frac{1}{2} + W_{-1}\left(\frac{-1}{2\sqrt{e}}\right)}$$

be the solution to the equation  $2x \log(x) = x - 1$ . Denote by  $\mathcal{M}(n)$  the value for which the function  $G(\cdot, n)$  reaches its maximum. Then the following hold:

- i)  $\lim_{n \rightarrow \infty} \frac{\mathcal{M}(n)}{n} = \theta = 0.284668\dots$ ; i.e,  $\mathcal{M}(n) \sim n \cdot \theta$ .
- ii)  $\lim_{n \rightarrow \infty} G([n \cdot \theta], n) = \theta(1 - \theta) = 0.2036321\dots$

*Proof.* We reason like in the previous theorem. First, observe that

$$\begin{aligned} G(r, n) &= \frac{2r(r-1)}{n^2} \sum_{k=r+1}^n \frac{(n-k)}{(k-1)(k-2)} = \frac{2r(r-1)}{n^2} \left[ \frac{n-2}{r-1} - \frac{n}{n-1} + \sum_{i=r}^{n-1} \frac{1}{i} \right] = \\ &= 2\frac{r}{n} \left( 1 - \frac{2}{n} \right) - 2\frac{r}{n} \left( \frac{r}{n-1} - \frac{1}{n-1} \right) - 2\frac{r}{n} \left( \frac{r}{n} - \frac{1}{n} \right) \sum_{i=r}^{n-1} \frac{1}{i}, \end{aligned}$$

where the latter sum is a Riemann sum. Consequently,

$$\lim_{n \rightarrow \infty} G(r, n) \simeq 2x - 2x^2 - 2x^2 \int_x^1 \frac{1}{t} dt = 2x(1 - x + x \log x) = g(x)$$

with  $x = \lim_{n \rightarrow \infty} r/n$ . It is easily seen that  $g(x) = 2x(1 - x + x \log x)$  reaches its absolute maximum at the point

$$\theta = -\frac{1}{2W_{-1}\left(-\frac{1}{2\sqrt{e}}\right)} = 0.284668\dots$$

Moreover, the value of this maximum is:

$$g(\theta) = \theta - \theta^2 = 0.203632\dots$$

Now, we prove the statements of the theorem.

- i) Since  $G(\cdot, n)$  reaches its maximum at  $\mathcal{M}(n)$  and  $x = \lim_{n \rightarrow \infty} r/n$ , it follows immediately that

$$\lim_{n \rightarrow \infty} \frac{\mathcal{M}(n)}{n} = \theta$$

- ii) It is clear that

$$\lim_{n \rightarrow \infty} G([n \cdot \theta], n) = g(\theta) = \theta(1 - \theta)$$

because  $\lim_{n \rightarrow \infty} [n \cdot \theta]/n = \theta$ .

□

Just like in the previous case, it has not been possible to determine the asymptotic formula of the type  $\mathcal{M}(n) \approx n \cdot \theta + \beta$ . Nevertheless, it is possible to conjecture that  $\mathcal{M}(n) \approx n \cdot \theta + \mathcal{O}(1)$ . Moreover, we conjecture that  $\mathcal{M}(n) \approx n \cdot \theta + \mu$ , where  $\mu \cong 1.4034$  is an approximation that has been experimentally found on the basis of the calculation of the solution for large values of  $n$ . In this case,  $[n \cdot \theta + \mu]$  is the exact value of  $\mathcal{M}(n)$  for all values of  $n$  up to 10000, without any exception.

**Remark.** The constant  $\theta = -\frac{1}{2W_{-1}\left(-\frac{1}{2\sqrt{e}}\right)} = 0.284668\dots$  also appears related to rumour theory [4, 9] and to Gabriel's Horn (see A101314 in OEIS).

6. CONCLUSIONS AND FUTURE WORK

In this paper, we have analyzed a very natural variant of the secretary problem that has probably not received much attention because its asymptotic analysis for the game with full information is very simple. However, its study has been found to be of interest in the case of an unknown random number of objects and with payoffs that are proportional to the number of inspected objects. We show a comparative table of the asymptotic optimal stopping rule  $m(\cdot)$  and the  $\mathbb{E}(\text{payoff})$  (the asymptotic mean payoff) in the classical problem and in the best/worst variant studied in this paper.

All the constants that appear in the table can be expressed in terms of the following three constants:

$$e = 2.71828\dots, \vartheta := \frac{W(-2e^{-2})}{2} = 0.20318\dots,$$

$$\theta := -\frac{1}{2W_{-1}(-\frac{1}{2\sqrt{e}})} = 0.28466\dots$$

	Classic		Best or Worst	
	$m(\cdot)$	$\mathbb{E}(\text{payoff})$	$m(\cdot)$	$\mathbb{E}(\text{payoff})$
$n$ objects	$ne^{-1}$	$e^{-1}$	$n/2$	$1/2$
$\lambda$ -Poisson	$\lambda/e$	$e^{-1}$	$\lambda/2$	$1/2$
Uniform $[1, n]$	$ne^{-1/2}$	$2e^{-2}$	$n\vartheta$	$2(\vartheta - \vartheta^2)$
Payoff = $(n - k)/n$	$n\vartheta$	$\vartheta - \vartheta^2$	$n\theta$	$\theta - \theta^2$
Payoff = $k/n$	$n/2$	$1/4$	$ne^{-1/2}$	$e^{-1}$

Although we have not addressed its study, it is not unreasonable to conjecture that the counterpart of the  $1/e$ -law of best choice [2] in this game is what we might call the  $1/2$ -law. Let us suppose that, independently of each other, all applicants have the same arrival time density  $f$  on  $[0, T]$  and let  $F$  denote the corresponding arrival time distribution function, i.e.

$$F(t) = \int_0^t f(s)ds, 0 \leq t \leq T.$$

The  $1/2$ -law: Let  $\tau$  be such that  $F(\tau) = 1/2$ . Consider the strategy of waiting and observing all applicants up to time  $\tau$  and then choosing, if possible, the first candidate after time  $\tau$  who is better than all the preceding ones.

Thus, this strategy, which we shall call the  $1/2$ -strategy, has the following properties:

- i) It yields for all  $N$  a success probability of at least  $1/2$ ,
- ii) It is the only strategy guaranteeing this lower success probability bound,  $1/2$ , and the bound is optimal,
- iii) It chooses, if there is at least one applicant, none at all with a probability of exactly  $1/2$ .

Finally, we wish to draw attention a question that has not been solved in this paper, namely that of how to establish the value of the constants that appear in the asymptotic formula of the optimal cutoff of the problems posed in Section 4. Although from a practical point of view, an approximation of these constants with 4 digit accuracy seems to provide the correct solution (except, perhaps, for some isolated values of  $n$ ), it would be very interesting to obtain them exactly as a

solution of certain transcendental equations, as was possible in Section 2 with the constant  $\frac{1}{4-2e^{2-2\vartheta}}$ , where  $\vartheta$  represents the rumour's constant.

## REFERENCES

- [1] J.N. Bearden. A new secretary problem with rank-based selection and cardinal payoffs. *Journal of Mathematical Psychology*, 50: 58-59. 2006.
- [2] F. Th. Bruss. A unified Approach to a Class of Best Choice problems with an Unknown Number of Options. *Annals of Probability*, 12(3): 882-891. 1984.
- [3] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Adv. Comput. Math.*, 5: 329-359. 1996.
- [4] D.J. Daley and D.G. Kendall. Stochastic rumours. *Journal of the Institute of Mathematics and Its Applications*, 1:42-55. 1965.
- [5] E. B. Dynkin. The optimum choice of the instant for stopping a markov process. *Soviet Mathematics - Doklady*, 4:627-629, 1963.
- [6] T.S. Ferguson. Who solved the secretary problem? *Statistical Science*, 4(3): 282-296. 1989.
- [7] T. S. Ferguson, J. P. Hardwick and M. Tamaki. Maximizing the duration of owning a relatively best object. *Contemporary Mathematics: Strategies for Sequential Search and Selection in Real Time*, American Mathematics Association (T. Ferguson and S. Samuels, eds), 125: 37-58. 1991.
- [8] J. Gilbert and F. Mosteller. Recognizing the maximum of a sequence. *J. Am. Statist. Assoc.*, 61, 35-73, 1966.
- [9] E. Lebensztayn, F. P. Machado and P. M. Rodriguez. Limit Theorems for a general stochastic rumour model. [arxiv.org/pdf/1003.4995](https://arxiv.org/pdf/1003.4995).
- [10] D. V. Lindley. Dynamic programming and decision theory. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 10(1):39-51, 1961.
- [11] E.L. Presman and I.M. Sonin. The best choice problem for a random number of objects. *Theory Prob. Applic.*, 17, 657-668. 1972.
- [12] K.A. Szjowski. A rank-based selection with cardinal payoffs and a cost of choice. *Sci. Math. Jpn.*, 69(2), 285-293. 2009.
- [13] R.J. Vanderbei. The Optimal Choice of a Subset of a Population. *Mathematics of Operations Research*, 5(4): 481-486. 1980.

*E-mail address:* bayon@uniovi.es

*E-mail address:* grau@uniovi.es

*E-mail address:* oller@unizar.es

*E-mail address:* mruiz@uniovi.es

*E-mail address:* pedrosr@uniovi.es

DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DE OVIEDO, AVDA. CALVO SOTELO S/N, 33007 OVIEDO, SPAIN

DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DE OVIEDO, AVDA. CALVO SOTELO S/N, 33007 OVIEDO, SPAIN

CENTRO UNIVERSITARIO DE LA DEFENSA DE ZARAGOZA, CTRA. HUESCA S/N, 50090 ZARAGOZA, SPAIN

DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DE OVIEDO, AVDA. CALVO SOTELO S/N, 33007 OVIEDO, SPAIN

DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DE OVIEDO, AVDA. CALVO SOTELO S/N, 33007 OVIEDO, SPAIN