

# A CONTINUOUS-TIME APPROACH TO ROBBINS' PROBLEM OF MINIMIZING THE EXPECTED RANK

F. THOMAS BRUSS \* \*\* AND  
YVIK C. SWAN, \* \*\*\* *Université Libre de Bruxelles*

## Abstract

Let  $X_1, X_2, \dots, X_n$  be independent random variables uniformly distributed on  $[0, 1]$ . We observe these sequentially and have to stop on exactly one of them. No recall of preceding observations is permitted. What stopping rule minimizes the expected rank of the selected observation? What is the value of the expected rank (as a function of  $n$ ) and what is the limit of this value when  $n$  goes to  $\infty$ ? This full-information expected selected-rank problem is known as Robbins' problem of minimizing the expected rank, and its general solution is unknown. In this paper we provide an alternative approach to Robbins' problem. Our model is similar to that of Gneden (2007). For this, we consider a continuous-time version of the problem in which the observations follow a Poisson arrival process on  $\mathbb{R}^+ \times [0, 1]$  of homogeneous rate 1. Translating the previous optimal selection problem in this setting, we prove that, under reasonable assumptions, the corresponding value function  $w(t)$  is bounded and Lipschitz continuous. Our main result is that the limiting value of the Poisson embedded problem exists and is equal to that of Robbins' problem. We prove that  $w(t)$  is differentiable and also derive a differential equation for this function. Although we have not succeeded in using this equation to improve on bounds on the optimal limiting value, we argue that it has this potential.

*Keywords:* Optimal stopping; secretary problem; Robbins' problem; Poisson embedding

2000 Mathematics Subject Classification: Primary 60G40

## 1. Introduction

### 1.1. Robbins' problem

Let  $X_1, X_2, \dots, X_n$  be a sequence of independent and identically distributed (i.i.d.) observations, sampled from the uniform distribution on  $[0, 1]$ . A decision maker observes the  $X_k$ s sequentially and has to stop on exactly one of them. No recall of preceding observations is permitted. What stopping rule minimizes the expected rank of the selected observation? What is the value  $v(n)$  of the minimal obtainable expected rank and what is its limit,  $v = \lim_{n \rightarrow \infty} v(n)$ ?

Throughout the paper, we will refer to this full-information expected selected-rank problem as Robbins' problem. This denomination was coined at the International Conference on Sequential Search and Selection in Real Time (Amherst, 1990) when Herbert Robbins brought the problem to public attention. Although the corresponding no-information expected-rank

---

Received 16 November 2006; revision received 30 October 2008.

\* Postal address: Département de Mathématiques, Faculté des Sciences, Université Libre de Bruxelles, Campus Plaine, CP 210. B-1050 Brussels, Belgium.

\*\* Email address: tbruss@ulb.ac.be

\*\*\* Email address: yvswan@ulb.ac.be

Research supported in part by a mandate of chargé de recherches du F.R.S.-FNRS, Belgium.

problem was solved in 1964 (see Chow *et al.* (1964)), some fundamental questions related to the full-information version are still open (see the survey paper Bruss (2005) and the references therein). This lack of progress is mainly due to the optimal strategy being fully history dependent, which means that the decision to stop on an arrival  $X_k$  depends on the full set of values  $\{X_1, \dots, X_{k-1}\}$ . The problem is therefore intrinsically infinite-dimensional, and any hope for progress lies in finding an alternative approach that bypasses this complexity.

Recently, Gnedin (2007) proposed a limit model for optimal stopping problems with rank-dependent loss, in which the process is chosen to be a homogeneous Poisson point process in the strip  $[0, 1] \times \mathbb{R}^+$  with intensity measure  $dt dx$ . Hence, in this model, time runs from 0 to 1, and there are infinitely many arrivals and a well-defined smallest arrival in any strip  $[s, s + \Delta s] \times \mathbb{R}^+$ . With this model, Gnedin obtained bounds on the stopping value for Robbins' problem and, in particular, showed that the full history dependence of the optimal decision process persists in the limit.

Our model is similar to Gnedin's in that we also consider a version of Robbins' problem for a random number of arrivals occurring according to a Poisson process. Our model is not a limit model, however. Our goal is different. We aim to construct an alternative approach which allows for direct comparison with Robbins' problem for finite  $n$ .

**1.2. The Poisson embedded Robbins' problem**

The problem is as follows. A decision maker observes opportunities occurring according to a planar Poisson process of homogeneous rate 1 on  $\mathbb{R}^+ \times [0, 1]$ . Here the first coordinate stands for time and the second for the corresponding value. He inspects each arrival and has to choose exactly one before a given time  $t > 0$ . Decisions are to be made immediately upon inspection, and no recall of preceding observations is permitted. The loss incurred for selecting an arrival of value  $X$  is defined as its absolute rank, that is, the number of observations in  $[0, t]$  which are not larger than  $X$ . If no decision has been reached before the given time  $t$  then the loss is equal to some nonnegative 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 nonanticipating strategy which minimizes his expected loss.

Let  $(T_1, X_1), (T_2, X_2), \dots$  denote the point arrival process. The random variables  $T_1 \leq T_2 \leq \dots$  are the arrival times of a homogeneous Poisson counting process  $(N(s))_{s \geq 0}$  of rate 1 with associated i.i.d. random values  $X_1, X_2, \dots$ . With this notation, the absolute rank of the  $k$ th arrival  $X_k$  is defined with respect to  $t$  by

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

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

$$\tilde{R}_k^{(t)} := R_k^{(t)} \mathbf{1}_{\{T_k \leq t\}} + \Pi(t) \mathbf{1}_{\{T_k > t\}}, \tag{1.1}$$

and the objective of the decision maker is to use a stopping time  $\tau$  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  $\{T_r\}_{r \geq 1}$  of arrival times of the point process, which satisfy  $\{\tau \leq s\} \in \mathcal{F}_s$ , where

$$\mathcal{F}_s = \sigma\{(N(u))_{0 \leq u \leq s}, (T_1, X_1), \dots, (T_{N(s)}, X_{N(s)})\},$$

and where it is understood that  $\mathcal{F}_s = \sigma\{(N(u))_{0 \leq u \leq s}\}$  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 1.1.** From now on we will *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 notation  $R_\tau, X_\tau$ , and  $T_\tau$  are well defined and will be used systematically throughout the paper.

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

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

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

**Remark 1.2.** 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  $\Pi(\cdot)$  is increasing and differentiable with bounded derivative. Hence, this function is Lipschitz continuous and satisfies

$$\lim_{t \rightarrow \infty} \frac{\Pi(t)}{t} \leq \kappa \quad \text{for some } \kappa \in (0, \infty). \tag{1.2}$$

**Remark 1.3.** Note that, for all  $\tau$ , the expected rank of an arrival selected by  $\tau$  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  $r_k = \sum_{j=1}^k \mathbf{1}_{\{X_j \leq X_k\}}$  is the relative rank of the  $k$ th 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. Properties of the value function

In this section we prove that the Poisson embedded problem is well defined and that the corresponding value function  $w(t)$  is bounded and Lipschitz continuous for sufficiently large  $t$ .

Let  $\tau \in \mathcal{T}$ , and let  $w_\tau(\cdot)$  be the expected loss incurred for using  $\tau$ , i.e.  $w_\tau(t) = E(\tilde{R}_\tau^{(t)})$ . Some interesting results are already obtained by considering  $w_\tau(t)$  for specific choices of  $\tau$  and using the bounds

$$0 \leq w(t) \leq w_\tau(t).$$

**Remark 2.1.** Note that, for all  $t$  such that  $\Pi(t) \geq 1$ , the value function satisfies  $w(t) \geq 1$ .

First consider the stopping rule  $\tau$  that stops on the first arrival after time  $t$ . Clearly,  $w_\tau(t) = \Pi(t)$ , and, thus,  $0 \leq \limsup_{t \rightarrow 0^+} w(t) \leq \limsup_{t \rightarrow 0^+} \Pi(t) = 0$ . Hence, although the problem is meaningless for  $t = 0$ , we see that, by posing  $w(0) = 0$ , the value function  $w(t)$  can be continuously prolonged at  $t = 0$ .

**2.1. Memoryless threshold rules**

A memoryless threshold rule  $\tau$  is defined as

$$\tau = \inf\{i \geq 1 \text{ such that } X_i \leq \varphi(T_i)\} \tag{2.1}$$

for some real-valued function  $\varphi(\cdot)$  defined on  $\mathbb{R}^+$ . Such rules are called memoryless because decisions depend only on the values of the arrivals and not otherwise on the history of the process. They have been studied in the discrete setting in Bruss and Ferguson (1993), (1996) and Assaf and Samuel-Cahn (1996), and in a continuous-time limit model in Gnedin (2007).

Although we can work in all generality with definition (2.1), it is relevant at this point to consider the threshold functions

$$\varphi(s) := \varphi_{t,c}(s) = \begin{cases} \frac{c}{t-s+c} & \text{if } 0 \leq s \leq t, \\ 1 & \text{otherwise.} \end{cases}$$

For all  $c > 1$ , the corresponding memoryless threshold rule is well defined and satisfies  $P(T_\tau < \infty) = 1$ . Also, since such a strategy does not stop before time  $s$  unless there has been an arrival below the threshold before that time, we immediately obtain from the planar Poisson process assumption

$$P(T_\tau > s) = e^{-\mu(s)},$$

where  $\mu(s) = \int_0^s \varphi(u) du$ ,  $s \leq t$ . This enables us to compute the associated value  $w_\tau(\cdot)$  explicitly.

Conditioning on the time of first acceptance, we obtain

$$w_\tau(t) = \int_0^t E(R_\tau^{(t)} | T_\tau = s) \varphi(s) e^{-\mu(s)} ds + \Pi(t) e^{-\mu(t)}. \tag{2.2}$$

Conditionally on the event  $\{T_\tau = s\}$ , the value of the accepted arrival is uniformly distributed on  $[0, \varphi(s)]$ . Hence,

$$E(R_\tau^{(t)} | T_\tau = s) = \frac{1}{\varphi(s)} \int_0^{\varphi(s)} E(R_\tau^{(t)} | T_\tau = s, X_\tau = x) dx.$$

Now consider the expected rank  $E(R_\tau^{(t)} | T_\tau = s, X_\tau = x)$ . The event  $\{T_\tau = s, X_\tau = x\}$  implies that there have been no previous smaller arrivals that were under the threshold. Therefore, the expected rank of the selected arrival will depend only on the expected number of arrivals in  $A_1$  and  $A_2$ , as illustrated in Figure 1.

Straightforward computations then yield

$$\begin{aligned} & E(R_\tau^{(t)} | T_\tau = s, X_\tau = x) \\ &= \begin{cases} 1 + x(t-s) & \text{if } 0 \leq x \leq \varphi(0), \\ 1 + x(t-s) + \int_0^{\varphi^{-1}(x)} \left( \frac{x - \varphi(u)}{1 - \varphi(u)} \right) du & \text{if } \varphi(0) \leq x \leq \varphi(s), \end{cases} \end{aligned} \tag{2.3}$$

where the second part of (2.3) holds because, conditionally on the event  $\{T_\tau = s\}$ , the value of an arrival occurring at some time  $0 \leq u \leq \varphi^{-1}(x)$  is uniformly distributed on  $[\varphi(u), 1]$ . Integrating (2.3) we obtain

$$E(R_\tau^{(t)} | T_\tau = s) = 1 + \frac{1}{\varphi(s)} \left( \int_0^{\varphi(s)} x(t-s) dx + \int_{\varphi(0)}^{\varphi(s)} \int_0^{\varphi^{-1}(x)} \frac{x - \varphi(u)}{1 - \varphi(u)} du dx \right),$$

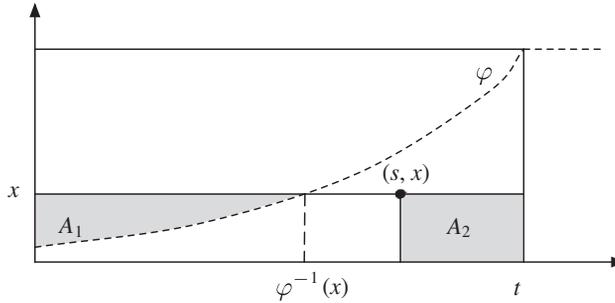


FIGURE 1: Conditionally on  $\{T_\tau = s, X_\tau = x\}$ , the expected rank of  $X_\tau = x$  is 1 plus the expected number of arrivals in  $A_1$  and  $A_2$ .

and, thus, from (2.2),

$$w_\tau(t) = 1 + (\Pi(t) - 1)e^{-\mu(t)} + \frac{1}{2} \int_0^t \varphi(s)^2(t - s)e^{-\mu(s)} ds + \frac{1}{2} \int_0^t \int_0^s \frac{(\varphi(s) - \varphi(u))^2}{1 - \varphi(u)} du e^{-\mu(s)} ds.$$

Now choose  $c > 1$ . Computing the above expression explicitly with  $\varphi(s) = c/(t - s + c)$ , we easily see that

$$(\Pi(t) - 1)e^{-\mu(t)} \rightarrow 0 \quad \text{as } t \rightarrow \infty,$$

i.e. the effect of the penalty function vanishes for large  $t$ . We can also check that  $w_\tau(t)$  is increasing for sufficiently large  $t$  and satisfies

$$\lim_{t \rightarrow \infty} w_\tau(t) = 1 + \frac{c}{2} + \frac{1}{c^2 - 1}.$$

This last expression is minimal for  $c = 1.9469\dots$  and the corresponding limiting value is  $2.33182\dots$ . Hence, since  $w(t) \leq w_\tau(t)$  for all  $t$ , we obtain the following result.

**Proposition 2.1.** *The value function of the Poisson embedded Robbins' problem is bounded on  $\mathbb{R}$  and satisfies*

$$1 \leq w(t) \leq 2.33183$$

for all sufficiently large  $t$ .

These values (the upper bound and the minimal  $c$  achieving this bound) are identical to those obtained in Bruss and Ferguson (1993) and Assaf and Samuel-Cahn (1996) for the discrete problem. They also coincide with those obtained in Gnedin (2007) for a Poisson embedded limit version of Robbins' problem. This is one of the common features of the original problem and the Poisson version. However, with respect to monotonicity of the value function, our result is much weaker.

### 2.2. Continuity of the value function

**Lemma 2.1.** *For all sufficiently large  $t$  and all  $\delta > 0$ ,*

$$w(t + \delta) - w(t) \geq -3\delta(\delta + 1). \tag{2.4}$$

*Proof.* Consider the Poisson embedded problem with horizon  $t + \delta$ . Recall the notation  $\tilde{R}_k^{(t)}$  from (1.1). By conditioning on the number  $N(\delta)$  of arrivals in  $(0, \delta)$  we obtain

$$w(t + \delta) \geq e^{-\delta} \inf_{\tau} E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 0) + \delta e^{-\delta} \inf_{\tau} E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 1), \tag{2.5}$$

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

We first consider the first term appearing on the right-hand side of (2.5). From the homogeneity assumptions on the arrival process, we see that solving the Poisson embedded Robbins' problem on  $[0, t + \delta]$  with no arrivals before time  $\delta$  is equivalent to solving the same problem on  $[0, t]$  with penalty  $\Pi(t + \delta)$ . Since  $\Pi(\cdot)$  is increasing, this implies that

$$\inf_{\tau} E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 0) \geq w(t). \tag{2.6}$$

Next consider the second term of (2.5). By conditioning on the value  $X$  of the (only) arrival in  $(0, \delta)$  we obtain

$$\begin{aligned} \inf_{\tau} E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 1) &= \inf_{\tau} \int_0^1 E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 1, X = x) dx \\ &\geq \int_0^1 \inf_{\tau} E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 1, X = x) dx. \end{aligned} \tag{2.7}$$

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 yields an expected loss given by

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

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

$$\inf_{\tau} E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 1, X = x) = \min\{1 + xt, E_t(x, \delta)\}. \tag{2.8}$$

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 (2.7) and (2.8),

$$\inf_{\tau} E(\tilde{R}_{\tau}^{(t+\delta)} \mid N(\delta) = 1) \geq \int_0^1 \min\{1 + xt, w(t)\} dx. \tag{2.9}$$

Combining (2.5), (2.6), and (2.9) then yields

$$w(t + \delta) \geq e^{-\delta} w(t) + \delta e^{-\delta} \int_0^1 \min\{1 + xt, w(t)\} dx. \tag{2.10}$$

Now choose  $t$  sufficiently large to ensure that  $w(t) \geq 1$ . Then there exists an  $x_0 \in [0, 1)$  for which  $1 + x_0t = w(t)$  and, thus,

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

From (2.10), 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 sufficiently large  $t$ , this implies that

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

and (2.4) follows.

**Lemma 2.2.** *There exists a constant  $L > 0$  such that, for all  $t$  and  $\delta$  positive,*

$$w(t + \delta) - w(t) \leq L\delta. \tag{2.11}$$

*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,  $w(t + \delta) = \inf_{\mathcal{T}} E(\tilde{R}_\tau^{(t+\delta)}) \leq \inf_{\mathcal{K}_t} E(\tilde{R}_\tau^{(t+\delta)})$ . 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

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

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

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

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

$$\begin{aligned} w(t + \delta) &\leq \inf_{\mathcal{K}_t} \{E(R_\tau^{(t)} \mathbf{1}_{\{T_\tau < t\}}) + \Pi(t) P(T_\tau > t)\} \\ &\quad + (\Pi(t + \delta) - \Pi(t)) \sup_{\mathcal{K}_t} P(T_\tau > t) + \delta \\ &= \inf_{\mathcal{K}_t} E(\tilde{R}_\tau^{(t)}) + (\Pi(t + \delta) - \Pi(t)) + \delta. \end{aligned}$$

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

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

The hypothesis on  $\Pi(\cdot)$  gives (2.11).

Lemmas 2.1 and 2.2 immediately yield the following theorem.

**Theorem 2.1.** *The value function  $w(t)$  is continuous on  $\mathbb{R}$  and Lipschitz continuous on  $(t_0, \infty)$  for some sufficiently large  $t_0$ .*

### 3. Existence of the asymptotic value

Although the Poisson embedded problem is interesting in its own right, we have developed this model in order to study the original  $n$ -arrival Robbins' problem. Our aim in this section is to show that the Poisson embedded model is the right setting for this endeavor.

Recall the statement of the original Robbins' problem from Section 1. For all  $n$ , its value is defined as

$$v(n) = \inf_{\tau} E(R_{\tau}^{(n)}),$$

where the infimum is taken over all adapted strategies and  $R_k^{(n)} = \sum_{j=1}^n \mathbf{1}_{\{X_j \leq X_k\}}$ . Bruss and Ferguson (1993), (1996) and Assaf and Samuel-Cahn (1996) proved that, for each  $n$ , there exists an optimal strategy  $\tau_n^*$  for which  $E(R_{\tau_n^*}^{(n)}) = v(n)$ , that  $v(n)$  is increasing, and that  $v(n)$  converges to a limit  $v$  satisfying  $1.9 < v < 2.33$ .

**Proposition 3.1.** *For all  $\varepsilon > 0$ , there exists a  $t^* > 0$  such that, for all  $t \geq t^*$ ,*

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

*Proof.* Fix  $\varepsilon > 0$ , 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

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

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

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

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

$$\inf_{\sigma \in \tilde{\mathcal{F}}} E(\tilde{R}_{\sigma}^{(t)} \mid N(t) = k) \geq v(\lfloor \Pi(t) \rfloor) \quad \text{for all } k > \Pi(t). \tag{3.3}$$

To see this, let  $v_Q(k) = \inf_{\sigma \in \tilde{\mathcal{F}}} E(\tilde{R}_{\sigma}^{(t)} \mid N(t) = k)$ . The same half-prophet argument as that used in 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  $k > \Pi(t)$ ,  $v_Q(k) \geq$

$v_Q(\lfloor \Pi(t) \rfloor)$ , where  $\lfloor x \rfloor$  denotes the largest integer not greater than  $x$ . 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 an  $m_0 = m_0(\varepsilon) \in \mathbb{N}$  such that  $v(m) > v - \varepsilon$  for all  $m \geq m_0$ . The monotonicity of  $\Pi(\cdot)$  implies that there exists a  $t_0 = t_0(\varepsilon)$  such that  $\Pi(t) > m_0$  for all  $t \geq t_0$ . Therefore, from (3.4),

$$w(t) \geq (v - \varepsilon) \sum_{k=m_0}^{\infty} P(N(t) = k) = (v - \varepsilon) P(N(t) \geq m_0) \quad \text{for all } t \geq t_0.$$

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

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

and, thus, since  $v \leq 3$ ,

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

**Corollary 3.1.** *If the limit  $w = \lim_{t \rightarrow \infty} w(t)$  exists then it satisfies  $w \geq v$ .*

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

**Lemma 3.1.** *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*

$$\lim_{n \rightarrow \infty} (n P(N(n) < n - n^\alpha)) = 0. \tag{3.5}$$

*Proof.* Here  $N(n)$  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  $\alpha$  between  $\frac{1}{2}$  and  $\frac{2}{3}$ , in order to ensure that  $n^{\alpha-1/2}$  increases to  $\infty$  as  $n \rightarrow \infty$  and that  $(n^{\alpha-1/2})^3/\sqrt{n} = n^{3\alpha-2}$  decreases to 0 as  $n \rightarrow \infty$ . Then we can apply a theorem on normal approximation (see Feller (1968, p. 193)) to obtain

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

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

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

Hence, altogether,

$$n P(N(n) < n - n^\alpha) \sim \frac{1}{\sqrt{2\pi}} n^{3/2-\alpha} \exp\left\{-\frac{1}{2}(n^{\alpha-1/2})^2\right\} < \frac{1}{\sqrt{2\pi}} \frac{n}{\exp\{n^{2\alpha-1/2}\}},$$

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

**Proposition 3.2.** *Let  $\alpha \in (\frac{1}{2}, 1)$  and  $t > 0$ . Define  $\beta_t = \lfloor t - t^\alpha \rfloor$ , and let  $\tau_{\beta_t}^*$  be the optimal stopping rule for the discrete problem with  $\beta_t$  arrivals. Then, for all  $\varepsilon > 0$  and all sufficiently large  $t$ ,*

$$w(t) \leq v(\beta_t) + E(X_{\tau_{\beta_t}^*})(t^\alpha + 1) + \varepsilon.$$

*Proof.* Any strategy for the discrete case with  $n$  arrivals can be extended in a natural way to define a (suboptimal) strategy for the continuous case, so that we can consider  $\tau_{\beta_t}^*$  as a stopping rule acting in continuous time on  $[0, t]$ . Let  $\sigma_{\beta_t}$  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  $\beta_t$  arrivals before the horizon  $t$ . Conditioning on the number of arrivals in  $[0, t]$ , we obtain

$$\tilde{w}(\beta_t) = \tilde{w}(\beta_t \mid N(t) < \beta_t) P(N(t) < \beta_t) + \tilde{w}(\beta_t \mid N(t) \geq \beta_t) P(N(t) \geq \beta_t), \tag{3.7}$$

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

First suppose that  $N(t) < \beta_t$ . Since  $\sigma_{\beta_t}$  acts on  $\beta_t$  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  $E(R_{\sigma_{\beta_t}}^{(t)} \mid N(t) < \beta_t) \leq v(\beta_t)$ . This yields

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

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

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

for some positive constant  $K$ . This last inequality, combined with the assumptions on  $\Pi(\cdot)$  (see (1.2)) and Lemma 3.1, proves that

$$\tilde{w}(\beta_t \mid N(t) < \beta_t) P(N(t) < \beta_t) < \frac{\varepsilon}{2} \tag{3.8}$$

for sufficiently large  $t$ .

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

$$\tilde{w}(\beta_t \mid N(t) \geq \beta_t) = v(\beta_t) + E(X_{\sigma_{\beta_t}}(N(t) - \beta_t) \mid N(t) \geq \beta_t). \tag{3.9}$$

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

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

and, thus, from (3.9),

$$\tilde{w}(\beta_t \mid N(t) \geq \beta_t) = v(\beta_t) + E(X_{\tau_{\beta_t}^*})(E(N(t) \mid N(t) \geq \beta_t) - \beta_t). \tag{3.10}$$

Furthermore,

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

Now, since  $P(N(t) < \beta_t) \rightarrow 0$  as  $t \rightarrow \infty$ , we know that

$$\frac{1}{1 - P(N(t) < \beta_t)} < 1 + 2 P(N(t) < \beta_t)$$

for sufficiently large  $t$ . Therefore,

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

and, from Lemma 3.1,

$$E(N(t) \mid N(t) \geq \beta_t) \leq t + \frac{\varepsilon}{2}$$

for all sufficiently large  $t$ . From (3.10), this yields

$$\tilde{w}(\beta_t \mid N(t) \geq \beta_t) \leq v(\beta_t) + E(X_{\tau_{\beta_t}^*}) \left( t - \beta_t + \frac{\varepsilon}{2} \right),$$

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

$$w(\beta_t \mid N(t) \geq \beta_t) \leq v(\beta_t) + E(X_{\tau_{\beta_t}^*})(t^\alpha + 1) + \frac{\varepsilon}{2} \tag{3.11}$$

for sufficiently large  $t$ .

Combining (3.7), (3.8), and (3.11), we obtain

$$\tilde{w}(\beta_t) \leq v(\beta_t) + E(X_{\tau_{\beta_t}^*})(t^\alpha + 1) + \varepsilon,$$

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

$$w(t) \leq v(\beta_t) + E(X_{\tau_{\beta_t}^*})(t^\alpha + 1) + \varepsilon$$

for all sufficiently large  $t$ . This completes the proof.

From Corollary 3.1 and Proposition 3.2, we see that, in order to prove both the existence of  $w$  and its equality with  $v$ , we need

$$\lim_{t \rightarrow \infty} t^\alpha E(X_{\tau_{\beta_t}^*}) = 0 \quad \text{for some } \alpha > \frac{1}{2}. \tag{3.12}$$

For this, we will use the following theorem.

**Theorem 3.1.** *Let  $\tau_n^*$  be the optimal strategy for the discrete  $n$ -arrival Robbins' problem. Then, for all  $p > 1$ ,*

$$E(X_{\tau_n^*}) \leq E(R_{\tau_n^*}) \left( \sum_{k=1}^n a_{n,k}^p \right)^{1/p}, \tag{3.13}$$

where

$$a_{n,k}^p = k^{2-p} \binom{n}{k} \int_0^1 x^{p+k-1} (1-x)^{n-k} dx. \tag{3.14}$$

*Proof.* Conditioning on the ranks we obtain

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

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

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

for all  $p$  and  $q$  such that  $1/p + 1/q = 1$ . Hence,

$$E(X_{\tau_n^*}) \leq \sum_{k=1}^n k^{-1/q} (E(X_{(k)}^p))^{1/p} (k P(R_{\tau_n^*} = k))^{1/q}.$$

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 right-hand side of the above inequality for the same choice of  $p$  and  $q$  gives

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

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

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

The known expressions for the moments of the  $k$ th order statistics then yield (3.14).

Now choose  $p$  to be an integer greater than 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 obtain

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

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 large  $n$ ,

$$A(n, p) \approx \left(\frac{1}{n+p}\right)^p. \tag{3.15}$$

Also, since  $(1 + p/(k - 1))^{k-1} \leq e^p$ , we obtain

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

so that, using Newton's binomial formula on  $(1 + (p - 1)/k)^p$ , we obtain

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

For fixed  $p > 1$ , this last function is in the order of  $n^2$ , and, hence, from (3.15),

$$\left(\sum_{k=1}^n a_k^p\right)^{1/p} = A(n, p)B(n, p) \approx K_p n^{2/p-1} \tag{3.16}$$

for some positive constant  $K_p$ .

Equation (3.16) suffices to prove (3.12) for all  $\alpha \in (0, 1)$ . Indeed, we know that  $E(R_{\tau_{\beta_t}^*}) \leq v$  for all  $t$ . Thus, from Theorem 3.1,

$$t^\alpha E(X_{\tau_{\beta_t}^*}) \leq E(R_{\tau_{\beta_t}^*}) t^\alpha \left(\sum_{k=1}^{\beta_t} a_k^p\right)^{1/p} \leq v t^\alpha \left(\sum_{k=1}^{\beta_t} a_k^p\right)^{1/p} \text{ for all } p > 1.$$

Estimate (3.16) then yields

$$t^\alpha E(X_{\tau_{\beta_t}^*}) \approx K_p t^\alpha \beta_t^{2/p-1}$$

for some positive constant  $K_p$ , and it is easy to check that, for all  $p > 2$  and all  $\alpha < (p - 2)/p$ ,  $t^\alpha \beta_t^{2/p-1} \rightarrow 0$  for  $t \rightarrow \infty$ . Since there are no upper bounds on the choice of  $p$ , we obtain the following theorem.

**Theorem 3.2.** *Let  $0 \leq \alpha < 1$ , and let  $\beta_t = \lfloor t - t^\alpha \rfloor$ . The optimal stopping rule  $\tau_{\beta_t}^*$  satisfies*

$$\lim_{t \rightarrow \infty} t^\alpha E(X_{\tau_{\beta_t}^*}) = 0 \text{ for all } \frac{1}{2} < \alpha < 1.$$

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

Corollaries 3.1 and 3.2 immediately yield the following theorem.

**Theorem 3.3.** *The limiting value for the Poisson embedded Robbins' problem exists and satisfies*

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

Note that Theorem 3.2 implies, in particular, that the  $n$ -optimal stopping rule  $\tau_n^*$  satisfies

$$\lim_{t \rightarrow \infty} n^\alpha E(X_{\tau_n^*}) = 0 \quad \text{for all } 0 \leq \alpha < 1.$$

This fails for  $\alpha = 1$ . In fact, estimate (3.13) does not even suffice to prove that  $nE(X_{\tau_n^*})$  is bounded. Now a natural parallel has been drawn between Robbins' problem and *Moser's problem* (see Moser (1956)), in which the objective of the decision maker is to minimize the expected value of the selected observation. Since ranks and values have limiting correlation 1 as  $n \rightarrow \infty$  (see Bruss and Ferguson (1993)), it is then natural to believe that the optimal strategies in both problems should have similar behaviors—at least asymptotically. The optimal strategy  $\hat{\tau}_n$  for Moser's problem satisfies  $\lim_{n \rightarrow \infty} (nE(X_{\hat{\tau}_n})) = 2$ . It therefore seems intuitive that  $nE(X_{\tau_n^*})$  should be bounded. Assaf and Samuel-Cahn (1996) assumed that the optimal rule satisfies this condition. We therefore call it the Assaf and Samuel-Cahn hypothesis. This hypothesis implies statement (3.12). We have found no rigorous proof of the hypothesis, and this interesting question remains open.

#### 4. A differential equation on the value function

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

$$w'(t) + w(t) = \int_0^1 \min\{1 + xt, w(t) + h(t, x)\} dx + \chi(t), \tag{4.1}$$

where  $\chi(t)$  tends to 0 as  $t$  tends to  $\infty$  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 (4.1) 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 (4.1), we need, for all  $t$ , the existence of a stopping rule  $\tau_t^*$  such that

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

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  $0 \leq T_i \leq t$  for some  $i \geq 1$ . Let  $r_i$  be the relative rank of  $X_i$ , and let  $\mathcal{F}_i$  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 by 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 is given 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  $i \geq 1$ , 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 2.1 prove that  $E(i, t)$  is continuous in  $t$ . Hence, the stopping rule  $\tau_i^*$  given by

$$\begin{aligned} \tau_i^* &= i && \text{if } r_i + X_i(t - T_i) \leq E(i, t) \\ \text{and } \tau_i^* &> i && \text{if } r_i + X_i(t - T_i) > E(i, t) \end{aligned}$$

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

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

$$P(T_{\tau_i^*} \geq t) = \lim_{\delta \rightarrow 0} P(T_{\tau_{t-\delta}^*} \geq t - \delta).$$

*Proof.* The probability of there being no arrivals in  $(t - \delta, t)$  tends to 1 as  $\delta \rightarrow 0$  independently of preceding arrivals. Hence, as  $\delta \rightarrow 0$ , a decision maker using either  $\tau_{t-\delta}^*$  or  $\tau_t^*$  will be almost surely confronted with the same set of arrivals before  $t$ . The continuity of  $\Pi(t)$  implies that, as  $\delta$  tends to 0, there exists almost surely a unique optimal limiting rule, and the statement follows.

**Remark 4.1.** Note that, since  $w(t)$  is bounded,  $P(T_{\tau_i^*} \geq t) \rightarrow 0$  as  $t \rightarrow \infty$ .

**Theorem 4.1.** *Let  $\tau_i^*$  be the optimal stopping rule with respect to the horizon  $t$ . The value function  $w(t)$  is differentiable and satisfies*

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

where  $\chi(t) = \Pi'(t) P(T_{\tau_i^*} > t)$  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.

$$w(t \mid x) = \inf_{\tau \in \mathcal{F}} \{E(\tilde{R}_{\tau}^{(t)} \mathbf{1}_{\{T_{\tau} \leq t\}} + \Pi(t) \mathbf{1}_{\{T_{\tau} > t\}} + \mathbf{1}_{\{X_{\tau} \geq x\}} \mathbf{1}_{\{T_{\tau} \leq t\}})\}. \tag{4.3}$$

**4.1. Proof of Theorem 4.1**

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

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

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

$$w(t) = (1 - \delta) E(\tilde{R}_{\tau_i^*}^{(t)} \mid N(\delta) = 0) + \delta E(\tilde{R}_{\tau_i^*}^{(t)} \mid N(\delta) = 1) + o(\delta). \tag{4.4}$$

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

**Lemma 4.1.** *Set*

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

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

$$\lim_{\delta \rightarrow 0^+} \frac{\chi(t, \delta)}{\delta} = \Pi'(t) P(T_{\tau_i^*} > t).$$

*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  $\tau_t^*$  would act on  $(\delta, t)$  under the condition that  $N(\delta) = 0$ . We have

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

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

$$\mathbb{E}(\tilde{R}_{\tilde{\tau}}^{(t)} \mid N(\delta) = 0) - w(t - \delta) \geq \Delta\Pi(t, \delta) \mathbb{P}(T_{\tau_t^*} > t \mid N(\delta) = 0). \tag{4.5}$$

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  $\tau_{t-\delta}^*$  on  $(\delta, t)$ . We have

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

Now we can easily check, from the definitions, that

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

Hence

$$\mathbb{E}(\tilde{R}_{\tau_t^*}^{(t)} \mid N(\delta) = 0) - w(t - \delta) \leq \Delta\Pi(t, \delta) \mathbb{P}(T_{\tau_{t-\delta}^*} > t - \delta) + o(\delta). \tag{4.6}$$

Combining (4.5) and (4.6), we then obtain

$$\Delta\Pi(t, \delta) \mathbb{P}(T_{\tau_t^*} > t \mid N(\delta) = 0) \leq \chi(t, \delta) \leq \Delta\Pi(t, \delta) \mathbb{P}(T_{\tau_{t-\delta}^*} > t - \delta) + o(\delta),$$

and, thus, from Proposition 4.1,

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

Recall the definition of  $w(t \mid x)$  from (4.3). 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  $t$ , we can check that  $w(t \mid x)$  is monotone decreasing in  $x$  on  $[0, 1]$  with

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

The following result holds.

**Lemma 4.2.** *For all  $t > 0$ ,*

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

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

$$E(\tilde{R}_{\tau_t^*}^{(t)} \mid N(\delta) = 1) = \int_0^1 E(\tilde{R}_{\tau_t^*}^{(t)} \mid N(\delta) = 1, X = x) dx, \tag{4.7}$$

where, by the definition of  $\tau_t^*$ ,

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

We can check that, for all  $t$ ,

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

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

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

and

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

We are now in a position to continue the proof of Theorem 4.1.

From Lemma 4.1 we know that  $E(\tilde{R}_{\tau_t^*}^{(t)} \mid N(\delta) = 0) = w(t - \delta) + \chi(t, \delta)$ , so that, after straightforward rearrangements, (4.4) yields

$$\frac{w(t) - w(t - \delta)}{\delta} - \frac{\chi(t, \delta)}{\delta} = -w(t - \delta) + E(\tilde{R}_{\tau_t^*}^{(t)} \mid N(\delta) = 1) + \frac{o(\delta)}{\delta}. \tag{4.8}$$

Now let  $\delta$  go to 0 on both sides of (4.8). We know, from the continuity of  $w(t)$  and Lemma 4.2, that the limit of the right-hand side exists. Therefore, the limit of the left-hand side must also exist. Also, from Lemma 4.1 we know that  $\chi(t) := \lim_{\delta \rightarrow 0} \chi(t, \delta)/\delta$  exists and is finite for all  $t$ . Hence,  $\lim_{\delta \rightarrow 0} (w(t) - w(t - \delta))/\delta$  also exists and, thus,  $w(\cdot)$  must be differentiable on  $\mathbb{R}$ . This completes the proof of Theorem 4.1.

**Remark 4.2.** Set  $h(t, x) = w(t \mid x) - w(t)$ . Equation (4.2) can be rewritten as

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

with  $\chi(t) = \Pi'(t) P(T_{\tau_t^*} > t)$ . This yields (4.1). Also, note that our assumptions on  $\Pi(\cdot)$  imply that  $\Pi'(t)$  is positive and uniformly bounded on  $\mathbb{R}$ . Hence,  $\chi(t) \leq K P(T_{\tau_t^*} > t)$  for some  $K > 0$  and, thus, from Remark 4.1,  $\chi(t) \rightarrow 0$  as  $t \rightarrow \infty$ . For all strategies we can think

of as being close to optimal,  $\chi(t)$  proves to decrease exponentially fast to 0. Hence, we suggest focusing interest on the simpler equation

$$w'(t) + w(t) = \int_0^1 \min\{1 + xt, w(t) + h(t, x)\} dx. \quad (4.9)$$

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

## 5. Conclusion and outlook

We have introduced an alternative version of Robbins' problem in continuous time which we have proved to be well defined. It is shown that this version bears important common features with the original problem, but that it also has specific characteristics of its own. Our main result—and this was our motivation—is that this new problem is asymptotically 'value-equivalent' with the discrete  $n$ -arrival problem, i.e.  $\lim_{n \rightarrow \infty} v(n) = \lim_{t \rightarrow \infty} w(t)$ . We have also established a differential equation which describes  $w(t)$  exactly, and suggested a simplified form of this equation to study the value.

Although we cannot solve the latter 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, (4.9) bypasses the full history dependence which lies at the heart of Robbins' problem. The idea is to substitute estimates of  $h(t, x)$  into (4.9) 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.

## Acknowledgements

The authors are grateful to Freddy Delbaen for several remarks, in particular his idea to combine Hölder's inequalities in the proof of Theorem 3.5 to prove (3.12). We also thank the anonymous referee for his/her insights and stimulating criticism.

## References

- ASSAF, D. AND SAMUEL-CAHN, E. (1996). The secretary problem: minimizing the expected rank with i.i.d. random variables. *Adv. Appl. Prob.* **28**, 828–852.
- BRUSS, F. T. (2005). What is known about Robbins' problem? *J. Appl. Prob.* **42**, 108–120.
- BRUSS, F. T. AND FERGUSON, T. S. (1993). Minimizing the expected rank with full information. *J. Appl. Prob.* **30**, 616–626.
- BRUSS, F. T. AND FERGUSON, T. S. (1996). Half-prophets and Robbins' problem of minimizing the expected rank. In *Athens Conf. Appl. Prob. Time Ser. Anal.* (Lecture Notes Statist. **114**), Vol. 1, Springer, New York, pp. 1–17.
- CHOW, Y. S., MORIGUTI, S., ROBBINS, H. AND SAMUELS, S. M. (1964). Optimal selection based on relative ranks. *Israel J. Math.* **2**, 81–90.
- FELLER, W. (1950). *An Introduction to Probability and its Applications*, Vol. 1, 3rd edn. John Wiley, New York.
- GNEDIN, A. V. (2007). Optimal stopping with rank-dependent loss. *J. Appl. Prob.* **44**, 996–1011.
- KÜHNE, R. AND RÜSCHENDORF, L. (2000). Approximation of optimal stopping problems. *Stoch. Process. Appl.* **90**, 301–325.
- MOSER, L. (1956). On a problem of Cayley. *Scripta Math.* **22**, 289–292.