A continuous time approach to Robbins' problem and open questions

F. Thomas Bruss and Yvik C. Swan

Université Libre de Bruxelles,
Département de Mathématique,
Campus Plaine - CP 210
B-1050 Brussels - Belgium
tbruss@ulb.ac.be, yvswan@ulb.ac.be

Abstract. Let $X_1, X_2, \ldots, X_n$ 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 infinity? This full-information expected selected rank problem is known as Robbins' problem, and its general solution is unknown. In this work 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 problem in this setting, we prove that, under reasonable assumptions, the corresponding value function is bounded, Lipschitz continuous and differentiable. We also derive a differential equation for this function. Our main result is that the limiting value of the Poisson embedded problem exists and is equal to that of Robbins' problem.

Keywords. Optimal stopping, Poisson embedding, Robbins problem, secretary problems.

1 Introduction

This paper summarizes the recent continuous-time approach to Robbins' problem of Bruss and Swan (2009) and discusses open questions related with this approach.

1.1 Robbins' problem

Let $X_1, X_2, \ldots, X_n$ be a sequence of iid 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 $v = \lim_{n\to\infty} v(n)$?

This full-information expected selected rank problem is known 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 problem was solved in 1964 (see Chow *et al.* (1964)), and significant results on the full-information version were obtained in the mid-nineties (see Bruss and Ferguson (1993, 1996) and Assaf and Samuel-Cahn (1996)), some fundamental questions are still open (see the survey paper Bruss (2005) and 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 $\{X_1, \ldots, X_{k-1}\}$. 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 zero to one, and there are infinitely many arrivals and a well-defined smallest one in any strip $[s, s + \Delta s] \times \mathbb{R}^+$. With this model, Gnedin obtains bounds on the stopping value for Robbins' problem and, in particular, shows 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, and 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 chose 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 is defined as its absolute rank, that is the number of observations in $[0,t]$ which are not larger. If no decision has been reached before the given time t, then the loss is equal to some non-negative 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 minimizes his expected loss.

Let $(T_1, X_1), (T_2, X_2), \ldots$ denote the point arrival process. The random variables $T_1 \leq T_2 \leq \ldots$ 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 $X_1, X_2, \ldots$. With this notation, the absolute rank of the kth arrival X_k is defined with respect to t as

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

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

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

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 $\{T_r\}_{r\geq 1}$ of arrival times of the point process, which satisfy $\{\tau \leq s\} \in \mathcal{F}_s$, where

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

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

Remark 1. We will here use the notation $\{\tau = k\}$ instead of $\{\tau = T_k\}$ to denote the event that the decision maker selects the kth arrival. Hence the notations R_τ, X_τ and T_τ are well defined and will be used systematically throughout.

The Poisson embedded Robbins' problem consists in studying the value function

$$w(t) = \inf_\tau E\left(\tilde{R}_\tau^{(t)}\right) = \inf_\tau E\left(R_\tau^{(t)} 1_{\{T_\tau \leq t\}} + \Pi(t) 1_{\{T_\tau > t\}}\right), \tag{2}$$

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 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 $\Pi(.)$ is increasing and differentiable with bounded derivative. Hence this function is Lipschitz-continuous and satisfies

$$\lim_{t\to\infty} \frac{\Pi(t)}{t} \leq \kappa, \tag{3}$$

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

Remark 3. 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 $r_k = \sum_{j=1}^{k} 1_{\{X_j \le X_k\}}$ 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).

Remark 4. For all $\tau \in \mathcal{T}$, let $w_\tau(t) = E(\tilde{R}_\tau^{(t)})$ denote the expected loss incurred by using τ. Then, by definition of $w(t)$,

$$0 \le w(t) \le w_\tau(t)$$

for any specific choice of stopping rule τ. Also note that $w(t) \ge 1$ for all t such that $\Pi(t) \ge 1$.

2 Properties of the value function

Proposition 1. *The value function of the Poisson embedded Robbins' Problem is bounded on* $\mathbb{R}$ *and satisfies*

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

for all t sufficiently large.

Outline of the proof: This result is readily obtained by considering a specific subclass of stopping rules of the form $\tau = \inf\{i \ge 1 \text{ such that } X_i \le \varphi(T_i)\}$, for $\varphi(\cdot)$ some real valued function defined on $\mathbb{R}^+$. Such rules are called *memoryless threshold rules* 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 by Bruss and Ferguson (1993, 1996) and Assaf and Samuel-Cahn (1996), and in a continuous time limit model by Gnedin (2007).

Straightforward computations of the value of these rules show that, for sufficiently well-behaved threshold functions $\varphi(\cdot)$,

$$\begin{aligned} w_\tau(t) = 1 + (\Pi(t) - 1)\, e^{-\mu(t)} + \tfrac{1}{2}\int_0^t \varphi(s)^2 (t-s) e^{-\mu(s)} ds \\ + \tfrac{1}{2}\int_0^t \int_0^s \frac{(\varphi(s) - \varphi(u))^2}{1 - \varphi(u)} du e^{-\mu(s)} ds, \end{aligned} \tag{4}$$

where $\mu(s) = \int_0^s \varphi(u)du,\ s \le t$. This result concurs with that obtained by Assaf and Samuel-Cahn (1996) for the corresponding rules in the discrete settings. Now choose

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

For all $c > 1$, the corresponding memoryless threshold rule is well defined and satisfies $P(T_\tau < \infty) = 1$. Writing out (4) for this choice of $\varphi(\cdot)$, one sees that $w_\tau(\cdot)$ is increasing and converges to the limit

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

This last expression is minimal in $c = 1.9469...$ and the corresponding limiting value is $2.33182...$. Hence the value function of the Poisson embedded Robbins' problem is bounded on $\mathbb{R}$ and satisfies

$$1 \le w(t) \le 2.33183 \tag{5}$$

for all t sufficiently large.

Remark 5. These values (the upper bound and the minimal c achieving this bound) are identical to those obtained by Bruss and Ferguson (1993) and Assaf and Samuel-Cahn (1996) for the discrete problem. They also coincide with those obtained by Gnedin (2007).

Theorem 1. *$w(t)$ is continuous on $\mathbb{R}$ and Lipschitz continuous on (t_0, ∞), for some t_0 sufficiently large.*

Outline of the proof: The proof of this result relies on upper and lower bounds on the increments $w(t+\delta) - w(t)$ for small δ and large t, namely

$$-3\delta(\delta+1) \leq w(t+\delta) - w(t) \leq L\delta,$$

where L is some positive constant independent of t and δ.

Theorem 2. *Let $\tau_t^\star$ be the optimal strategy with respect to the horizon t. The value function $w(t)$ is differentiable and satisfies*

$$w'(t) + w(t) = \int_0^1 \min\{1 + xt, w(t \mid x)\} dx + \chi(t), \tag{6}$$

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

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

Outline of the proof: Though (6) is relatively easy to derive heuristically, obtaining an exact equation proves to be a lot of work. We first use the continuity of $w(t)$ to justify, for all t, the existence of an optimal stopping rule $\tau_t^\star$ such that $w(t) = w_{\tau_t^\star}(t)$. We can then write

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

Next we prove that

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

and

$$\lim_{\delta \to 0^+} \frac{E(\tilde{R}_{\tau_t^\star}^{(t)} \mid N(\delta) = 0) - w(t-\delta)}{\delta} = \Pi'(t) P(T_{\tau_t^\star} > t).$$

The statement then follows.

3 Existence of the asymptotic value

Our main result is that the classical and the continuous-time Robbins' problems are asymptotically equivalent, in that their limiting values exist and are equal. The proof of this result relies on the following three statements.

Proposition 2. $\lim_{t \to \infty} w(t) \geq v$.

Outline of the proof: Consider a "half-prophet" Q, who is asked to solve the Poisson embedded Robbins' problem but is allowed to know in advance the number of observations. Let $w_Q(t)$ be her value. Then clearly $w(t) \geq w_Q(t)$. Moreover, it is easy to see that

$$w_Q(t) \geq \sum_{k=0}^{\infty} P(N(t) = k) \inf_\sigma E(\tilde{R}_\sigma^{(t)} \mid N(t) = k).$$

Now, since Q knows in advance the number of arrivals, she is in a position to use the optimal stopping rule for the discrete problem so that her optimal value, $\inf_\sigma E(\tilde{R}_\sigma^{(t)} \mid N(t) = k)$, can be expressed in terms of $v(k)$. After dealing with the impact of the penalty function, the claim follows.

Proposition 3. *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*

$$\lim_{t\to\infty} w(t) \le v + \lim_{t\to\infty} t^\alpha E(X_{\tau^\star_{\beta_t}}).$$

Outline of the proof: Our approach consists in studying how well the optimal rule $\tau^\star_{\beta_t}$ fares in the continuous setting. Straightforward conditioning yields

$$w_{\tau^\star_{\beta_t}}(t) = w_{\tau^\star_{\beta_t}}(t \mid N(t) < \beta_t)P(N(t) < \beta_t) + w_{\tau^\star_{\beta_t}}(t \mid N(t) \ge \beta_t)P(N(t) \ge \beta_t), \quad (7)$$

where $w_{\tau^\star_{\beta_t}}(t \mid E)$ denotes the expected loss under $\tau^\star_{\beta_t}$ conditioned on the event E. First note that the first term of the rhs of (7) is uniformly bounded by $\Pi(t)$. Hence our hypothesis on the penalty function suffice to guarantee that it becomes arbitrarily small as $t \to \infty$. Next we prove that

$$w_{\tau^\star_{\beta_t}}(t \mid N(t) \ge \beta_t) = v(\beta_t) + E(X_{\tau^\star_{\beta_t}})\big(E(N(t) \mid N(t) \ge \beta_t) - \beta_t\big). \quad (8)$$

A large deviations argument for the second term of (8) then yields the claim.

Theorem 3. *Let $\tau^\star_n$ be the optimal strategy for the discrete n-arrival Robbins' problem. Then, for all $p > 1$,*

$$E(X_{\tau^\star_n}) \le E(R_{\tau^\star_n}) \left(\sum_{k=1}^n a^p_{n,k}\right)^{\frac{1}{p}},$$

where

$$a^p_{n,k} = k^{2-p}\binom{n}{k}\int_0^1 x^{p+k-1}(1-x)^{n-k}dx. \quad (9)$$

Outline of the proof: Conditioning on the ranks we obtain $E(X_{\tau^\star_n}) = E(E(X_{\tau^\star_n} \mid R_{\tau^\star_n})) = \sum_{k=1}^n E(X_{(k)}\mathbf{1}_{\{R_{\tau^\star_n}=k\}})$, where $X_{(k)}$ is the kth smallest order statistic of the sample $X_1, \dots, X_n$. The expectations cannot be factorized. Nevertheless, applying Hölder's inequality *twice* (an idea we owe to F. Delbaen) yields

$$E(X_{\tau^\star_n}) \le \left(\sum_{k=1}^n a^p_{n,k}\right)^{1/p} E(R_{\tau^\star_n}),$$

where $a_{n,k} = k^{-1/q}(E(X^p_{(k)}))^{1/p}$. Theorem 3 follows.

For fixed p, the coefficients (9) can be shown to behave like $n^{2/p-1}$ when n is large. Hence, for each α, one can find p such that $n^\alpha n^{2/p-1} \to 0$, and thus $\lim_{t\to\infty} t^\alpha E(X_{\tau^\star_{\beta_t}}) = 0$. This yields the following.

Theorem 4. *The limiting value for the Poisson embedded Robbins' problem exists and satisfies*

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

4 Open questions

In this final section, we state a number of interesting questions which arose during our exploration of the mysteries Robbins' problem.

Question 1: Is $w(t)$ increasing?

Bruss and Ferguson (1993) proved, through a crafty "half-prophet" argument, that $v(n)$ is increasing in n. Hence, although the value function $w(t)$ need not be increasing for small values of t (because of the penalty function), it seems reasonable that it should, for an increasing penalty function, have the same behavior as its discrete counterpart. Interestingly, the "half-prophet" trick fails to hold in the continuous setting, and we have found no alternative argument to substantiate the statement that $w(t)$ is increasing.

Question 2: How useful is the integro-differential equation (6)?

Obtaining bounds on v is an important part of contributing to a solution of Robbins' problem, and the best bounds known so far are

$$1.908 < v < 2.327,$$

(see Assaf and Samuel-Cahn (1996) and Bruss and Ferguson (1996)). Now although (6) is not a differential equation in the usual sense, it does open the way for an improvement on those bounds. For instance, it it is easy to show from (6) that, for all t sufficiently large,

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

where

$$H(t,c) = \int_0^{c/t+c} (1+xt) \mathrm{dx} + \int_{c/t+c}^1 (w(t \,|\, x) - w(t)) \mathrm{dx}.$$

Hence estimates on $h(t,x) = w(t \,|\, x) - w(t)$ yield estimates on w, and thus on v. We believe that herein lies hope for improvement on the known bounds on v although, in the present state of our research, this remains a mere shift in the difficulty. We have so far obtained no significant improvement on the bounds given above.

Question 3: Is $nE(X_{\tau_n^\star})$ bounded?

Theorem 3 implies, in particular, that the n-optimal strategy $\tau_n^\star$ satisfies $\lim_{n\to\infty} n^\alpha E(X_{\tau_n^\star}) = 0$ for all $0 \leq \alpha < 1$. This fails for $\alpha = 1$. In fact, the estimate (3) does not even suffice to prove that $nE(X_{\tau_n^\star})$ is bounded. Now a natural parallel has been drawn between Robbins' problem and *Moser's problem* (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 one as $n \to \infty$ (Bruss and Ferguson (1993)), it is then natural to believe that the optimal strategies in both problems should have similar behaviors, at least asymptotically. Since the optimal strategy $\hat{\tau}_n$ for Moser's problem satisfies $\lim_{n\to\infty} (nE(X_{\hat{\tau}_n})) = 2$, it seems intuitive that $nE(X_{\tau_n^\star})$ is – at the least – bounded. We have found no proof of this assertion, and call it the Assaf–Samuel-Cahn conjecture.

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, *Springer Lecture Notes in Stat.* **114**, Vol. 1 in honor of J.M. Gani, 1–17.

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.

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

Gnedin, A. V. (2007). Optimal Stopping with Rank-Dependent Loss, *J. Appl. Prob.*, **44**, 4, 996–1011.

Kühne, R. and Rüschendorf, L. (2000). Approximation of optimal stopping problems, *Stoch. Proc. Appl.*, **90**, 301–325.

Moser, L. (1956). On a problem of Cayley, *Scripta Math.* **22**, 289–292.