A contribution to the study of Robbins' Problem

by Y. C. Swan

Mémoire présenté le 31 Mars 2010 à l'Académie Royale
des Sciences, des Lettres et des Beaux-Arts

Université Libre de Bruxelles
Faculté des Sciences

Contents

Introduction

I would like to see this problem solved before I die.
H. Robbins, Amhersty, June 1990

WHERE WE INTRODUCE the sequential selection problem known as *Robbins' problem (of optimal stopping).*

1 Historical background

Robbins' problem belongs to an interesting class of optimal selection problems which are sometimes referred to as *Secretary Problems*. These are problems whose statement starts roughly as follows. *"A decision maker sequentially observes a given number, say n, of realizations of random variables. At each time $i \geq 1$ he must decide whether or not to reject the current observation, say X_i, and examine the next observation, or to accept X_i and therefore reject all subsequent observations. His objective is to maximize a specified payoff function.* Solving a secretary problem requires describing (i) the form of the optimal rule, (ii) the value of this optimal rule (as a function of the number of observations), and (iii) the limiting value and behavior as n grows to infinity.

Behind the apparent triviality of the above problem lies a fundamental mathematical question that has been in the mathematical community since the days of the English mathematician Arthur Cayley (1821-1895). Answering such questions brings to light the "unreasonable efficiency of mathematics", since such problems deal with the question of optimal action under worst circumstances. In other words, solving a secretary problem answers the question of what one can do to get the best out of a situation when we have little or no information about what is going to happen to us.

The interest in such problems was kindled by Martin Gardner, who posed its simplest version as a recreational problem in his column of the *Scientific American* (see Gardner 1960). Rapidly thereafter the problem spread its way around the mathematical community, each new author bringing a different light on the implications and ramifications which lie behind this seemingly anecdotic mathematical game. From that time on "it has been taken up and developed by a number of eminent probabilists and statisticians [...]" and it has spawned a whole class of problems which now "[...] constitute a 'field' of study within mathematics-probability-optimization[1]."

Now although the theory of optimal stopping (as described in Chow et al (1971)) provides in principle a solution to (nearly) all such problems through the method of *backward induction*, obtaining tractable descriptions of the optimal stopping rules and values requires far more than the simple application of a standard tool. In fact, each variation on the above problem requires developing ad hoc methods, and therefore demands for elegant and innovative thinking. This perhaps explains the intense research on these problems by such eminent mathematicians as Dynkin, Ferguson, Robbins, Samuels and many others (see the survey papers by Freeman (1983), Petruccelli (1988) or Samuels (1991)).

2. The classical secretary problems

Although it would be fruitless to recall all the existing versions of the above problem, it is necessary to inscribe our problem within its context, i.e. that of the *the four classical* secretary problems. Of these, the first three were solved successively by Lindley (1961), Chow, Moriguti, Robbins and Samuels (1964), and Gilbert and Mosteller (1966). The fourth problem, Robbins' problem, remains to this date unsolved.

2.1 The no-information best-choice problem

Consider a situation where an employer has advertised an opening for a secretary. There are a known number, n, of applicants, and the employer interviews them one at a time. He is very specific about the qualities that are needed for the job so that, after each interview, he can rank the present applicant with respect to all previous applicants with no ties. The applicant

[1]Ferguson (1984).

must be told immediately after each interview whether or not he has been hired and there cannot be any regrets later on. Moreover if the first $n-1$ applicants have been rejected, then the employer is forced to hire the last one.

What selection strategy will maximize the probability of the employer selecting the best candidate? What is the maximal probability of choosing the best candidate? In particular, what is the limiting value of this probability when the number of applicants becomes infinite?

These questions are the essence of the problem that has come to be known as the *classical no-information secretary problem* (abbreviated CSP), where the terminology "no-information" refers to the fact that the decisions of the employer must be based solely on the relative ranks of the different observations and not on their specific values.

The first solution of the CSP to be published in a scientific journal is due to Lindley (1961). It is obtained by simple backward recursion and states that if, for $r = 1, 2, \ldots, n-1$, we define

$$a_r = \frac{1}{r} + \frac{1}{r+1} + \ldots + \frac{1}{n-1},$$

then the optimal action is to ignore any candidate who is not the best so far, and, if the rth candidate is the relative best at the time at which it is observed, then he should be chosen if $a_r < 1$ and rejected if $a_r > 1$ (see also e.g. Gilbert and Mosteller (1966) or the survey papers Freeman (1983), Petruccelli (1988), or Samuels (1991)). Thus, if $r^\star$ is the first integer for which $a_{r-1} \geq 1 > a_r$, the optimal strategy is to reject the first $r^\star - 1$ applicants and then to accept the first applicant thereafter that is better than all previous applicants.

The probability of the employer selecting the best candidate with this policy is given by $(r^\star - 1)a_{r^\star - 1}/n$, and integral approximation yields that for n large, $r^\star/n \approx e^{-1}$. With this result, one shows that the asymptotic optimal win probability is given by $1/e = 0.368....$ We see that with very little information, the employer still has a surprisingly high probability of obtaining the overall best applicant. This result is perhaps even more striking when one notices that when there are, for example, 100 candidates, then it is (approximately) optimal to reject the first 36 applicants and to hire the first candidate thereafter who is relative best. The probability of selecting a record candidate is even much higher than $1/e$, typically over 60%.

Remark 1 *An alternative solution is due to Dynkin (1963), who considers this problem as an application of the theory of Markov stopping times. In this setting, the optimal strategy is given by the one-stage look-ahead rule. There was a third different solution obtained by Rasche (1975). This solution is a corollary of the more general Odds-Theorem of optimal stopping, see Bruss (2000).*

2.2 The full-information best-choice problem

Consider the following situation. "An urn contain n tags, identical except that each has a different number printed on it. The tags are to be successively drawn at random without replacement (the $n!$ permutations are equally likely). Knowing the number of tags, a player must choose just one of the tags, his object being to choose the one with the largest number. The player's behavior is restricted because after each tag is drawn he must either choose it, thus ending the game, or permanently reject it. The problem is to find the strategy that maximizes the probability of obtaining the largest tag and to evaluate that probability[2]."

Let us suppose that the numbers on each tag (say $X_1, \ldots, X_n$) are a random sample from some specified continuous distribution (which we can take to be the uniform continuous distribution on $[0, 1]$ since we are only interested in the comparative quality of each tag). Clearly, this problem is equivalent to the CSP if the player only considers the relative ranks of the numbers on the tags, since the hypothesis of continuity of the distribution guarantees the absence of ties. But suppose now that at each stage i, the player is allowed to know the values $X_1, \ldots, X_i$. Then his decisions are to be based on a more informative data set, and thus the optimal win probability should be better than in the classical no-information problem.

As an illustration of this fact, let us consider the case $n = 2$ and let X_1, X_2 be the first and second numbers examined, respectively. In the no-information problem, the player does not have much of a choice, since the first arrival will always be of relative rank one (it is obviously the best so far) and thus the player will simply win with one chance out of two. Suppose now that the player is allowed full information on the problem and let us

[2]This presentation of the problem is the same as that in Gilbert and Mosteller (1966). It is equivalent to the presentation made by Gardner, see Gardner (1960), under the name 'game of Googol'.

choose any number x between 0 and 1. We define the rule

$$\tau_x = \begin{cases} 1 \text{ if } X_1 > x \\ 2 \text{ otherwise.} \end{cases}$$

If both X_1 and X_2 are greater or smaller than x, this rule selects the larger of the two with probability $1/2$. If not, then this rule necessarily selects the maximum of the two. Hence the win probability with the rule τ_x is given by

$$\mathrm{P}[X_{\tau_x} = \max(X_1, X_2)] = 1/2 + x - x^2,$$

which is always greater than $1/2$ and equal to $3/4$ for $x = 0.5$.

The optimal strategy for all n (say $\tau_n^\star$) was obtained by Gilbert and Mosteller (1966). These authors show that it is defined through a sequence of thresholds, which they call *decision numbers*, $b_0 = 0, b_1, b_2, \ldots$, not depending on n such that

$$\tau_n^\star = \min_{1 \le i \le n} \left\{ i : \; X_i = \max_{j \le i} X_j \; \text{ and } \; X_i \ge b_{n-i} \right\}.$$

Each decision number b_m, $m = 1, 2, \ldots$ is solution to

$$\sum_{j=1}^{m} j^{-1} b_m^{-j} = 1 + \sum_{j=1}^{m} j^{-1},$$

where, as one would expect from the example we gave above, $b_1 = 1/2$. These numbers form an increasing sequence which goes to one as the number of draws becomes large.

Now let $w_n = \mathrm{P}[X_{\tau_n^\star} = \max\{X_1, \ldots, X_n\}]$ denote the win probability under the optimal strategy. Samuels (1982) showed that w_n is strictly decreasing in n and that

$$\lim_{n \to \infty} w_n \approx 0.580164...$$

Hence we see that there is an improvement of roughly 58% from the no-information to the full-information problem.

2.3 The no-information expected rank problem

We consider the same situation as in the classical secretary problem, in which an employer interviews n candidates for a job under the restriction that, at each interview, the only information he can work on is the relative ranks of the preceding applicants. Now suppose that instead of maximizing the probability of selecting the best, we consider the objective of minimizing the total expected rank of the selected candidate, where the overall best candidate is given rank one, the second best two, etc., and the worst rank n.

This objective is arguably more realistic than that of the CSP. Indeed, maximizing the probability of accepting only the best candidate implies a utility function that takes the value 1 if the best is accepted and 0 otherwise. Such 'nothing-but-the-best'[3] objectives are therefore very restrictive in comparison to real-life problems in which one could imagine that an employer would also be satisfied with a less perfect candidate. In this respect, a more appropriate utility function would be that which takes the value $n-i$ if the ith best candidate is accepted; maximizing the expected value of this utility function corresponds to minimizing the expected rank of the selected observation.

With this in mind, it is now easy to surmise that the optimal strategy for the best-choice problem is no longer optimal with respect to this new objective. Indeed, although this policy selects an arrival which has absolute rank 1 with a high probability, it also has a major drawback: it suffices for the overall best candidate to appear among the first $r^\star - 1$ applicants to ensure that this strategy never stops and thus selects the last candidate. Since the last candidate has expected rank $\frac{n+1}{2}$, we infer that this policy must be suboptimal in this setting.

The optimal strategy for this problem can for example be obtained by the method of *backward induction* (see e.g. Chow et al. (1971)). Labeling the relative ranks of each applicant by $r_1, r_2, \ldots r_n$ respectively, a direct application of this method shows (see e.g. Lindley (1961) or Chow et al. (1966)) that the optimal strategy is given by a sequence of thresholds $s_1 \leq s_2 \leq \ldots \leq s_n = n$ such that it is optimal to stop on the first applicant whose relative rank satisfies $r_i \leq s_i$. However, it turns out that the recurrence equations which define the s_i's are extremely complicated and thus this result lends little insight into the asymptotic value of the optimal expected rank.

A heuristic argument given in Lindley (1961) indicated that, by approximating these recurrence relations by a differential equation, the optimal expected rank should approach a finite limit as n goes to infinity. Chow et al. (1964) were able to make this rigorous and obtained then the limiting form of the expected rank under the optimal policy. For this they showed that the minimum expected rank for the n arrival problem is a strictly in-

[3] This appellation is due to Lindley (1961).

creasing function of n, and that it converges to

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

It is here of interest to point out that H. Robbins was co-author of this paper.

2.4 The full-information expected rank problem

This problem has the same formulation as the full-information best choice problem but instead of maximizing the probability of obtaining the best X_i, the objective is now to minimize the expected rank of the selected observation. One sees that this problem fits perfectly in the two-by-two pattern of the classical secretary problems. Surprisingly, although the three previous problems had been solved by the mid 60's, it was not until Professor Herbert Robbins (1915-2001) kindled the interest of the mathematical community (at the International Conference on Search and Selection in Real Time in 1990) that results were published on this problem. For this reason among others, it has been named in his honor (see Bruss and Ferguson (1993) and Assaf and Samuel-Cahn (1996); for a review see Bruss (2005)).

We defer a precise definition of this problem to Chapter 1. However, we can already deduce an upper bound on the optimal expected rank from the previous remarks. Indeed, a player with full-information can only do better than a player with no information, since he can always choose to use a strategy which only considers the relative ranks of the arrivals. Therefore, letting $v(n)$ be the value of the optimal expected rank for the n-arrival full-information expected rank problem, we know that

$$\lim_{n\to\infty} v(n) \leq 3.8695...$$

Now, in light of the fact that the passage from no-information to full-information in the best choice problems yielded a 58% increase in the asymptotic win probability, it seems reasonable to hope that the improvement for the rank problem should be of the same order, i.e. that $\lim w(n) \approx 2.44$. We will see that the improvement is, in fact, better.

3. Moser's problem

Before moving on, we need to relate Robbins' problem to one final example of selection problem which we will refer to as *Moser's problem*, in honor of Professor Leo Moser, who was the first to obtain its solution (see Moser (1956)). This problem is an extension of a problem posed by Cayley in 1875 (see Cayley (1875)). Although it is not *per se* a 'secretary problem', we will see that it yields some necessary intuitions on Robbins' problem.

The problem is as follows. A player observes sequentially n random variables $X_1, \dots, X_n$ which are known to be independent, identically and uniformly distributed on $[0, 1]$. His cost for stopping at time j is equal to the value of the observation, and no recall of preceding observations is permitted. His objective it to use an adapted stopping rule τ which minimizes $\mathrm{E}[X_\tau]$.

Moser (1956) shows that if we define recursively the sequence $(a_k)_{k\geq 0}$ by $a_0 = 1$ and

$$a_{j+1} = a_j - \frac{1}{2}a_j^2, \ j \geq 1,$$

then it is optimal to stop on X_j if $X_j \leq a_{n-j}$, i.e. the optimal strategy $\hat{\tau}_n$ is

$$\hat{\tau}_n = \min\{k \geq 1 \mid X_k \leq a_k\}.$$

The optimal return with this strategy is given by $\mathrm{E}[X_{\hat{\tau}_n}] = a_n$. These thresholds are asymptotically equal to

$$a_k \approx \frac{2}{n-k+1} \wedge 1,$$

and for large n,

$$a_n \approx \frac{2}{n + \log n + O(1)},$$

so that

$$\lim_{n\to\infty} \mathrm{E}[nX_{\hat{\tau}_n}] = 2.$$

In the sequel, we will often refer to this problem, and to the optimal rule $\hat{\tau}_n$.

4. Acknowledgments and outline of the mémoire

The following work is the outcome of our efforts on the full-information expected rank problem. The original research which is recalled in the next few pages was conducted in close collaboration with Professor F. T. Bruss from the ULB, and the author of this manuscript gratefully acknowledges that without Professor Bruss' insight and experience, nothing would have been achieved. We are also grateful to Prof. F. Delbaen, from ETH Zurich,

and to Prof. M. Drmota from TU Wien for their help and interesting discussions. Part of this work has been published, in 2009, under the title *"A continuous time approach to Robbins' Problem of minimizing the expected rank"*. Also recorded are a number of results which, though not yet published, we deem of interest.

We must, sadly, warn our reader that he will not find in these next few pages a solution to the problem. We do hope, however, that after reading this account, our reader will have the flash of intuition that we have not had and perhaps find the missing piece of information that yields the solution. Despite our lack of success, we can nevertheless proudly present a number of fundamental results on a problem which has so far baffled all those who have tackled it; moreover we believe that herein lies the hope for a definite answer.

The outline of this mémoire is as follows. In Chapter 1 we present a number of known results as well as a fundamental inequality linking expected values and expected ranks in optimal selection problems. Chapter 2 describes continuous-time approach to the problem. Our third chapter provides the link between the results of the first two chapters, while the fourth and concluding chapter provides some analytic results which we believe to contain the solution. A list of references concludes the manuscript.

Chapter 1

The classical Robbins' Problem

WHERE WE SUMMARIZE the known results on Robbins' Problem, and provide a crucial relationship between ranks and values.

1.1 Definition and notations

A player observes sequentially n i.i.d random variables $X_1, \ldots, X_n$ distributed uniformly on $[0,1]$ and has to chose exactly one of them. The objective of the player is to minimize the expected rank of the chosen observation, where the best observation is given rank one, the second best rank two, etc., and the worst rank n. However, once a value is rejected, it cannot be recalled afterwards, so that at time k, only X_k can be selected, and the data on which the decision is made are the values of the arrivals up to time k. Let $\mathcal{F}_k = \sigma(X_1, \ldots, X_k)$ be the σ-algebra generated by $X_1, \ldots, X_k$. The *relative rank* of an arrival X_k is defined by

$$r_k = \sum_{j=1}^{k} \mathbf{1}_{\{X_j \le X_k\}},$$

and the *(absolute) rank* of the kth observation is defined by

$$R_k^\star = \sum_{j=1}^{n} \mathbf{1}_{\{X_j \le X_k\}}.$$

Since $R_k^\star$ is not $\mathcal{F}_k$-measurable, we replace it by

$$R_k = \mathrm{E}[R_k^\star | \mathcal{F}_k] = r_k + (n-k)X_k.$$

The objective of the player is to use a non anticipating strategy τ which minimizes $\mathrm{E}[R_\tau]$ (where this problem is clearly equivalent to that of minimizing $\mathrm{E}[R^\star_\tau]$ since the corresponding expressions are equal for all stopping rules τ). Now let $T_n = \{\tau : \{\tau = k\} \in \mathcal{F}_k,\ \forall k = 1, 2, \ldots, n\}$ be the set of stopping rules adapted to $X_1, \ldots, X_n$, and define the *value function* for n arrivals by

$$v(n) = \inf_{\tau \in T_n} E[R_\tau]. \tag{1.1}$$

Robbin's problem consists in studying the value function $v(n)$ defined by equation (1.1), the stopping rule $\tau^\star = \tau^\star_n$ which achieves $v(n)$ and the asymptotic value

$$v = \lim_{n\to\infty} v(n). \tag{1.2}$$

1.2 Memoryless threshold rules

Before studying $v(n)$ and the optimal rule in T_n, it is appropriate to introduce the subclass of *memoryless threshold rules.* These are stopping rules which are defined for each n through a sequence of constants (called *threshold constants*) $0 \le a_{n,1} \le a_{n,2} \le \ldots \le a_{n,n} = 1$ by

$$\tau_n = \min\{k : X_k \le a_{n,k}\}. \tag{1.3}$$

The restriction $a_{n,n} = 1$ is necessary in order to ensure that rules defined by (1.3) stop for at least one of the arrivals X_k, and hence for exactly one.

Assaf and Samuel-Cahn (1996) prove that for any discrete memoryless threshold rule defined by a sequence which is not monotone increasing, there exists a rule determined by a monotone increasing sequence which yields a better value. Thus only monotone increasing sequences need to be considered. This statement is, of course, rather transparent.

Now let M_n be the set of all such rules, and for all $\tau_n \in M_n$, denote the value of Robbins' problem under τ_n by

$$V(\tau_n) = \mathrm{E}[R_{\tau_n}]. \tag{1.4}$$

Straightforward computations yield the following lemma (see Bruss and Ferguson (1993) and Assaf and Samuel-Cahn (1996)).

Lemma 1.1 *Consider the threshold sequence* $0 < a_1 \le a_2 \le \ldots \le a_n = 1$ *and let* τ_n *be the corresponding strategy. Then if* $a_{k-1} < 1$,

$$V(\tau_n) = 1 + \frac{1}{2}\sum_{k=1}^{n-1}(n-k)a_k^2\prod_{j=1}^{k-1}(1-a_j) + \frac{1}{2}\sum_{k=1}^{n}\prod_{j=1}^{k-1}(1-a_j)\sum_{j=1}^{k-1}\frac{(a_k-a_j)^2}{1-a_j} \tag{1.5}$$

where $0/0$ *should be interpreted as 0 in the last sum.*

We define the restricted value function $V(n)$ (with a capital 'V') as the optimal value of $V(\tau_n)$ among all $\tau_n \in M_n$ (i.e. it is the minimal expected rank attainable through a memoryless threshold rule) and we define the restricted asymptotic value

$$V = \lim_{n\to\infty} V(n), \tag{1.6}$$

if it exists. Note that $V(n)$ gives an upper bound on $v(n)$ for all n, and hence $v \leq V$.

The following results are well-known, and due to Bruss and Ferguson (1993, 1996), and Assaf and Samuel-Cahn (1996). They are therefore stated without proof or explanation.

Theorem 1.2 *There exists an optimal rule* $\tau_n^\star \in M_n$*, i.e. there exists a memoryless threshold rule* $\tau_n^\star$ *for which* $V(\tau_n^\star) = V(n)$*. Moreover this rule is uniquely defined.*

Theorem 1.3 $V(n)$ *is an increasing and bounded function of* n *and hence the limit* $V = \lim_{n\to\infty} V(n)$ *exists and is finite.*

Theorem 1.4 *For any stopping rule* τ_n *let*

$$U_n(\tau_n) = \left(2\mathrm{E}[nX_{\tau_n}](1 + \mathrm{E}[\tau_n/n])\right)^{\frac{1}{2}}. \tag{1.7}$$

Also let $U^\star = \liminf U_n(\tau_n^\star)$ *where* $\tau_n^\star$ *is the optimal memoryless threshold rule. Then*

$$V = U^\star, \tag{1.8}$$

and

$$V \geq \liminf_{n\to\infty} \inf_{\tau_n \in M_n} U_n(\tau_n). \tag{1.9}$$

These results are, obviously, of great importance for our problem. We will return to them, and to their consequences, in Section 1.4.

1.3 The optimal rule

Backward induction (see Chow. et al (1971)) guarantees the existence of an optimal strategy $\tau_n^\star \in T_n$ for all n, and provides, in principle, a way to compute it. However, even for small values of $n \geq 3$, computing the optimal strategy through backward induction is a formidable task which does not give any intuition on the asymptotic value (see Assaf and Samuel-Cahn (1996) for the case $n = 3$).

Bruss and Ferguson (1996) prove that the optimal rule is a threshold rule of the form

$$\tau_n^\star = \inf\{1 \leq k \leq n : X_k \leq p_k^{(n)}(X_1, X_2, \ldots, X_k)\},$$

where the functions $p_k^{(n)}(.)$ are *fully history dependent* in the sense that for each k, the value of the corresponding threshold depends on every arrival $X_1, \ldots, X_k$. They also show that no nontrivial statistic of $X_1, \ldots, X_k$ is sufficient to achieve the optimal value $v(n)$, and hence the optimal thresholds have an unbounded number of arguments and "figure in the list of most undesirable mathematical objects[1]".

Now although this property of the optimal rule seems to demonstrate that the problem is not tractable, Bruss and Ferguson (1993) also prove that this problem possesses some monotonic features. More precisely they show that the optimal thresholds are stepwise-monotone-increasing in the sense that for each n and for all $k = 1, \ldots, n-2$,

$$0 \leq p_k^{(n)}(X_1, X_2, \ldots, X_k) \leq p_{k+1}^{(n)}(X_1, X_2, \ldots, X_k, X_{k+1}) < p_n^{(n)} = 1$$

almost surely. They also show that the value function $v(n)$ is increasing in n. In particular this proves that the limiting value $v = \lim_{n\to\infty} v(n)$ exists, since, as we have already seen, the value function is bounded above for all n by the value function in the corresponding no-information problem.

1.4 Bounds on the value functions

We know that $M_n \subset T_n$ and thus the asymptotic value $\lim_{n\to\infty} V(\tau_n)$ for any sequence of memoryless threshold rules $(\tau_n)_{n\geq 1}$ gives an upper bound on v. For example, by considering a version of Moser's rule of the form

$$\tau_n = \min\left\{k \geq 1 : X_k \leq \frac{2}{n-k+2}\right\},$$

[1] Bruss (2005)

Bruss and Ferguson (1993) show by use of integral approximation of (1.5) that

$$v \leq \lim_{n\to\infty} V(\tau_n) = \frac{7}{3}.$$

Assaf and Samuel-Cahn (1996) go one step further and obtain a limiting form of (1.5) for certain thresholds. This allows them to compute the limiting value $\lim_{n\to\infty} V(\tau_n)$ for rules of the form

$$\tau_n = \min\left\{k \geq 1 : X_k \leq \frac{g(k)}{n-k+c}\right\},$$

with $g(z) = \sum_{j=0}^{m} c_j z^j$. Direct computation with these expressions prove that taking $m = 2$ and $c_0 = 1.77, c_1 = 0.54$ and $c_2 = -0.27$ yields

$$V \leq \lim_{n\to\infty} V(\tau_n) = 2.3267;$$

this is the best upper bound they obtain.

Such small improvements are however not unexpected. Indeed, Bruss and Ferguson (1993) further extrapolate that the limiting value is given by $V = 2.32659$. Our own explorations of the above rules have yielded a memoryless rule which achieves 2.32661. Also, standard results from calculus of variations applied to equation (1.7) of Theorem 1.4 allow Assaf and Samuel-Cahn to show that $\inf_{\tau_n \in M_n} U_n(\tau_n) = 2.29558...$, so that we now know

$$2.29558... \leq V \leq 2.3267....$$

Interestingly the statement of Theorem 1.4 can be further strengthened to show that for all sequences of stopping rules τ_n satisfying $\lim E[R_{\tau_n}] < \infty$ one has

$$\lim E[R_{\tau_n}] \geq \liminf U(\tau_n) = \liminf \left(2E[nX_{\tau_n}](1 + E[\tau_n]/n)\right)^{1/2}. \quad (1.10)$$

Although this last result, in a sense, hints towards the main result of this chapter (see Theorem 1.5), a little reflection shows that this inequality is in no way strong enough to yield any positive result other than that obtained by Assaf and Samuel-Cahn. In fact, it does not even suffice to prove that $nE(X_{\tau_n^\star})$ is bounded! It nevertheless brings some intuition. For instance, plugging $\hat{\tau}_n$, the optimal strategy for Moser's problem, into (1.10) allows us to conclude

$$\lim U(\hat{\tau}_n) = \sqrt{16/3} = 2.3094.$$

These results concur to suggest that the quest for the optimal memoryless value is essentially over.

The bounds given above obviously provide upper bounds on v, a value which we know exist. They do not tell us, however, what we need to know, i.e. they do not allow for proving (or disproving) the equality $v = V$. This question is still open, although Gnedin (2007) has recently proved that, in a limiting Poisson model similar to the one we will present in the next chapter, the full-history dependence persists. Such a result only holding within the framework of Gnedin's (specific) model, it is of limited interest to us here.

The best information on this aspect of the problem is due to Bruss and Ferguson (1993), who consider a sequence of truncated games which yield systematically lower payoff than the original problem. Although these truncated problems represent considerable simplifications over the original problem, the authors report that the computational aspects involved with this approach are still severe. Some computations were nevertheless carried out for $m = 1$ through $m = 5$, and all they were able to conclude is that

$$v \geq 1.908.$$

1.5 Ranks and values

A natural question, to which we have already hinted, is that of the link between Moser's and Robbins' problem. Of course we know that, for finite n, minimizing the value is different from minimizing the rank since we know that the optimal rule for minimizing the value is memoryless, and can be strictly improved upon in the rank problem by taking into account the history of the process. Naturally, it is also important to verify whether or not this still holds true as n grows. A first step towards answering this fundamental question is to determine whether, for $\tau_n^\star$ the optimal (non threshold) rule we have $nE[X_{\tau_n^\star}$ bounded. Now although such an affirmation is rather transparent (indeed, how could one imagine a strategy which minimizes the rank without selecting a value which is at least smaller than $1/n$?), we have not been able to prove it. The best we have obtained is the following.

Theorem 1.5 *Let $\tau_n^\star$ be the optimal strategy for the discrete n-arrival Robbins' Problem. Then, for all $p > 1$,*

$$E(X_{\tau_n^\star}) \leq E(R_{\tau_n^\star}) \left(\sum_{k=1}^{n} a_{n,k}^p \right)^{\frac{1}{p}}, \tag{1.11}$$

where

$$a_{n,k}^p = k^{2-p}\binom{n}{k}\int_0^1 x^{p+k-1}(1-x)^{n-k}dx. \tag{1.12}$$

Proof: Conditioning on the ranks we obtain

$$E(X_{\tau_n^\star}) = E(E(X_{\tau_n^\star} \mid R_{\tau_n^\star})) = \sum_{k=1}^n E(X_{(k)}\mathbf{1}_{\{R_{\tau_n^\star}=k\}})$$

where $X_{(k)}$ is the kth smallest order statistic of the sample $X_1,\ldots,X_n$. Note that the expectation cannot be factorized. Applying Hölder's inequality yields

$$E(X_{\tau_n^\star}) \le \sum_{k=1}^n (E(X_{(k)}^p))^{1/p}(P(R_{\tau_n^\star}=k))^{1/q},$$

for all p, q such that $\frac{1}{p}+\frac{1}{q}=1$. Hence

$$E(X_{\tau_n^\star}) \le \sum_{k=1}^n k^{-1/q}(E(X_{(k)}^p))^{1/p}(kP(R_{\tau_n^\star}=k))^{1/q}.$$

Now let $a_{n,k} = k^{-1/q}(E(X_{(k)}^p))^{1/p}$. Applying Hölder's inequality this time to the sum on the rhs of the above inequality for the same choice of p and q gives

$$E(X_{\tau_n^\star}) \le \left(\sum_{k=1}^n a_{n,k}^p\right)^{1/p}\left(\sum_{k=1}^n kP(R_{\tau_n^\star}=k)\right)^{1/q} = \left(\sum_{k=1}^n a_{n,k}^p\right)^{1/p} E(R_{\tau_n^\star})^{1/q}.$$

Since $E(R_{\tau_n^\star}) \ge 1$ for all n, we obtain

$$E(X_{\tau_n^\star}) \le \left(\sum_{k=1}^n a_{n,k}^p\right)^{1/p} E(R_{\tau_n^\star}).$$

The known expressions for the moments of the kth order statistics then yield (1.12). ■

Note that nowhere do we use the properties of $\tau^\star$ in the proof of Theorem 1.5. Hence this result is valid for *every* stopping rule, and we get the following (rather intuitive) result.

Corollary 1.6 *Let $0 \leq \alpha < 1$, and let τ_n be a sequence of stopping rules for which* $\lim_n E(R_\tau) < \infty$. *Then*

$$\lim_{n\to\infty} n^\alpha E(X_{\tau_n}) = 0$$

for all $0 < \alpha < 1$.

Proof: It suffices to show that the constants $\left(\sum_{k=1}^n a_k^p\right)^{1/p}$ are well behaved for n large. For this, first fix some integer $p > 1$. Computing the $a_{n,k}^p$ explicitly, we obtain

$$a_{n,k}^p = k^{1-p} \frac{n!}{(n+p)!} \frac{(k+p-1)!}{(k-1)!}.$$

Applying Stirling's approximation

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/(12n+1)} \leq n! \leq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/(12n)}$$

we get

$$\sum_{k=1}^n a_{n,k}^p = A(n,p)B(n,p),$$

where

$$A(n,p) = \sqrt{\frac{n}{n+p}}\left(\frac{n}{n+p}\right)^n e^p \left(\frac{1}{n+p}\right)^p f_1(n,p),$$

with $f_1(n,p) \approx 1$ for all p when n is large, and

$$B(n,p) = p! + e^{-p}\sum_{k=2}^n k\sqrt{1+\frac{p}{k-1}}\left(1+\frac{p}{k-1}\right)^{k-1}\left(1+\frac{p-1}{k}\right)^p f_2(k,p),$$

with $f_2(k,p) \leq 1$ for all k and p.

It is easy to check that, for n large,

$$A(n,p) \approx \left(\frac{1}{n+p}\right)^p. \tag{1.13}$$

Also, since $\left(1+\frac{p}{k-1}\right)^{k-1} \leq e^p$, we get

$$B(n,p) \leq p! + \sqrt{1+p}\sum_{k=2}^n k\left(1+\frac{p-1}{k}\right)^p$$

so that, using Newton's Binomial formula on $\left(1+\frac{p-1}{k}\right)^p$, one obtains

$$B(n,p) \leq p! + \sqrt{1+p}\left(\sum_{k=2}^{n} k + \sum_{l=1}^{p} \binom{p}{l} (p-1)^l \sum_{k=2}^{n} \frac{1}{k^{l-1}}\right).$$

For fixed $p > 1$ this last function is in the order of n^2, and hence, from (1.13),

$$\left(\sum_{k=1}^{n} a_k^p\right)^{1/p} = A(n,p)B(n,p) \approx K_p n^{2/p-1} \tag{1.14}$$

for some positive constant K_p. ■

These results will play a fundamental rôle in the sequel, and we will return to them in Chapter 3.

Chapter 2

Embedding Robbins' Problem

Where we consider a variation on Robbins' problem with a random number of arrivals.

2.1 Definition and notations

We first consider the following variation on the original problem. A decision maker observes opportunities occurring according to a planar Poisson process of homogeneous rate 1. He inspects each option when the opportunity arises and has to chose exactly one before a given time t. Decisions are to be made immediately after each arrival, and no recall of preceding observations is permitted. The loss incurred by selecting an arrival is defined at time t as the total number of observations in $[0, t]$ which are smaller than the selected observation. If no decision has been reached before the given time t, then his loss is equal to some function of t, say $\Pi(t)$. At all times the decision maker has the knowledge of the full history of the process, and his objective is to use a non anticipating strategy which will minimize the expected loss.

Formally, this problem can be translated as follows. Fix $t > 0$. Let $(T_1, X_1), (T_2, X_2), \ldots$ denote a point arrival process with $T_1 \leq T_2 \leq \ldots$ being the arrival times of a homogeneous Poisson counting process $(N(s))_{s \geq 0}$ of rate 1, and $X_1, X_2, \ldots$ the associated i.i.d. random values . With this notation, define the absolute rank of the kth arrival X_k (with respect to t)

by

$$R_k^{(t)} = \sum_{j=1}^{N(t)} \mathbf{1}_{\{X_j \le X_k\}},$$

where the sum is set to 0 if $N(t) = 0$. The loss incurred for selecting X_k at time T_k is then

$$\tilde{R}_k^{(t)} := R_k^{(t)} \mathbf{1}_{\{T_k \le t\}} + \Pi(t) \mathbf{1}_{\{T_k > t\}}, \tag{2.1}$$

and the objective of the decision maker is to use a stopping time τ which minimizes $E(\tilde{R}_\tau^{(t)})$. Since we only allow stopping upon inspection, the set of adapted strategies is restricted to the collection $\mathcal{T}$ of all random variables with values in the set $\{T_r\}_{r\ge1}$ of arrival times of the point process, which satisfy $\{\tau \le s\} \in \mathcal{F}_s$ where

$$\mathcal{F}_s = \sigma \left\{ (N(u))_{0\le u\le s}, (T_1, X_1) \ldots, (T_{N(s)}, X_{N(s)}) \right\},$$

and where it is understood that $\mathcal{F}_s = \sigma\{(N(u))_{0\le u\le s}\}$ for all s for which $N(s) = 0$. Such stopping rules are called "canonical stopping times" in Kühne and Rüschendorf (2000) or Gnedin (2007).

Remark 2 *We will from now onwards* always *use the notation* $\{\tau = k\}$ *instead of* $\{\tau = T_k\}$ *to denote the event that the decision maker selects the* k*th arrival. Hence the notations* R_τ, X_τ *and* T_τ *are well defined and will be used systematically throughout the paper.*

The Poisson embedded Robbins' Problem consists in studying the value function $w(t)$ defined by

$$w(t) = \inf_\tau E\left(\tilde{R}_\tau^{(t)}\right) = \inf_\tau E\left(R_\tau^{(t)} \mathbf{1}_{\{T_\tau \le t\}} + \Pi(t) \mathbf{1}_{\{T_\tau > t\}}\right), \tag{2.2}$$

including its asymptotic value $w = \lim_{t\to\infty} w(t)$, if it exists, as well as the stopping rule $\tau_t^\star$ which achieves this value.

Remark 3 *The function* $\Pi(t)$ *reflects the loss incurred for selecting no observation before time* t*. We call it the* penalty function*. Although we keep this function unspecified throughout the text, we suppose that* $\Pi(0) = 0$ *and that* $\Pi(.)$ *is increasing and differentiable with bounded derivative. Hence this function is Lipschitz-continuous and satisfies*

$$\lim_{t\to\infty} \frac{\Pi(t)}{t} \le \kappa, \tag{2.3}$$

for some $\kappa \in (0, \infty)$.

Remark 4 *Note that for all τ, the expected rank of an arrival selected by τ before the horizon t satisfies*

$$E(R_\tau^{(t)}) = E(E(R_\tau^{(t)} \mid \mathcal{F}_{T_\tau})) = E(r_\tau + (t - T_\tau)X_\tau)$$

where $r_k = \sum_{j=1}^{k} \mathbf{1}_{\{X_j \le X_k\}}$ is the relative rank of the kth observation. Hence, although the absolute ranks $R_k^{(t)}$ are not measurable with respect to $\mathcal{F}_{T_k}$, the problem of minimizing the loss among all adapted stopping rules is well defined via that of minimizing $E(r_\tau + (t - \tau)X_\tau)$, as already seen in Bruss and Ferguson (1993) and Assaf and Samuel-Cahn (1996).

2.2 Memoryless threshold rules

For each $t > 0$, we define the set of *threshold functions* on $[0, t]$ as the set of all functions $g_t : \mathbb{R} \to [0, 1] : s \mapsto g_t(s)$ such that $g_t(s) = 1 \; \forall s \ge t$. To each such function we associate (uniquely) a *memoryless threshold rule* σ_t

$$\sigma_t = \inf\{i,\; i = 1, 2, \ldots \text{ such that } X_i \le g_t(T_i)\}, \tag{2.4}$$

and a value

$$W(\sigma_t) = \mathrm{E}[\tilde{R}_{\sigma_t}] = \mathrm{E}[R_{\sigma_t}\mathbf{1}_{\{T_{\sigma_t} \le t\}} + \Pi(t)\mathbf{1}_{\{T_{\sigma_t} > t\}}]. \tag{2.5}$$

Let $\mathcal{M}_t$ be the set of all such rules (note the indexing in t). We define the restricted value function $W(t)$ as the minimal value of $W(\sigma_t)$ obtainable on $\mathcal{M}_t$, i.e.

$$W(t) = \inf_{\sigma_t \in \mathcal{M}_t} W(\sigma_t). \tag{2.6}$$

We also denote the restricted asymptotic value (if it exists) by

$$W = \lim_{t \to \infty} W(t). \tag{2.7}$$

Clearly, for all t, we have $\mathcal{M}_t \subset \mathcal{T}$. Therefore, as in the discrete problem, the restricted optimal value $W(t)$ gives an upper bound on the optimal value $w(t)$ for all t, and thus the corresponding limits, if they exist, must satisfy $w \le W$.

Our next result shows that, as in the discrete case, only increasing thresholds need to be considered. Intuitively, this simply translates the fact that if it is optimal to accept an arrival of value x at time s, then it should also be optimal to accept an arrival of smaller value at later times $s' \ge s$.

Proposition 2.1 *Let $\sigma_t^\star \in \mathcal{M}_t$ be the optimal memoryless threshold rule and let $g_t^\star(.)$ be the corresponding threshold function. Then for all $0 \leq s \leq s' \leq t$, $g_t^\star(s) \leq g_t^\star(s')$.*

Proof: Suppose that there is an arrival X_i at time T_i, and define $E(i,t)$ as the minimal expected rank obtainable with memoryless strategies which stop almost surely after the ith arrival, i.e.

$$E(i,t) = \inf_{\sigma_t \in \mathcal{M}_t, \sigma_t > i} \mathrm{E}[\tilde{R}_{\sigma_t}^{(t)} \mid \mathcal{F}_i],$$

where the infimum is taken over the set of all stopping rules $\sigma_t \in \mathcal{M}_t$ such that $\mathrm{P}[\sigma_t > i] = 1$. We know that it is optimal to stop on an arrival (T_i, X_i) if and only if

$$\mathrm{E}[\tilde{R}_i^{(t)} \mid \mathcal{F}_i] \leq E(i,t).$$

Since $\mathrm{E}[\tilde{R}_i^{(t)} \mid \mathcal{F}_i] = r_i + (t - T_i)X_i$, this implies that it is optimal to stop on (T_i, X_i) if and only if it satisfies

$$r_i + (t - T_i)X_i \leq E(i,t). \tag{2.8}$$

Now suppose that $\sigma_t^\star$ is optimal but that $g_t^\star(.)$ is not increasing on $[0,t]$ (as illustrated in Figure 1).

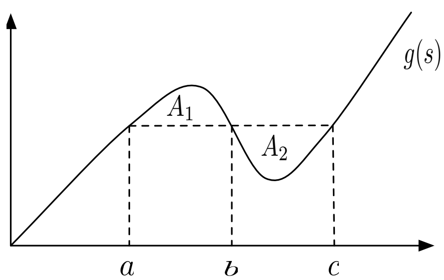

Figure 1: *The threshold function is not monotone increasing and hence we can define the areas A_1 and A_2.*

Since $g_t^\star(.)$ is not increasing, it must be possible to choose areas A_1, A_2 and a, b, c as illustrated in Figure 1. By definition of $\sigma_t^\star$, any arrival in A_1 will be accepted, and any arrival in A_2 will be rejected. Now suppose that the ith arrival (T_i, X_i) lies in A_1. Then this arrival is accepted and must be an optimal choice in the class of memoryless strategies, so that it satisfies equation (2.8), which yields

$$r_i + (t - T_i)X_i \leq E(i,t). \tag{2.9}$$

If the next arrival (T_{i+1}, X_{i+1}) lies in A_2, then, although $X_{i+1} \leq X_i$, it will not be selected by $\sigma_t^\star$. Hence it is not optimal to stop on this arrival, and

$$r_{i+1} + (t - T_{i+1})X_{i+1} > E(i+1, t). \tag{2.10}$$

Since $X_{i+1} \leq X_i$, we must have $r_{i+1} \leq r_i$. Also, under fixed history up to time $i-1$, we see that $E(i,t) \leq E(i+1,t)$. Therefore, from (2.9) and (2.10) we obtain

$$E(i,t) \leq E(i+1,t) < r_{i+1} + (t - T_{i+1})X_{i+1} \leq r_i + (t - T_i)X_i,$$

which in turn yields

$$r_i + (t - T_i)X_i > E(i,t). \tag{2.11}$$

Hence, if the optimal strategy $\sigma_t^\star$ is defined through a non monotone increasing threshold function, we see that there is a positive probability of there being a realization of the process for which equations (2.9) and (2.11) must hold at the same time. This yields a contradiction. ■

From now on we will only consider threshold functions $g_t(.)$ that are monotone increasing on $[0,t]$. Let $g_t(.)$ be such a threshold function on $[0,t]$ and take $\sigma_t \in \mathcal{M}_t$ to be the corresponding memoryless threshold rule (see equation (2.4)). Define the function $\mu_t(s)$ for $s \geq 0$ by

$$\mu_t(s) = \int_0^s g_t(u)du. \tag{2.12}$$

From the properties of homogeneous Poisson processes we see that for all $s \in [0,t]$,

$$\mathrm{P}[T_{\sigma_t} \geq s] = e^{-\mu_t(s)},$$

and hence the density of T_{σ_t} is given on $(0,t)$ by

$$f_{T_{\sigma_t}}(s) = g_t(s)e^{-\mu_t(s)}. \tag{2.13}$$

Hence, conditioning on the arrival time, we see that for all $\sigma_t \in \mathcal{M}_t$,

$$\mathrm{E}[R_\tau^{(t)} \mathbf{1}_{\{T_{\sigma_t} \leq t\}}] = \int_0^t \mathrm{E}[R_\tau^{(t)} \mid T_{\sigma_t} = s] f_{T_{\sigma_t}}(s)ds. \tag{2.14}$$

We now use this equation to obtain an integral version of (1.5) for memoryless threshold rules.

Proposition 2.2 *Let $g_t(s)$ be a continuous increasing threshold function, and let $\sigma_t \in \mathcal{M}_t$ be the corresponding memoryless threshold rule. Let $\mu_t(s) = \int_0^s g_t(u)du$. Then*

$$\begin{aligned} W(\sigma_t) &= 1 + (\Pi(t) - 1)e^{-\mu_t(t)} + \tfrac{1}{2}\int_0^t g_t(s)^2(t-s)e^{-\mu_t(s)}ds \\ &\quad + \tfrac{1}{2}\int_0^t \int_0^s \frac{(g_t(s) - g_t(u))^2}{1 - g_t(u)} du\, e^{-\mu_t(s)}ds \end{aligned} \tag{2.15}$$

Proof: Recall that for a memoryless threshold rule σ_t defined by a function g_t, the density of T_{σ_t} is given on $(0, t)$ by

$$f_{T_{\sigma_t}}(s) = g_t(s)e^{-\mu_t(s)}. \tag{2.16}$$

Now choose $s \in (0, t)$ and suppose that $T_{\sigma_t} = s$. Then, conditionally to $X_{N(s)} = x \in [0, g_t(s)]$, the relative rank $r_{N(s)}$ is given by the number of arrivals in A_1 and A_2 (see Figure 1) , and

$$\mathrm{E}[R_{\sigma_t}^{(t)} \mid T_{\sigma_t} = s, X_{\sigma_t} = x] =$$

$$\begin{cases} 1 + x(t-s) & \text{if } 0 \le x \le g_t(0) \\ 1 + x(t-s) + \displaystyle\int_0^{g_t^{-1}(x)} \frac{x - g_t(u)}{1 - g_t(u)} du & \text{if } g_t(0) \le x \le g_t(s). \end{cases} \tag{2.17}$$

where the second part of (2.17) holds because we know that if we haven't stopped before s then there can have been no arrivals under the curve before s so that, conditionally to $T_{\sigma_t} = s$, the value of any arrival occurring at time $0 \le u \le g_t^{-1}(x)$ is uniformly distributed on $[g_t(u), 1]$.

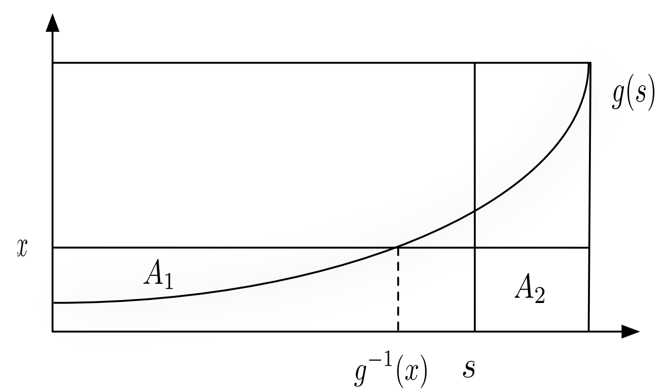

Figure 2: Smaller arrivals can only occur in A_1 and A_2.

Now, conditionally to $T_{\sigma_t} = s$, we know that the arrivals are distributed uniformly on $[0, g_t(s)]$. Therefore, integrating (2.17) yields

$$\begin{aligned} \mathrm{E}[R^{(t)}_{\sigma_t} \mid T_{\sigma_t} = s] &= 1 + \tfrac{1}{g_t(s)} \int_0^{g_t(s)} x(t-s)dx \\ &\quad + \tfrac{1}{g_t(s)} \int_{g_t(0)}^{g_t(s)} \int_0^{g_t^{-1}(x)} \frac{x - g_t(u)}{1 - g_t(u)} du dx. \end{aligned} \tag{2.18}$$

Using

$$W(\sigma_t) = \int_0^t \mathrm{E}[R^{(t)}_{\sigma_t} | T_{\sigma_t} = s] f_{T_{\sigma_t}}(s) ds + \Pi(t) \mathrm{P}[T_{\sigma_t} \geq t],$$

straightforward rearrangement and integration of (2.18) yields (2.15). ■

Finally take $g_n(.)$ and σ_n as above. If there have been no satisfactory arrivals for σ_n before time n, then the loss of the decision maker is given by $\tilde{R}_{\sigma_n} = \Pi(n)$ and thus

$$W(\sigma_n) \geq \Pi(n) \mathrm{e}^{-\mu_n(n)}, \tag{2.19}$$

where $\mu_n(s)$ is, as before, defined for all $0 \leq s \leq n$ by $\mu_n(s) = \int_0^s g_n(u)du$. Since we are interested in optimal values, and since the penalty function is chosen to be asymptotically linear in the horizon t, we see from (2.19) that we can restrict our attention without loss of generality to sequences of threshold functions which satisfy

$$\lim_{n \to \infty} n \mathrm{e}^{-\mu_n(n)} = 0. \tag{2.20}$$

Now let $(g_n(.))_{n \geq 1}$ be a sequence of strict monotone increasing threshold functions on $[0, n]$ (i.e. for each n, the function $g_n(.)$ is a strictly increasing threshold function with horizon n) and let $\sigma_n \in \mathcal{M}_n$ be the corresponding sequence of memoryless threshold strategies. Clearly, for all n, we have

$$w(n) \leq W(n) \leq W(\sigma_n).$$

We shall show that, under general conditions on the threshold sequence, we can use equation (2.15) to obtain $\lim_{n \to \infty} W(\sigma_n)$.

We first define the functions

$$h_n(u) = n g_n(nu)$$

for $u \in [0,1]$. A change of variables in (2.15) yields

$$W(\sigma_n) = 1 + (\Pi(n) - 1)e^{-\mu_t(n)} + \frac{1}{2}\int_0^1 d_n^1(s)ds + \frac{1}{2}\int_0^1 \int_0^s d_n^2(s,u)duds \tag{2.21}$$

where

$$\begin{cases} d_n^1(s) = h_n(s)^2(1-s)e^{-\int_0^s h_n(v)dv} \\ \\ d_n^2(s,u) = \dfrac{(h_n(s) - h_n(u))^2}{1 - h_n(u)/n}\, e^{-\int_0^s h_n(v)dv} \end{cases}$$

Now suppose that the sequence $g_n(.)$ satisfies (2.20) and that, for all $u \in (0,1)$, the sequence $h_n(u)$ converges. We can define the limit function

$$g(u) = \lim_{n\to\infty} h_n(u) = \lim_{n\to\infty} ng_n(nu).$$

Note that this function is unbounded in $u = 1$. In order to interchange the limit and the integration appearing in (2.21), we need some stronger assumptions on the sequence of thresholds $g_n(.)$. We will impose two conditions.

(C1) For every $s \in (0,1)$, $h_n(s)$ increases monotonically as it approaches $g(s)$.

(C2) The sequence of functions $h_n(s)$ is uniformly convergent on every interval $[0,a]$, for $a < 1$.

With these assumptions, a version of the dominated convergence theorem applies to $d_n^1(.)$ and $d_n^2(.,.)$ so that

$$\lim_{n\to\infty}\int_0^1 d_n^1(s)ds = \int_0^1 \lim_{n\to\infty} d_n^1(s)ds = \int_0^1 d_1(s)ds,$$

with

$$d_1(s) = g(s)^2(1-s)e^{-\int_0^s g(u)du},$$

and also

$$\lim_{n\to\infty}\int_0^1\int_0^s d_n^2(s,u)duds = \int_0^1\int_0^s d_2(s,u)duds,$$

with

$$d_2(s,u) = (g(s) - g(u))^2\, e^{-\int_0^s g(u)du}.$$

Hence, taking the limit for n going to infinity in (2.15), we get

$$\lim_{n\to\infty} W(\sigma_n) = 1 + \frac{1}{2}\int_0^1 d_1(s)ds + \frac{1}{2}\int_0^1\int_0^s d_2(s,u)duds =: L(g).$$

Likewise, if for each $n \geq 1$ we define the threshold sequence $a_i = g_n(i)$, $i = 1, \ldots, n$ and let $\tau_n \in M_n$ be the discrete stopping rule defined through these thresholds, then (1.5) applies and gives $V(\tau_n)$ as the Lebesgue integral of suitably chosen step functions. Under the same conditions on the sequence $g_n(.)$ as above, we see that we can take the limit for n going to infinity of $V(\tau_n)$ and this also yields $L(g)$.

This explains the following proposition.

Proposition 2.3 *Let $(g_n(.))_{n\geq 1}$ be a sequence of threshold functions satisfying (2.20), such that $\lim_{n\to\infty} n g_n(u)$ exists and is finite for all $u \in (0,1)$. Define the function*

$$g(u) = \lim_{n\to\infty} n g_n(u).$$

Let $\sigma_n \in \mathcal{M}_n$ be the sequence of memoryless threshold rules (for the Poisson embedded problem with horizon n) defined, for each n, by $g_n(.)$ and let $\tau_n \in M_n$ be the sequence of memoryless threshold rules (for the discrete n-arrival Robbins' problem) defined, for each n, by the threshold sequence $(g_n(i))_{i=1,\ldots,n}$. Then, under assumptions C1 and C2,

$$\lim_{n\to\infty} W(\sigma_n) = \lim_{n\to\infty} V(\tau_n) = L(g) \tag{2.22}$$

where

$$\begin{aligned} L(g) &= 1 + \tfrac{1}{2}\int_0^1 g(u)^2(1-u)e^{-\int_0^u g(x)dx}du \\ &\quad + \tfrac{1}{2}\int_0^1 \int_0^u (g(u)-g(v))^2 dv\, e^{-\int_0^u g(x)dx}du. \end{aligned} \tag{2.23}$$

Remark 5 *This is the same integral expression as that obtained by Assaf and Samuel-Cahn (1996), see Section 1.4.*

Example 1 *Let $g_n(s) = \dfrac{c}{n-s+c}$, with $c > 1$. This sequence satisfies the conditions imposed above with $g(u) = \dfrac{c}{1-u}$. Applying (2.22) yields*

$$L(c) = 1 + \frac{c}{2} + \frac{1}{c^2-1}.$$

Therefore for all $c > 1$ and all $t \in [0,\infty)$ we get $w(t) \leq W(t) \leq L(c)$. This expression is minimal for $c = 1.94697$ and yields the upper bound

$$w(t) \leq 2.33183.$$

This upper bound has already been obtained in Bruss and Ferguson (1993) and Assaf and Samuel Cahn (1996) for the discrete n arrival problem.

2.3 Properties of the value functions

The bounds obtained for $W(t)$ can obviously be extended to $w(t)$. This explains the following result, which we can now state without proof.

Proposition 2.4 *The value functions are bounded on* $]0, \infty]$, *and satisfy*

$$1 \leq w(t) \leq W(t) \leq 2.33182$$

for all t sufficiently large.

Lemma 2.5 *For all t sufficiently large and all $\delta > 0$,*

$$w(t+\delta) - w(t) \geq -3\delta(\delta+1). \tag{2.24}$$

Proof: Consider the Poisson embedded problem with horizon $t+\delta$. Recall the notation $\tilde{R}_k^{(t)}$ from equation (2.1). By conditioning on the number $N(\delta)$ of arrivals in $(0, \delta)$ we get

$$w(t+\delta) \geq e^{-\delta} \inf_\tau E(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 0) + \delta e^{-\delta} \inf_\tau E(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 1), \tag{2.25}$$

where we have neglected the case $N(\delta) \geq 2$.

We first consider the first term appearing in the rhs of (2.25). From the homogeneity assumptions on the arrival process, one sees that solving the Poisson embedded Robbins' Problem on $[0, t+\delta]$ with no arrivals before time δ is equivalent to solving the same problem on $[0, t]$ with penalty $\Pi(t+\delta)$. Since $\Pi(\cdot)$ is increasing this implies

$$\inf_\tau E\left(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 0\right) \geq w(t). \tag{2.26}$$

Next consider the second term of (2.25). By conditioning on the value X of the (only) arrival in $(0, \delta)$, we get

$$\begin{aligned} \inf_\tau E\left(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 1\right) &= \inf_\tau \int_0^1 E\left(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 1, X = x\right) dx \\ &\geq \int_0^1 \inf_\tau E\left(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 1, X = x\right) dx. \end{aligned} \tag{2.27}$$

From the optimality principle we know that an optimal action given $\{X = x\}$ is to select this arrival if and only if its expected rank is smaller than the optimal value obtainable by refusing it. Selecting x yields an expected loss of $1 + xt$, and refusing it an expected loss given by

$$E_t(x, \delta) := \inf_{\tau, T_\tau > \delta} \left\{ E\left(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 1, X = x \right) \right\},$$

where the infimum is taken over all strategies for which $T_\tau > \delta$ almost surely. Hence

$$\inf_\tau E\left(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 1, X = x \right) = \min\left\{1 + xt, E_t(x, \delta)\right\}. \tag{2.28}$$

As above, the homogeneity of the arrival process and the hypothesis on $\Pi(\cdot)$ guarantee that $E_t(x, \delta) \geq w(t)$ so that, from (2.27) and (2.28),

$$\inf_\tau E\left(\tilde{R}_\tau^{(t+\delta)} \mid N(\delta) = 1 \right) \geq \int_0^1 \min\left\{1 + xt, w(t)\right\} dx. \tag{2.29}$$

Combining (2.25), (2.26) and (2.29) then yields

$$w(t + \delta) \geq e^{-\delta} w(t) + \delta e^{-\delta} \int_0^1 \min\left\{1 + xt, w(t)\right\} dx. \tag{2.30}$$

Now choose t sufficiently large to ensure that $w(t) \geq 1$. Then there exists $x_0 \in [0, 1)$ for which $1 + x_0 t = w(t)$ and thus

$$\int_0^1 \min\left\{1 + xt, w(t)\right\} dx = \int_0^{x_0} (1 + xt) dx + w(t)(1 - x_0).$$

From (2.30) this then yields

$$w(t + \delta) \geq e^{-\delta} w(t) + \delta e^{-\delta} \left(w(t) - \frac{(w(t) - 1)^2}{2t} \right).$$

Now use $e^{-\delta} \geq 1 - \delta$, to obtain

$$w(t + \delta) \geq w(t) - \delta^2 w(t) - \delta e^{-\delta} \frac{(w(t) - 1)^2}{2t}.$$

Since $w(t) \leq 3$ for t sufficiently large, this implies

$$w(t + \delta) \geq w(t) - 3\delta^2 - 2\delta \frac{e^{-\delta}}{t},$$

and equation (2.24) follows. ■

Lemma 2.6 *There exists a constant $L > 0$ such that, for all t and δ positive,*

$$w(t+\delta) - w(t) \leq L\delta. \tag{2.31}$$

Proof: Let $\mathcal{K}_t$ be the subset of $\mathcal{T}$ consisting of all strategies which disregard any event occurring in $(t, t+\delta)$. Clearly $w(t+\delta) = \inf_{\mathcal{T}} E\left(\tilde{R}_\tau^{(t+\delta)}\right) \leq \inf_{\mathcal{K}_t} E\left(\tilde{R}_\tau^{(t+\delta)}\right)$. Now take $\tau \in \mathcal{K}_t$. Then $\mathbf{1}_{\{T_\tau \leq t+\delta\}} = \mathbf{1}_{\{T_\tau \leq t\}}$ almost surely. Since the rank of the selected arrival (evaluated with respect to the number of observations in $(0, t+\delta)$) cannot increase from t to $t+\delta$ by more than the number of arrivals in $(t, t+\delta)$, this yields

$$E\left(R_\tau^{(t+\delta)}\mathbf{1}_{\{T_\tau \leq t+\delta\}}\right) \leq E\left(R_\tau^{(t)}\mathbf{1}_{\{T_\tau \leq t\}}\right) + \delta.$$

This inequality holds for all $\tau \in \mathcal{K}_t$ and thus

$$w(t+\delta) \leq \inf_{\mathcal{K}_t}\left\{E\left(R_\tau^{(t)}\mathbf{1}_{\{T_\tau < t\}}\right) + \Pi(t+\delta)P(T_\tau > t)\right\} + \delta.$$

Adding and subtracting $\inf_{\mathcal{K}_t}\{(\Pi(t) - \Pi(t+\delta))P(T_\tau > t)\}$ to the rhs of this last equation, and using the fact that the sum of infima is smaller than the infimum of a sum, we get

$$\begin{aligned} w(t+\delta) &\leq \inf\nolimits_{\mathcal{K}_t}\left\{E\left(R_\tau^{(t)}\mathbf{1}_{\{T_\tau < t\}}\right) + \Pi(t)P(T_\tau > t)\right\} \\ &\qquad + (\Pi(t+\delta) - \Pi(t))\sup\nolimits_{\mathcal{K}_t} P(T_\tau > t) + \delta \\ &= \inf\nolimits_{\mathcal{K}_t} E\left(\tilde{R}_\tau^{(t)}\right) + (\Pi(t+\delta) - \Pi(t)) + \delta. \end{aligned}$$

Since, by definition, $\inf_{\mathcal{K}_t} E\left(\tilde{R}_\tau^{(t)}\right) = w(t)$ we get

$$w(t+\delta) \leq w(t) + (\Pi(t+\delta) - \Pi(t)) + \delta.$$

The hypothesis on $\Pi(\cdot)$ give (2.31). ■

Lemmas 2.5 and 2.6 immediately yield the following.

Theorem 2.7 *The value function $w(t)$ is continuous on $\mathbb{R}$ and Lipschitz continuous on (t_0, ∞), for some t_0 sufficiently large.*

If we restrict our attention to $\mathcal{M}_t$, i.e. the set of memoryless threshold rules on $[0, t]$, we see that the proof of Proposition 2.7 holds for $W(t)$ with only minor changes. Hence we obtain

Proposition 2.8 *The value function $W(t)$ restricted to the class of memoryless threshold rules is uniformly continuous on $[0, \infty)$.*

We now prove the existence of optimal strategies for the Poisson embedded Problem. This is intuitively clear since we have shown that the value functions $w(t)$ and $W(t)$ are well defined and bounded, so that we should be able to compare the expected rank of each arrival to the best obtainable value and thus decide at each arrival whether or not it is optimal to stop. The point is that this comparison is possible at any arrival time, and so leads to an almost surely unique optimal strategy.

Proposition 2.9 *For each t there exists a stopping rule $\tau_t^\star$ in $\mathcal{T}$ such that*

$$w(t) = \mathrm{E}\left[\tilde{R}^{(t)}_{\tau_t^\star}\right], \tag{2.32}$$

and a stopping rule $\sigma_t^\star \in \mathcal{M}_t$ such that

$$W(t) = \mathrm{E}\left[\tilde{R}^{(t)}_{\sigma_t^\star}\right]. \tag{2.33}$$

Proof: Fix $t > 0$, and suppose that there is an arrival of value X_i at time T_i, $0 < T_i < t$, $i \geq 1$. Let $E(i, t)$ be the expected loss incurred by refusing this arrival and continuing optimally thereafter, i.e. $E(i, t)$ is the minimal expected rank obtainable under the history $\mathcal{F}_i$ by using strategies which stop almost surely after the ith arrival. It is given by

$$E(i, t) = \inf_{\tau \in \mathcal{T}, \tau > i} \mathrm{E}[\tilde{R}^{(t)}_\tau \mid \mathcal{F}_i],$$

where the infimum is taken over the set of all stopping rules $\tau \in \mathcal{T}$ such that $\mathrm{P}[\tau > i] = 1$. For all $i \geq 0$ and every history $\mathcal{F}_i$, we see that $E(i, t)$ is well defined for all horizons $t > 0$. Using arguments similar to those appearing in the proof of Proposition 2.7 we see that it satisfies the upper bound

$$E(i, t) \leq r_i + w(t - T_i) + \Pi(t) - \Pi(T_i).$$

From the optimality principle, we know that it can only be optimal to stop on an arrival X_i if the expected loss incurred by selecting X_i is smaller than the expected loss incurred by refusing X_i. Hence, if we define the rule $\tau_t^\star$ by

$$\begin{cases} \tau_t^\star = i & \text{if} \quad \mathrm{E}[\tilde{R}^{(t)}_i \mid \mathcal{F}_i] \leq E(i, t) \\ \tau_t^\star > i & \text{if} \quad \mathrm{E}[\tilde{R}^{(t)}_i \mid \mathcal{F}_i] > E(i, t), \end{cases} \tag{2.34}$$

then $\tau_t^\star$ belongs to $\mathcal{T}$ and must be optimal for each time t.

Minor adaptations of these arguments show that the same result holds for the restricted problem. ∎

2.4 Variations on the embedding

The painstaking care with which we have set up the above Poisson embedding has a number of advantages, first and foremost of which is the fact that we can now define a number of variations on this embedding which profit from the framework we have just established. Among these variations, the following stand out.

First, instead of considering a fixed penalty $\Pi(t)$, one could decide that if no decision has been reached before time t, the decision maker is forced to choose the first arrival thereafter. His loss is therefore defined by

$$R_\tau = \sum_{i=1}^{N(t)} 1_{\{X_i \leq X_\tau\}}$$

where it is now implicit that $X_\tau \sim U[0,1]$ if $T_\tau \geq t$. We will refer to this version of the problem as *Variation A*.

Another option (which we will refer to as *Variation B*) consists in randomizing the number of arrivals without embedding the problem within a Poisson process. In other words, we could consider a situation where a *prophet* draws a number N from a given known discrete distribution, and then draws N independent uniform random samples. The prophet then presents these values to the decision maker without telling him the value of N. If no decision has been reached before the Nth arrival is presented, the prophet warns the decision maker upon visualizing the Nth arrival that he is obliged to choose it. The objective if therefore to minimize the loss

$$R_\tau = \sum_{i=1}^{N} 1_{\{X_i \leq X_\tau\}},$$

where it is now implicit that $X_\tau \sim U[0,1]$ if $\tau = N$. Letting the mean of N vary towards infinity, we see that we are in a situation similar to the above.

Going over the proofs of the different results of this chapter, one sees that most carry through basically without change for the above two variations on the problem. Moreover, some properties are nicer in these problems. Indeed, consider the second variation, and let $\tilde{v}(n)$ denote the minimal expected loss when $E[N] = n$. Then the following holds.

Lemma 2.10 *The function $\tilde{v}(n)$ is increasing.*

Proof: Clearly we can write

$$v(n+1) \geq \inf E[R_\tau \mid N \leq n] P[N \leq n] + \inf E[R_\tau \mid N \geq n] P[N \geq n]$$

Now obviously we have $\inf E[R_\tau \mid N \leq n] = v(n)$, since otherwise $v(n)$ would not be the optimal value. Also, we have $\inf E[R_\tau \mid N \geq n] \geq v(n)$. Indeed, suppose the existence of a half-prophet who will tell the decision maker the values of the $N - n + 1$ worst arrivals. In this case, the decision maker can ignore all these values and concentrate on acting optimally with n arrivals. This he can achieve with a value at best of $v(n)$. Whence the claim. ∎

Chapter 3

Comparison of the classical and the embedded problems

WHERE WE PROVE THAT the variations presented in the previous chapter are, at least asymptotically, equivalent to the original problem.

3.1 Equivalence of the Poisson embedded and the classical problems

We now prove that the asymptotic value for the Poisson embedded problem exists, and is equal to that of the

Proposition 3.1 *For all $\epsilon > 0$ there exists $t^\star > 0$ such that for all $t \geq t^\star$,*

$$w(t) > v - \epsilon.$$

Proof: Fix $\epsilon > 0$ and consider the Poisson embedded problem with horizon t. Suppose that the decision maker (say Q) is told in advance the number of arrivals which will occur in $[0, t]$. Let $w_Q(t)$ be the corresponding expected optimal value. Since Q is facing our problem with more information, he can only do better than us so that

$$w_Q(t) \leq w(t). \tag{3.1}$$

Conditioning on the number of arrivals in $[0,t]$ yields

$$\begin{aligned} w_Q(t) &= \inf_{\sigma\in\tilde{\mathcal{T}}} \sum_{k=0}^{\infty} P(N(t)=k) E(\tilde{R}_\sigma^{(t)} \mid N(t)=k) \\ &\geq \sum_{k=0}^{\infty} P(N(t)=k) \inf_{\sigma\in\tilde{\mathcal{T}}} E(\tilde{R}_\sigma^{(t)} \mid N(t)=k). \end{aligned}$$

Now consider the minimal expected rank obtainable by Q conditionally to $\{N(t)=k\}$. On the one hand, if $k \leq \Pi(t)$, the best Q can do is apply $\tau_k^\star$, the strategy that is optimal for exactly k arrivals; hence, for all $k \leq \Pi(t)$,

$$\inf_{\sigma\in\tilde{\mathcal{T}}} E(\tilde{R}_\sigma^{(t)} \mid N(t)=k) = v(k). \tag{3.2}$$

On the other hand, if $k > \Pi(t)$, this equality does not hold since Q is solving Robbins' Problem for k arrivals with the knowledge that he can always obtain at the worst a penalty of $\Pi(t)$, i.e. he is in a better position than a player in the discrete setting with k arrivals. However we have

$$\inf_{\sigma\in\tilde{\mathcal{T}}} E(\tilde{R}_\sigma^{(t)} \mid N(t)=k) \geq v(\lfloor \Pi(t) \rfloor), \tag{3.3}$$

for all $k > \Pi(t)$. To see this, let $v_Q(k) = \inf_{\sigma\in\tilde{\mathcal{T}}} E(\tilde{R}_\sigma^{(t)} \mid N(t)=k)$. The same half-prophet argument as that used by Bruss and Ferguson (1993) to prove the monotonicity of $v(n)$ applies in this setting, and shows that $v_Q(k)$ must be an increasing function of k. Hence, for all $k > \Pi(t)$, $v_Q(k) \geq v_Q(\lfloor \Pi(t) \rfloor)$, and thus (3.3) holds. Combining (3.1), (3.2) and (3.3) we obtain

$$w(t) \geq \sum_{k=0}^{\lfloor \Pi(t) \rfloor} P(N(t)=k)v(k) + \sum_{k=\lfloor \Pi(t) \rfloor+1}^{\infty} P(N(t)=k)v(\lfloor \Pi(t) \rfloor). \tag{3.4}$$

We know that $v(k)$ increases to v. Hence there exists $m_0 = m_0(\epsilon) \in \mathbb{N}$ such that $v(m) > v - \epsilon$ for all $m \geq m_0$. The monotonicity of $\Pi(\cdot)$ implies that there exists $t_0 = t_0(\epsilon)$ such that $\Pi(t) > m_0$ for all $t \geq t_0$. Therefore, from (3.4),

$$w(t) \geq (v-\epsilon) \sum_{k=m_0}^{\infty} P(N(t)=k) = (v-\epsilon) P(N(t) \geq m_0)$$

for all $t \geq t_0$.

Since $P(N(t) \geq m_0) \to 1$ as $t \to \infty$ for all m_0, there exists t_1 such that for all $t \geq t_1$, $P(N(t) \geq m_0) \geq 1 - \epsilon$. Therefore, for all $t \geq \max\{t_0, t_1\}$,

$$w(t) \geq (v - \epsilon)(1 - \epsilon) = v + \epsilon^2 - (\epsilon + v\epsilon),$$

and thus, since $v \leq 3$,

$$w(t) \geq v - 4\epsilon.$$

■

Corollary 3.2 *If the limit $w = \lim_{t\to\infty} w(t)$ exists, then it satisfies $w \geq v$.*

The arguments in this proof also hold if we restrict our attention to the set of memoryless strategies $\mathcal{M}_t$. Hence Proposition 3.1 holds for the asymptotic memoryless values W and V, and we obtain the following corollary.

Corollary 3.3 *Let W be the minimal expected rank obtainable through memoryless thresholds in the Poisson embedded Robbins' Problem, and let V be its discrete counterpart. Then*

$$V \leq W.$$

To obtain an inequality in the other direction, we first need a preparatory lemma on the tail probabilities for Poisson processes.

Lemma 3.4 *Let $N(n)$ be the number of arrivals of a Poisson process of rate 1 on $[0, n] \times [0, 1]$, and let $\frac{1}{2} < \alpha < 1$. Then*

$$\lim_{n\to\infty} \big(nP(N(n) < n - n^\alpha)\big) = 0. \tag{3.5}$$

Proof: $N(n)$ is a Poisson random variable of mean and variance n so that, by the central limit theorem, $(N(n) - n)/\sqrt{n}$ converges in law to a standard normal distribution $\mathcal{N}(0, 1)$. Now choose some α between $\frac{1}{2}$ and $\frac{2}{3}$, in order to ensure that $n^{\alpha-\frac{1}{2}}$ increases to ∞ as $n \to \infty$ and that $(n^{\alpha-\frac{1}{2}})^3/\sqrt{n} = n^{3\alpha-2}$ decreases to 0 as $n \to \infty$. Then we can apply a theorem on normal approximation (see Feller (1968), p.193) to get

$$P\left(\frac{N(n) - n}{\sqrt{n}} > n^{\alpha-\frac{1}{2}}\right) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{n^{\alpha-\frac{1}{2}}} e^{-\frac{1}{2}(n^{\alpha-\frac{1}{2}})^2}. \tag{3.6}$$

For n sufficiently large we can then use the approximate symmetry of the distribution of $(N(n) - n)/\sqrt{n}$ to obtain, from (3.6),

$$P\left(N(n) < n - n^{\alpha}\right) = P\left(\frac{N(n) - n}{\sqrt{n}} < -n^{\alpha - \frac{1}{2}}\right) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{n^{\alpha - \frac{1}{2}}} e^{-\frac{1}{2}(n^{\alpha - \frac{1}{2}})^2}.$$

Hence altogether

$$nP\left(N(n) < n - n^{\alpha}\right) \sim \frac{1}{\sqrt{2\pi}} n^{\frac{3}{2} - \alpha} e^{-\frac{1}{2}(n^{\alpha - \frac{1}{2}})^2} < \frac{1}{\sqrt{2\pi}} \frac{n}{e^{\frac{1}{2}n^{2\alpha - 1}}}$$

which tends to 0 as n tends to infinity. This establishes (3.5) for all $\alpha \in \left(\frac{1}{2}, \frac{2}{3}\right)$, and the extension to $\alpha \in (\frac{1}{2}, 1)$ is immediate. Hence the result. ∎

Proposition 3.5 *Let $\alpha \in (\frac{1}{2}, 1)$ and $t > 0$. Define $\beta_t = \lfloor t - t^{\alpha} \rfloor$ and let $\tau^{\star}_{\beta_t}$ be the optimal strategy for the discrete problem with β_t arrivals. Then, for all $\epsilon > 0$ and all t sufficiently large,*

$$w(t) \leq v(\beta_t) + E\left(X_{\tau^{\star}_{\beta_t}}\right)(t^{\alpha} + 1) + \epsilon. \tag{3.7}$$

Proof: Any strategy for the discrete case with n arrivals can be extended in a natural way to define a (sub-optimal) strategy for the continuous case, so that we can consider $\tau^{\star}_{\beta_t}$ as a strategy acting in continuous time on $[0, t]$. Let σ_{β_t} denote this strategy and let $\tilde{w}(\beta_t)$ be its corresponding value, i.e. $\tilde{w}(\beta_t)$ is the expected rank obtained by using a strategy which is optimal if and only if there are exactly β_t arrivals before the horizon t. Conditioning on the number of arrivals in $[0, t]$, we get

$$\tilde{w}(\beta_t) = \tilde{w}(\beta_t | N(t) < \beta_t) P(N(t) < \beta_t) + \tilde{w}(\beta_t | N(t) \geq \beta_t) P(N(t) \geq \beta_t), \tag{3.8}$$

where $\tilde{w}(\beta_t \mid E)$ denotes the expected loss under σ_{β_t} conditioned on the event E.

First suppose that $N(t) < \beta_t$. Since σ_{β_t} acts on β_t arrivals, there is a positive probability that no arrival is selected within the given time. Hence we must distinguish two cases. On the one hand, if $T_{\sigma_{\beta_t}} > t$, the player loses the penalty. On the other hand, if $T_{\sigma_{\beta_t}} \leq t$, his loss is given by some function $E(R^{(t)}_{\sigma_{\beta_t}} \mid N(t) < \beta_t) \leq v(\beta_t)$. This yields

$$\tilde{w}(\beta_t | N(t) < \beta_t) \leq v(\beta_t) P(T_{\sigma_{\beta_t}} \leq t) + \Pi(t) P(T_{\sigma_{\beta_t}} > t),$$

and thus, since $v(\cdot)$ is bounded,

$$\tilde{w}(\beta_t | N(t) < \beta_t) \leq \Pi(t) + K$$

for some positive constant K. This last inequality, combined with the assumptions on $\Pi(\cdot)$ (see (2.3)) and Lemma 3.4, proves that

$$\tilde{w}(\beta_t|N(t) < \beta_t)P(N(t) < \beta_t) < \frac{\epsilon}{2} \tag{3.9}$$

for t sufficiently large.

Next suppose that $N(t) \geq \beta_t$. Then, since the β_t-optimal strategy stops almost surely not later than the β_tth arrival,

$$\tilde{w}(\beta_t|N(t) \geq \beta_t) = v(\beta_t) + E(X_{\sigma_{\beta_t}}(N(t) - \beta_t)|N(t) \geq \beta_t). \tag{3.10}$$

Now, given $N(t) \geq \beta_t$, $X_{\sigma_{\beta_t}}$ is independent of $N(t)$ and of $X_{\beta_t+1}, X_{\beta_t+2}, \ldots$. Hence

$$E(X_{\sigma_{\beta_t}} \mid N(t) \geq \beta_t) = E(X_{\sigma_{\beta_t}}) = E(X_{\tau^\star_{\beta_t}}),$$

and thus, from (3.10),

$$\tilde{w}(\beta_t \mid N(t) \geq \beta_t) = v(\beta_t) + E(X_{\tau^\star_{\beta_t}})\big(E(N(t) \mid N(t) \geq \beta_t) - \beta_t\big). \tag{3.11}$$

Furthermore,

$$E(N(t)|N(t) \geq \beta_t) = \sum_{k=\beta_t}^{\infty} k \frac{P(N(t) = k)}{P(N(t) \geq \beta_t)} \leq \frac{E(N(t))}{1 - P(N(t) < \beta_t)}.$$

Now, since $P(N(t) < \beta_t) \to 0$ as $t \to \infty$, we know that $\frac{1}{1-P(N(t)<\beta_t)} < 1 + 2P(N(t) < \beta_t)$ for t sufficiently large. Therefore

$$E(N(t)|N(t) \geq \beta_t) \leq E(N(t))(1 + 2P(N(t) < \beta_t)) = t + 2tP(N(t) < \beta_t),$$

and, from Lemma 3.4,

$$E(N(t)|N(t) \geq \beta_t) \leq t + \frac{\epsilon}{2},$$

for all t sufficiently large. From (3.11) this yields

$$\tilde{w}(\beta_t|N(t) \geq \beta_t) \leq v(\beta_t) + E(X_{\tau^\star_{\beta_t}})(t - \beta_t + \frac{\epsilon}{2}),$$

and thus, since $t - \beta_t \leq t^\alpha + 1$,

$$\tilde{w}(\beta_t|N(t) \geq \beta_t) \leq v(\beta_t) + E(X_{\tau^\star_{\beta_t}})(t^\alpha + 1) + \frac{\epsilon}{2} \tag{3.12}$$

for t sufficiently large. Combining (3.8), (3.9) and (3.12), we obtain

$$\tilde{w}(\beta_t) \leq v(\beta_t) + E(X_{\tau^\star_{\beta_t}})(t^\alpha + 1) + \epsilon,$$

and thus, since $w(t) \leq \tilde{w}(\beta_t)$,

$$w(t) \leq v(\beta_t) + E(X_{\tau^\star_{\beta_t}})(t^\alpha + 1) + \epsilon$$

for all t sufficiently large. ■

From Corollary 3.2 and Proposition 3.5 we see that, in order to prove both the existence of w and its equality with v, we need

$$\lim_{t\to\infty} t^{\alpha} E\left(X_{\tau^{\star}_{\beta_t}}\right) = 0, \tag{3.13}$$

for some $\alpha > \frac{1}{2}$. This obviously follows from Theorem 1.5. We have therefore proved the following

Corollary 3.6 *If the limit $w = \lim_{t\to\infty} w(t)$ exists then it satisfies $w \leq v$.*

Corollaries 3.2 and 3.6 immediately yield

Theorem 3.7 *The limiting value for the Poisson embedded Robbins' Problem exists and satisfies*

$$w = \lim_{t\to\infty} w(t) = v.$$

3.2 Equivalence of the variations on the problem

We conclude this chapter by noting that the above equivalence can be transposed without further ado to most reasonable embeddings of the problem, such as those introduced at the end of Chapter 2. More precisely, one sees that our proof of the equivalence can be carried through to Variation A and, more generally, we can prove the following.

Theorem 3.8 *Let P_n be a family of discrete probability distributions on the positive integers such that if $X \sim P_n$, then $E[X] = n$. Suppose that the family P_n has sufficiently thin tails for Lemma 3.4 to hold. Finally consider Variation B with $N \sim P_n$. Then*

$$\lim_{n\to\infty} \tilde{v}(n) = v.$$

The proof runs along the same lines as that for the Poisson embedded problem, the only difference being that Proposition 3.1 is now significantly easier to establish since there is no need to differentiate with respect to the value of the penalty function. Again the key step is the inequality between expected values and expected ranks which we established in Chapter 1.

Chapter 4

Randomizing Stopping Problems

WHERE WE EXPLORE the consequences of the results given in the previous chapters.

4.1 Calculus of variations and optimal memoryless rules

Considering some of the equations given throughout this manuscript, one sees that some questions could be quite elegantly answered by methods of calculus of variations. The first section of this final chapter aims to give an account of what can be achieved through this tool.

4.1.1 An astonishingly simple solution to Moser's problem

Recall the statement of Moser's problem from the Introduction to this mémoire. Now consider a Poisson embedding of the form given by Variation A. Using the same notations and setup as before, we can therefore consider a Poisson version of Moser's problem defined by the value function

$$M(t) = \inf_{\tau} E(X_\tau),$$

where the infimum is to be chosen among all canonical stopping times for this problem.

The arrivals being independent of each other, it is clear that an optimal decision at time T_k depends only on the value of the arrival and on the length of the time interval $[T_k, t]$. Hence the optimal rule for this problem $\hat{\tau}_t$, say, is of the form

$$\tau = \inf\{s > 0 : X_s \leq g^\star(s)\},$$

where $g^\star(\cdot)$ is some threshold function (see Section 1.2). Direct integration along the same lines as in the proof of Proposition 2.4 then yields

$$M(t) = \frac{1}{2} \inf_g \left\{ \int_0^t g^2(u) \exp\left(- \int_0^u g(v)dv \right) du + \exp\left(- \int_0^t g(v)dv \right) \right\}.$$

Minimizing this last equation can, to the best of our knowledge, only be performed numerically. The second term in the rhs is, nevertheless, of lesser importance than the first since we expect that, for the strategy to be good, it selects an arrival with a probability which tends rapidly to 1. It therefore makes sense to consider the suboptimal stopping rule which minimizes

$$L(u, G(u), g(u)) := \int_0^t g^2(u) \exp\left(-G(u)\right) du$$

with $G(u) = \int_0^u g(v)dv$. Setting $L(u, x, v) = e^{-x}v^2$, we see that the problem has become a standard minimization problem in calculus of variations. The solution $g^\star$ must therefore be a solution to the Euler-Lagrange equations

$$\partial_x L(u, G(u), g(u) = d_t \partial_v L(u, G(u), g(u)) \tag{4.1}$$

which satisfies the border condition $g(t) = 1$. Solving (4.1) is, in this case, easy. It yields

$$g^\star(u) = \frac{2}{t - s + 2},$$

which is exactly the form of the asymptotic optimal threshold sequence for the discrete version of Moser's problem.

In particular this proves that

$$M(t) \leq 2\frac{t+1}{(2+t)^2},$$

and all known results about Moser's problem are easily re-obtained.

4.1.2. How close are Moser's and Robbins' problem?

Recall equation (2.15). The same arguments as in the previous section allow for solving, in principle, Robbins' problem when the quest for optimal rules is restricted to the set of memoryless thresholds. Indeed, minimizing the rank through memoryless threshold rules is equivalent to finding a function g which satisfies the constraints on threshold functions and minimizes the functional

$$I(g) = \int_0^t \left[(t-u)g^2(u) + \int_0^u \frac{(g(u)-g(v))^2}{1-g(v)} dv \right] \exp\left(- \int_0^u g(v)dv \right) du.$$

Hence the optimal g is, in principle, given by the tools of calculus of variations. The corresponding equations are, however, much less tractable than for Moser's problem, and we know of no explicit solution. Rewriting $I(g)$, we get

$$\begin{aligned} I(g) = \quad & t\int_0^t g^2(u) \exp\left(- \int_0^u g(v)dv \right) du \\ & + \int_0^t \left(\int_0^u \frac{(g(v)-g(u))^2}{1-g(v)} - g(u)^2 dv \right) \exp\left(- \int_0^u g(v)dv \right) du, \end{aligned}$$

which proves that, in the context of our embedded memoryless problems, the ranks satisfy

$$E(R_\tau) = 1 + \frac{1}{2}\left(tE(X_\tau) + \mathcal{I}(\tau) \right),$$

with $\mathcal{I}(\tau) = \int_0^t \left(\int_0^u \frac{(g(v)-g(u))^2}{1-g(v)} - g(u)^2 dv \right) \exp\left(- \int_0^u g(v)dv \right) du.$

It is easy to show that

$$0 \leq \lim_{t\to\infty} \inf_g \mathcal{I}(\tau) \leq \frac{2}{3}.$$

Better estimates have, so far, eluded our sagacity.

4.2 A differential equation on the value function

In this final section we prove that $w(\cdot)$ is a differentiable function which satisfies the relationship

$$w'(t) + w(t) = \int_0^1 \min\{1+xt, w(t)+h(t,x)\}dx + \chi(t), \tag{4.2}$$

where $\chi(t)$ tends to zero as t tends to infinity, and $h(t,x)$ is a continuous function depending on the value of an arrival selected before time t. Although this is not a differential equation in the usual sense, it is a capsule that contains an infinite dimensional problem in a closed form. Moreover, our results from the previous section show that a solution to this equation is tantamount to a solution to Robbins' Problem. Of course the presence of two unknown functions in (4.2) does not allow for obtaining explicit solutions. This equation, however, does open the way for a numerical study of the behavior of $w(t)$ in terms of the function $h(t,x)$.

Before proceeding to the proof of (4.2), we need, for all t, the existence of a strategy $\tau_t^\star$ such that

$$w(t) = w_{\tau_t^\star}(t),$$

i.e. we need the existence – for every horizon – of an optimal strategy. This follows from the optimality principle and the continuity of $w(t)$.

To see this, fix $t \in \mathbb{R}^+$ and suppose that there is an arrival of value X_i at time $0 \le T_i \le t$ for some $i \ge 1$. Let r_i be the relative rank of X_i and $\mathcal{F}_i$ be shorthand for the history of the process up to time T_i. Then, from the optimality principle, we know that it is optimal to select (T_i, X_i) if and only if the expected loss incurred for selecting (T_i, X_i), is smaller than the expected loss incurred for refusing it. The former is given by

$$E(\tilde{R}_i^{(t)} \mid \mathcal{F}_i) = r_i + X_i(t - T_i),$$

and the latter by

$$\inf_{\tau\in\mathcal{T},\tau>i} E(\tilde{R}_\tau^{(t)} \mid \mathcal{F}_i) =: E(i,t),$$

where the infimum is taken over the set of all stopping rules $\tau \in \mathcal{T}$ such that $P(T_\tau > T_i) = 1$. The function $E(i,t)$ is well defined for all $i \ge 1$, every history $\mathcal{F}_i$ and all horizons t. Also, for all i and fixed history $\mathcal{F}_i$, the same arguments as for Theorem 2.7 prove that $E(i,t)$ is continuous in t. Hence the stopping rule $\tau_t^\star$ given by

$$\begin{cases} \tau_t^\star = i & \text{if} \quad r_i + X_i(t - T_i) \le E(i,t) \\ \tau_t^\star > i & \text{if} \quad r_i + X_i(t - T_i) > E(i,t) \end{cases}$$

is well defined at all stages of the process and is optimal for each horizon t.

Proposition 4.1 *For all $t > 0$,*

$$P(T_{\tau_t^\star} \ge t) = \lim_{\delta\to 0} P(T_{\tau_{t-\delta}^\star} \ge t - \delta).$$

Proof: The probability of there being no arrivals in $(t-\delta,\ t)$ tends to one as $\delta \to 0$ independently of preceding arrivals. Hence, as $\delta \to 0$, a decision maker using either $\tau^\star_{t-\delta}$ or $\tau^\star_t$ will be a. s. confronted with the same set of arrivals before t. The continuity of $\Pi(t)$ implies that, as δ tends to zero, there exists a. s. a unique optimal limiting rule, and the statement follows. ■

Remark 6 *Note that, since $w(t)$ is bounded, $P(T_{\tau^\star_t} \geq t) \to 0$ as $t \to \infty$.*

Theorem 4.2 *Let $\tau^\star_t$ be the optimal strategy with respect to the horizon t. The value function $w(t)$ is differentiable and satisfies*

$$w'(t) + w(t) = \int_0^1 \min\{1 + xt, w(t \mid x)\}dx + \chi(t), \qquad (4.3)$$

where $\chi(t) = \Pi'(t)P(T_{\tau^\star_t} > t)$ and $w(t \mid x)$ is the optimal value conditioned on a first arrival at time 0 of value x which cannot be selected, i.e.

$$w(t \mid x) = \inf_{\tau \in \mathcal{T}} \left\{ E\left(R^{(t)}_\tau \mathbf{1}_{\{T_\tau \leq t\}} + \Pi(t)\mathbf{1}_{\{T_\tau > t\}} + \mathbf{1}_{\{X_\tau \geq x\}}\mathbf{1}_{\{T_\tau \leq t\}} \right) \right\}. \quad (4.4)$$

Proof: Fix $\delta > 0$. Conditioning on $N(\delta)$, the number of arrivals in $(0, \delta)$, we get

$$\begin{aligned} w(t) &= P(N(\delta)=0)\, E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta)=0) + P(N(\delta)=1)\, E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta)=1) \\ &\quad + P(N(\delta)\geq 2)\, E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta)\geq 2). \end{aligned}$$

Since $w(t)$ is bounded, so must be $E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) \geq 2)$ and thus, for δ sufficiently small,

$$w(t) = (1-\delta)\, E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta)=0) + \delta\, E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta)=1) + o(\delta). \qquad (4.5)$$

We now need to study the behavior of $E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta)=0)$ and $E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1)$ for $\delta \to 0$. For the sake of clarity, we will separate these results in two preparatory lemmas.

Lemma 4.3 *Set*

$$\chi(t,\delta) = E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta)=0) - w(t-\delta).$$

Then $\chi(t,\delta) \geq 0$ for all $t > 0$ and all $\delta > 0$, and

$$\lim_{\delta \to 0^+} \frac{\chi(t,\delta)}{\delta} = \Pi'(t)P(T_{\tau^\star_t} > t).$$

Proof: Set $\Delta\Pi(t,\delta) = \Pi(t) - \Pi(t-\delta)$. First consider the Poisson embedded Robbins' Problem with horizon $t-\delta$, and let $\tilde{\tau}$ be a strategy acting on $(0, t-\delta)$ as $\tau_t^\star$ would act on (δ, t) under the condition that $N(\delta) = 0$. We have

$$\begin{aligned} w_{\tilde{\tau}}(t-\delta) &= E(R^{(t)}_{\tau_t^\star} \mathbf{1}_{\{T_{\tau_t^\star} \le t\}} \mid N(\delta) = 0) + \Pi(t-\delta) P(T_{\tau_t^\star} > t \mid N(\delta) = 0) \\ &= E(\tilde{R}^{(t)}_{\tau_t^\star} \mid N(\delta) = 0) - \Delta\Pi(t,\delta) P(T_{\tau_t^\star} > t \mid N(\delta) = 0). \end{aligned}$$

Hence, since $w(t-\delta) \le w_{\tilde{\tau}}(t-\delta)$,

$$E(\tilde{R}^{(t)}_{\tau_t^\star} \mid N(\delta) = 0) - w(t-\delta) \ge \Delta\Pi(t,\delta) P(T_{\tau_t^\star} > t \mid N(\delta) = 0). \tag{4.6}$$

Next consider the Poisson embedded problem with respect to the horizon t, and let $\tilde{\sigma}$ be a strategy that ignores every arrival, if any, in $(0,\delta)$ and applies $\tau^\star_{t-\delta}$ on (δ, t). We have

$$\begin{aligned} E(\tilde{R}^{(t)}_{\tilde{\sigma}} \mid N(\delta) = 0) &= E(R^{(t-\delta)}_{\tau^\star_{t-\delta}} \mathbf{1}_{\{\tau^\star_{t-\delta} \le t-\delta\}}) + \Pi(t) P(T_{\tau^\star_{t-\delta}} > t-\delta) \\ &= w(t-\delta) + \Delta\Pi(t,\delta) P(T_{\tau^\star_{t-\delta}} > t-\delta). \end{aligned}$$

Now one can easily check, from the definitions, that

$$E(\tilde{R}^{(t)}_{\tau_t^\star} \mid N(\delta) = 0) \le E(\tilde{R}^{(t)}_{\tilde{\sigma}} \mid N(\delta) = 0) + o(\delta).$$

Hence

$$E(\tilde{R}^{(t)}_{\tau_t^\star} \mid N(\delta) = 0) - w(t-\delta) \le \Delta\Pi(t,\delta) P(T_{\tau^\star_{t-\delta}} > t-\delta) + o(\delta). \tag{4.7}$$

Combining (4.6) and (4.7), we then get

$$\Delta\Pi(t,\delta) P(T_{\tau_t^\star} > t \mid N(\delta) = 0) \le \chi(t,\delta) \le \Delta\Pi(t,\delta) P(T_{\tau^\star_{t-\delta}} > t-\delta) + o(\delta),$$

and thus, from Proposition 4.1,

$$\lim_{\delta \to 0^+} \frac{\chi(t,\delta)}{\delta} = \Pi'(t) P(T_{\tau_t^\star} > t).$$

■

Recall the definition of $w(t \mid x)$ from equation (4.4). For fixed $x \in [0, 1]$, the same arguments as those used to prove the continuity of $w(t)$ apply to $w(t \mid x)$. Hence, for each x, $w(t \mid x)$ is continuous in t. Also, for fixed t, one can check that $w(t \mid x)$ it is monotone decreasing in x on $[0, 1]$, with

$$w(t) + 1 \geq w(t \mid 0) \geq w(t \mid x) \geq w(t \mid 1) = w(t) \text{ for all } 0 \leq x \leq 1.$$

The following holds.

Lemma 4.4 *For all $t > 0$,*

$$\lim_{\delta \to 0^+} E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1) = \int_0^1 \min\{1 + xt, w(t \mid x)\}dx.$$

Proof: Fix $\delta > 0$ and let X denote the value of the (unique) arrival in $(0, \delta)$. Conditioning on X yields

$$E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1) = \int_0^1 E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1, X = x)dx, \tag{4.8}$$

where, by definition of $\tau^\star_t$,

$$\begin{aligned} &E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1, X = x) \\ &\qquad = \min\left\{1 + x(t - \delta), \inf\nolimits_{\tau \in \mathcal{T}, T_\tau > \delta} E\left(\tilde{R}^{(t)}_{\tau} \mid N(\delta) = 1, X = x\right)\right\}. \end{aligned}$$

One can check that, for all t,

$$w(t - \delta \mid x) \leq \inf_{\tau \in \mathcal{T}, T_\tau > \delta} E(\tilde{R}^{(t)}_{\tau} \mid N(\delta) = 1, X = x) \leq w(t - \delta \mid x) + \Delta\Pi(t, \delta).$$

Hence, from the continuity of $w(t \mid x)$,

$$\lim_{\delta \to 0} \left(\inf_{\tau \in \mathcal{T}, T_\tau > \delta} E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1, X = x)\right) = w(t \mid x),$$

and

$$\begin{aligned} \lim\nolimits_{\delta \to 0^+} E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1) &= \lim\nolimits_{\delta \to 0^+} \int_0^1 E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1, X = x)dx \\ &= \int_0^1 \lim_{\delta \to 0^+} E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1, X = x)dx \\ &= \int_0^1 \min\{1 + xt, w(t \mid x)\}dx. \end{aligned}$$

■

Proof of Theorem 4.2, continued

From Lemma 4.3 we know that $E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 0) = w(t-\delta) + \chi(t,\delta)$, so that, after straightforward rearrangements, equation (4.5) yields

$$\frac{w(t) - w(t-\delta)}{\delta} - \frac{\chi(t,\delta)}{\delta} = -w(t-\delta) + E(\tilde{R}^{(t)}_{\tau^\star_t} \mid N(\delta) = 1) + \frac{o(\delta)}{\delta}. \quad (4.9)$$

Now let δ go to zero on both sides of (4.9). We know, from the continuity of $w(t)$ and Lemma 4.4, that the limit of the rhs exists. Therefore the limit of the lhs must also exist. Also, from Lemma 4.3, we know that $\chi(t) := \lim_{\delta\to 0} \chi(t,\delta)/\delta$ exists and is finite for all t. Hence $\lim_{\delta\to 0}(w(t) - w(t-\delta))/\delta$ also exists and thus $w(\cdot)$ must be differentiable on $\mathbb{R}$. This completes the proof of Theorem 4.2. ∎

Remark 7 *Set $h(t,x) = w(t \mid x) - w(t)$. Equation (4.3) can be rewritten*

$$w'(t) + w(t) = \int_0^1 \min\{1 + xt, w(t) + h(t,x)\}dx + \chi(t), \quad (4.10)$$

with $\chi(t) = \Pi'(t)P(T_{\tau^\star_t} > t)$. This yields (4.2). Also note that our assumptions on $\Pi(\cdot)$ imply that $\Pi'(t)$ is positive and uniformly bounded on $\mathbb{R}$. Hence $\chi(t) \leq KP(T_{\tau^\star_t} > t)$ for some $K > 0$ and thus, from remark 6, $\chi(t) \to 0$ as $t \to \infty$. For all strategies we can think of as being close to optimal, $\chi(t)$ proves to decrease exponentially fast to zero. Hence we suggest to focus interest on the simpler equation

$$w'(t) + w(t) = \int_0^1 \min\{1 + xt, w(t) + h(t,x)\}dx. \quad (4.11)$$

This is still an equation in two unknown functions, and the challenge is to find a good estimate for $h(t,x)$.

If $w(t \mid x)$ were known, then equation (4.3) would be solvable, at least numerically, and this solution would yield the value function for the Poisson embedded problem. This defines a new secondary aim into Robbins' problem, namely to estimate the difference between $w(t \mid x)$ and $w(t)$ sufficiently precisely in order to be able to use (4.3) to obtain estimates on $w(t)$. This problem turns out to share the same difficulties as Robbins' problem itself. We are, however, able to give some rough estimates on $w(t \mid x)$.

Proposition 4.5 *For all* $x \in [0,1]$,

$$w(t) \leq w(t|x) \leq w(t) + 1 - e^{-(1-x)t}$$

Proof: The first inequality is evident. To show the second one, we need to express $w(t \mid x)$ in terms of $w(t)$. This is done in the following way.

$$w(t|x) =$$

$$w(t|x) + \inf_{\tau \in \mathcal{T}_t} \{-\mathrm{P}[X_\tau \geq x, T_\tau < t]\} - \inf_{\tau \in \mathcal{T}_t} \{-\mathrm{P}[X_\tau \geq x, T_\tau < t]\}$$

$$\leq w(t) + \sup_{\tau \in \mathcal{T}_t} \{\mathrm{P}[X_\tau \geq x, T_\tau < t]\}.$$

Now let $\tau^\star = \inf_i \{X_i > x\}$. This is a stopping time which stops on the first arrival, if any, over x. It is clear that $\tau^\star$ yields the supremum appearing in the previous inequality, therefore

$$\begin{aligned} w(t|x) &\leq w(t) + \mathrm{P}[X_{\tau^\star} \geq x, T_{\tau^\star} < t] \\ &\leq w(t) + \mathrm{P}[T_{\tau^\star} < t] \\ &= w(t) + 1 - \mathrm{P}[\text{ there is no arrival in } [0,t] \times [x,1]] \\ &= w(t) + 1 - e^{-(1-x)t} \end{aligned}$$

■

We define the difference function

$$h(t,x) = w(t \mid x) - w(t).$$

For each $x \in [0,1]$, $h(t,x)$ is the difference between two continuous functions and thus is continuous in t. Moreover, this function is decreasing in x and satisfies

$$0 \leq h(t,x) \leq 1 - e^{-(1-x)t}.$$

Since estimates on $h(t,x)$ yield estimates on $w(t \mid x)$, it is natural to consider $h(t,x)$ for specific strategies. In this spirit, for every strategy $\tau \in \mathcal{T}$, we define the function

$$h_\tau(t,x) = \mathrm{P}[X_\tau > x, T_\tau < t].$$

This function has some interesting properties.

Proposition 4.6 *Let $g_t(.)$ be a threshold function and let τ be the corresponding memoryless threshold rule. Let $h_\tau(t,x)$ be defined by $h_\tau(t,x) = \mathrm{P}[X_\tau > x, T_\tau < t]$. Then*

$$h_\tau(t,x) = \begin{cases} 1 - e^{-\mu_t(t)} - x\displaystyle\int_0^t e^{-\mu_t(s)}ds & 0 \le x \le g_t(0) \\ \\ 1 - e^{-\mu_t(t)} - x\displaystyle\int_{g_t^{-1}(x)}^t e^{-\mu_t(s)}ds & g_t(0) \le x \le 1 \end{cases}$$

Moreover this functions satisfies

$$\begin{array}{ll} h_\tau(t,x) = 1 - e^{-\mu(t)} - x\mathrm{E}[T_\tau] & \text{if } \ 0 \le x \le g_t(0) \\ h_\tau(t,x) > 1 - e^{-\mu(t)} - x\mathrm{E}[T_\tau] & \text{if } \ g_t(0) < x \le 1 \end{array}$$

Proof: From the definition of τ we know that, conditionally to $T_\tau = s \in [0,t)$, X_τ is distributed uniformly on $[0, g_t(s)]$. Hence, using the density of T_τ which is given by (2.16), we see that if $0 \le x \le g_t(0)$, then

$$\begin{aligned} \mathrm{P}[X_\tau > x, T_\tau < t] &= \int_0^t \frac{g_t(s) - x}{g_t(s)} f_{T_\tau}(s)ds \\ &= 1 - e^{-\mu(t)} - x\int_0^t e^{-\mu_t(s)}ds \\ &= 1 - e^{-\mu(t)} - x\mathrm{E}[T_\tau]. \end{aligned}$$

Likewise, if $g_t(0) \le x \le 1$, then we see from the definition of a threshold rule that $\mathrm{P}[X_\tau > x \mid T_\tau = s]$ will be identically nil for all $s \in [0, g_t^{-1}(x)]$. Therefore

$$\begin{aligned} \mathrm{P}[X_\tau > x, T_\tau < t] &= \int_0^{g_t^{-1}(x)} f_{T_\tau}(s)ds + \int_{g_t^{-1}(x)}^t \frac{g_t(s) - x}{g_t(s)} f_{T_\tau}(s)ds \\ &= 1 - e^{-\mu_t(t)} - x\int_{g_t^{-1}(0)}^t e^{-\mu_t(s)}ds \\ &\ge 1 - e^{-\mu_t(t)} - x\int_0^t e^{-\mu_t(s)}ds. \end{aligned}$$

This yields the result. ∎

The function $\chi(t)$ is a nuisance parameter of equation (4.3). However, it is uniformly bounded by an $o(1/t)$, and will not play any role asymptotically.

Now choose some constant $c > 1$. Then for all $s > 0$ in which the differential equation is satisfied, we can write

$$\begin{aligned} w'(s) + w(s) &\leq \int_0^{\frac{c}{c+s}} (1 + xs)\mathrm{dx} + \int_{\frac{c}{c+s}}^1 (w(s) + h(s,x))\mathrm{dx} + \chi(s) \\ &\leq w(s) - \tfrac{c}{c+s} w(s) + H(s,c) \end{aligned}$$

where

$$H(s,c) = \int_0^{\frac{c}{c+s}} (1 + xs)\mathrm{dx} + \int_{\frac{c}{c+s}}^1 h(s,x)\mathrm{dx} + \chi(s). \tag{4.12}$$

Hence

$$w'(s) + \frac{c}{s+c} w(s) \leq H(s,c), \tag{4.13}$$

Multiplying both sides of (4.13) by $(c+s)^c$, we get

$$((c+s)^c w(s))' \leq (c+s)^c H(s,c),$$

which after integration yields

$$w(t) \leq (c+t)^{-c} \int_0^t (c+s)^c H(s,c)\mathrm{ds}. \tag{4.14}$$

Example 2 *We saw that $h(t,x) \leq 1 - e^{-(1-x)t}$. Applying (4.14) to $1 - e^{-(1-x)t}$ and taking the limit of this expression for $t \to \infty$ yields the trivial upper bound $w \leq \infty$. In fact, one can show that any upper bound on $h(t,x)$ which is not asymptotically equivalent to zero will always yield from (4.14) a trivial upper bound on $w(t)$.*

Example 3 *If τ is the memoryless threshold strategy defined by the threshold function $g_t(s) = \frac{c}{t-s+c}$ (see Example 1), then an explicit computation of $h_\tau(t,x) = \mathrm{P}[X_{N(\tau)} > x, T_\tau < t]$ yields*

$$h_\tau(t,x) = \begin{cases} 1 + g_t(0)^c(\frac{c}{c+1}x - 1) - \frac{c}{c+1}\frac{x}{g_t(0)} & 0 \leq x \leq g_t(0) \\ \\ g_t(0)^c(\frac{c}{c+1}x - 1) + \frac{1}{c+1}\left(\frac{g_t(0)}{x}\right)^c & g_t(0) \leq x \leq 1 \end{cases} \tag{4.15}$$

Applying (4.14) to $h_\tau(t,x)$, and taking the limit for $t \to \infty$ we see that this yields, as before, the upper bound

$$w \leq 1 + \frac{c}{2} + \frac{1}{c^2 - 1}.$$

4.3 Conclusion

Although we cannot solve the differential equation explicitly, we hope that it will prove to be a starting point for a numerical analysis of the behavior of $w(t)$, and hence of $v(n)$. Indeed, equation (4.11) bypasses the full-history dependence which lies at the heart of Robbins' Problem. The idea is to plug estimates of $h(t, x)$ into (4.11) and to study the corresponding solutions. To facilitate this approach, we have, throughout the paper, avoided specifying the penalty function $\Pi(t)$ in order to leave room for the choice of initial conditions on "candidate" solutions. The key to success for improvements on the known bounds on v should therefore be a sufficiently close estimate of $h(t, x)$. Unfortunately, the estimates we have obtained so far are not precise enough. The problem remains a challenge but, as we see it, with a new focus.

Bibliography

[1] Assaf, D. and Samuel-Cahn, E. (1996). *The secretary problem; minimizing the expected rank with i.i.d. random variables*, Adv. Appl. Prob., Vol. 28, pp. 828-852.

[2] Bruss, F. T. (1987). *On an optimal selection problem by Cowan and Zabcyk*, J. Appl. Probab. 24, pp. 918-928.

[3] Bruss, F.T. (2000). *Sum the odds to one and stop*, Ann. Prob. 28, pp. 1384-1391.

[4] Bruss, F.T. (2005). *What is known about Robbins' Problem?*, J. Appl. Prob. 42, pp. 108 - 120.

[5] Bruss, F. T. and Ferguson, T. S. (1993). *Minimizing the expected rank with full information*, J. Appl. Prob. 30, pp. 616 - 626.

[6] Bruss, F. T. and Ferguson, T. S. (1996). *Half-Prophets and Robbins' Problem of Minimizing the expected rank*, Springer Lecture Notes in Stat. 114, Vol. 1 in honor of J.M. Gani, pp. 1-17.

[7] Bruss, F. T. and Rogers, L. C. G. (1991). *Embedding optimal selection problems in a Poisson process*, Stoch. Proc. Appl. 38, pp. 267-278.

[8] Bruss, F. T. and Paindaveine, D. (2000). *Selecting a sequence of last successes in independent trials*, J. Appl. Prob. 24, pp. 389 - 399.

[9] Bruss, F. T. and Swan, Y. C. (2009). A continuous-time approach to Robbins' problem of minimizing the expected rank, *J. Appl. Prob.*, **46**, 1–18.

[10] Cayley, A. (1875). *Mathematical questions and their solutions*, Educational times 22, pp. 18-19.

[11] Chow, Y.S., Moriguti, S., Robbins, H. and Samuels, S. M. (1964). *Optimal selection based on relative ranks*, Israel Journal of Mathematics, Vol. 2, pp. 81-90.

[12] Chow, Y.S., Robbins, H. and Siegmund, D. (1971). *Great expectations: The Theory of Optimal Stopping*, Houghton Mifflin Company, Toronto, London.

[13] Cowan, R. and Zabczyk, J. (1978). *An optimal selection problem associated with the Poisson process*, Theory Probab. Appl. 23, pp. 584-592.

[14] Cunningham, F. (1967). *Taking Limits under the Integral Sign*, Mathematics Magazine, Vol. 40, No. 4, pp. 179-186.

[15] Dynkin, E. B. (1963). *The optimum choice of the instant for stopping a Markov process*, Soviet Math. Dokl. 4, pp. 627-629.

[16] Feller, W. (1968). *An Introduction to Probability and its Applications*, Volume I, Third Edition, John Wiley and Sons, Inc., New York - London - Sydney.

[17] Ferguson, T. S. (1989). *Who solved the secretary problem?*, Statistical Science 4, pp. 282-289 (with discussion).

[18] Ferguson, T. S. (2000). *Optimal Stopping and Applications*, unpublished manuscript, available at http://www.math.ucla.edu/ tom/Stopping/Contents.html.

[19] Freeman, P. R. (1983). *The secretary problem and its extensions: a review*, Int. Statist. Rev. 51, pp. 189-206.

[20] Gardner, M. (1960). *Mathematical games*, Scientific American 202, pp. 152 and 178-179.

[21] Gilbert, J. P. and Mosteller, F. (1966). *Recognizing the maximum of a sequence*, J. Amer. Statist. Assov. 61, pp. 35-73.

[22] Gnedin, A. V. (2002). *Best choice from the planar Poisson process*, arXiv:math.PR/0209050 v2.

[23] Gnedin, A. V. (2006). *Recognising the Last Record of a Sequence*, arXiv:math.PR/0602278 v1 13.

[24] Gnedin, A. V. (2007). *Optimal Stopping with Rank-Dependent Loss*, J. Appl. Prob., Vol. 44, Number 4, pp. 996-1011.

[25] Gnedin, A. V. and Sakaguchi, M. (1992). *On a best-choice problem related to the Poisson process*, Contemporary Math. 125, pp. 59-64.

[26] Heinonen, J. (2004). *Lectures on Lipschitz Analysis*, Lectures at the 14th Jyväskylä Summer School, available at http://www.math.jyu.fi/research/ber.html.

[27] Karlin, S. (1962). *Stochastic models and optimal policy for selling an asset*, Chapter 9 of *Studies in Applied Probability and Management Science*, Ed. by K. Arrow, S. Karlin and W. Scarf, Stanford University Press.

[28] Lindley, D. V. (1961). *Dynamic programming and decision theory*, Appl. Statistics . 10, pp. 39-51.

[29] Moser, L. (1956). *On a problem of Cayley*, Scripta Math. 22, pp. 289-292.

[30] Petruccelli, J. D. (1988). *Secretary Problem*, Encyclopedia of Statistical Sciences, Vol. 8, pp. 326-329, S. Kotz and N. Johnson, eds. Wiley-Interscience, New York.

[31] Rasche, M. (1975). *Allgermeine Stopprobleme*, Technical report, Institut für Mathematische Statistik, Universität Münster.

[32] Samuels, S. M. (1982). *Exact solutions for the full information best choice problem*, Purdue Univ. Stat. Dept. Mimea Series, pp. 82-17.

[33] Sakaguchi, M. (1976). *Optimal stopping problems for randomly arriving offers*, Math. Japonicae 21, pp. 201-217.

[34] Samuels, S. M. (1991). *Secretary problems.* In Handbook of sequential analysis, volume 118 of Statist. Textbooks Monogr., pp. 381-405. Dekker, New York.