An Optimal Stopping Rule for Approaching a Border that Should not be Crossed

Andrzej Z. Grzybowski, *Member, IAENG*

***Abstract*—We consider an optimal stopping problem with some of the features of the blackjack type games. Let X_1, $X_2,\ldots,X_N$ be a finite sequence of nonnegative random variables. A decision maker observes sequentially the values and decides whether to stop or to continue. If he decides to stop at the moment k he obtains a payoff dependent on the sum $X_1+\ldots+X_k$. The greater the sum, the more the decision maker gains, unless the sum exceeds a positive number T – a limit given in the problem. If so, the decision maker loses all or part of his payoff. A special case of such a problem is considered in details. In this case a simple optimal stopping rule is found. Some examples and practical questions are discussed as well.**

***Index Terms*— Markov chain, optimal stopping rule, sequential decision making.**

I. Introduction

There are some problems in real world applications where the decision maker wants to get as close as possible to a given limit but the limit should be crossed under the threat of some kind of punishment. Especially interesting are those in which the limit is approached by random moves (steps). Some of the problems can be modeled as follows. Let $X_1, X_2, \ldots, X_N$ be a finite sequence of independent nonnegative random variables. A decision maker observes sequentially the values of the variables and decides whether to stop or to continue. If he decides to stop at the moment k he gains a value $W(\sum_{i=1}^{k} X_i)$, where $W: R_+ \to R$ is a given function. The function is positive and increasing on the interval $(0, T]$ and is decreasing for arguments greater than T. It implies that W achieves its only maximum for $\sum_{i=1}^{k} X_i = T$.

Such a problem can be a model for various real world situations which can be observed in economics, finance, politics or social life. We describe just two of them.

The first example is *service with work time limit*.

A decision maker controls a mechanism (or an individual) which should not work longer than a given time period T. He has several jobs to process in sequential order with the ith job requiring a random time X_i for its execution. After each job he must decide whether to start next one or to stop. Every initiated job must be completed. The longer the mechanism works, the more the decision maker gains but if the work time exceeds the limit T the decision maker will be punished.

The second problem is *blackjack type games*.

The games are played on a points system that gives numeric values to every card in a single deck of playing cards. The cards are given to a player sequentially until he decides to stop. His score is the sum of the values in hand. The player with the highest total score wins as long as it doesn't exceed a given limit number. If a player's cards exceed the limit then the player loses and his/her bet is taken by the dealer. Such games were considered in [3]

One specific model of such a situation will be considered in detail in the sequel.

II. Problem Statement

Let $X_1, X_2, \ldots, X_N$ be a sequence of i.i.d. random variables having an exponential distribution with the density function:

$$f(t) = \lambda \exp(-\lambda t)\mathbf{1}_{[0,\infty)}(t), \quad \lambda > 0 \tag{1}$$

The decision maker observes the sequence $X_1, X_2, \ldots, X_N$ and at each stage decides whether to stop or to continue. If he decides to stop at the moment k his gain depends on $Y_k = \sum_{i=1}^{k} X_i$. The payoff function W is given by the following equation:

$$W(Y_k) = \begin{cases} B \cdot Y_k, & Y_k \le T \\ B \cdot T - F(Y_k - T), & Y_k > T \end{cases} \tag{2}$$

with B>0, T>0. The *punishment* function F appearing in (2) has the form $F(x) = K \cdot x + C$, with $K \ge 0, C \ge 0$, and $K+C > 0$.

Our tusk is to find an optimal stopping rule for this problem i.e. stopping rule which maximizes the expected payoff for a decision maker

III. Some General Definitions and Results

Before we solve our optimal stopping problem we need to present some necessary formal definitions and fundamental results from the theory of optimal stopping. They can be found e.g. in [1], [2], [4].

Let $X_1, X_2, \ldots$ be a sequence of independent random variables. Let F_n denote the σ-algebra generated by the random variables $X_1, X_2, \ldots, X_n$ in an underlying probability space $(\Omega,\mathsf{F},\mathsf{P})$ *A stopping rule* is a random variable τ with values in a set of natural numbers such that $\{\tau=n\} \in \mathsf{F}_n$ for n=1,2,… and $\mathsf{P}(\tau<\infty)$=1. Let $\mathsf{M}(n,N)$ be a class of all stop-

Manuscript received February 27, 2009.
A. Z. Grzybowski is with the Czestochowa University of Technology, 42-200 Czestochowa, Poland (e-mail: azgrzybowski@ gmail.com).

ping rules τ such that $P(n \le \tau \le N)=1$. The class $\mathsf{M}(1,N)$ will be denoted $\mathsf{M}(N)$

Let (Y_n,F_n), $n=1,2,\ldots$, be a homogenous Markov chain with values in a state space (Υ,B). Let $W: R_+ \to R$ be a Borel measurable function which values $W(y)$ will be interpreted as the gain for a player when he stops the chain (Y_n,F_n) at the state y. Assume that for a given state y and for a given stopping rule τ the expectation $E(W(Y_\tau)|Y_1=y)$ exists. Then it is natural to interpret the value - denoted by $E_yW(Y_\tau)$ - as the mean gain corresponding to a chosen stopping rule τ.

Let us define a function V_N by the equation:

$$V_N(y) = \sup_{\tau\in \mathsf{M}_W(N)} E_yW(Y_\tau) \tag{3}$$

where $\mathsf{M}_W(N)$ is a set of all stopping rules belonging to $\mathsf{M}(N)$ for which the expectations $E_yW(Y_\tau)$ exist for all $y\in \Upsilon$ and are larger than $-\infty$. The value $V_N(y)$ is called a *value* of the problem of optimal stopping when the *initial state* of the process is y.

A stopping rule $\tau^*\in \mathsf{M}_W(N)$ which for all $y\in\Upsilon$ satisfies the condition

$$E_yW(Y_{\tau*})= V_N(y) \tag{4}$$

is called an *optimal stopping rule*.

Let B denote a class of *all* Borel measurable functions W for which the expectations $E_yW(Y_2)$ exist for all $y\in\Upsilon$. Let us define an operator Q operating on functions $W \in \mathsf{B}$ by

$$QW(y) = \max\{W(y), E_yW(Y_2)\} \tag{5}$$

The following theorem, which can be found in [4], provides us with the solution to the optimal stopping problem in the considered case.

Theorem Assume that $W \in \mathsf{B}$. Then:

i. $V_n(y) = Q^nW(y)$, $n=1,2,\ldots$

ii. $V_n(y) = \max\{W(y), E_yV_{n-1}(Y_2)\}$, where $V_0(y) = W(y)$

iii. A stopping rule τ_n^* defined by

$$\tau_n^* = \min\{1 \le k \le n : V_{n-k}(Y_k) = W(Y_k)\}$$

is an optimal stopping rule in a class $\mathsf{M}_W(n)$

If $E_y|W(Y_k)| < \infty$, for $k=1,\ldots,n$, then the stopping rule τ_n^* is optimal in the class $\mathsf{M}(n)$

IV. Problem Solution

It is easy to see that our problem is a special case of the above general problem. So, in order to solve it we apply the Theorem.

First we need to find the form of $V_n(y)=Q^n(y)$, $n=1,2,\ldots N$. For every $y\in(0,T]$, by (1),(2) and (5), we have

$$\begin{aligned} V_1(y) &= QW(y) = \max\{W(y), E_yW(Y_2)\} = \\ &\max\{W(y), E_yW(y+X_2)\} = \max\{W(y), I_1(y)\} \end{aligned} \tag{6}$$

with

$$\begin{aligned} I_1(y) &= \int_0^\infty W(y+x)f(x)dx = B\int_0^{T-y}(y+x)\lambda\exp(-\lambda x)dx + \\ &\int_{T-y}^{\infty}[B\cdot T - K\cdot(y+x-T) - C]\cdot\lambda\exp(-\lambda x)dx = \\ &\frac{B}{\lambda}(y\lambda - \exp(\lambda(y-T))+1) - \frac{1}{\lambda}(K+C\lambda)\exp(\lambda(y-T)) \end{aligned} \tag{7}$$

It is easy to verify, that (for any given parameters T, λ, B, K, C characterizing our problem) a point t_1 for which the functions W and I_1 have equal values is exactly the same as the point at which the function I_1 takes its only maximum on the interval $(0,T]$. Moreover, the following conditions hold:

$I_1(y) > W(y)$ for $y\in(0, t_1)$ and $I_1(y) < W(y)$ for $y\in(t_1,T]$

The value of t_1 is given by the formula:

$$t_1 = T - \frac{1}{\lambda}\ln\left(\frac{B+K+C\lambda}{B}\right) \tag{8}$$

It also follows from (2) that $I_1(y)\le W(y)$ for $y > T$. In view of Theorem 1 it implies that one step before the end of observations the decision maker should continue the observations if he is at any state y which is less than t_1 and should stop otherwise.

Let $I_n(y)$ denote the expectation $E_yV_{n-1}(y+X_2)$, $n=2,\ldots,N$. Now, with the help of mathematical induction, we show that the following lemma is true

Lemma. Let t_1 be given by the formula (8). Then for any natural number n the function I_n satisfies the following conditions:

i. $I_n(y) > W(y)$ for $y\in(0, t_1)$

ii. $I_n(y) \le W(y)$ for $y\in[t_1,\infty)$

Proof. It was already shown that conditions *i* and *ii* hold for $n=1$. Now let us assume that the conditions hold for I_{n-1}. Then, by the definition of V_{n-1} and the induction assumption, for $y\in(0, t_1)$ we have:

$$\begin{aligned} I_n(y) &= \int_0^\infty V_{n-1}(y+x)f(x)dx = \\ &\int_0^{t_1-y} I_{n-1}(y+x)f(x)dx + \int_{t_1-y}^\infty W(y+x)f(x)dx > \\ &\int_0^\infty W(y+x)f(x)dx = I_1(y) > W(y) \end{aligned} \tag{9}$$

It implies that condition *i.* is satisfied.

Again by the induction assumption, when $y\in[t_1,\infty)$ we obtain:

$$\begin{aligned} I_n(y) &= \int_0^\infty V_{n-1}(y+x)f(x)dx = \\ &\int_0^\infty W(y+x)f(x)dx = I_1(y) \le W(y) \end{aligned} \tag{10}$$

Thus the condition *ii.* also holds and the proof of the lemma is completed.

It follows from the lemma immediately that for $n=1,\ldots,N$ functions V_n have the form:

$$V_n(y) = I_n(y)\cdot\mathbf{1}_{[0,t_1)}(y) + W(y)\cdot\mathbf{1}_{[t_1,\infty)}(y) \tag{11}$$

where t_1 is given by (8).

The following proposition provides us with the solution of

ISBN:978-988-18210-1-0

our problem.

Proposition. Let us consider a sequence $X_1, X_2, \dots, X_N$ of i.i.d. random variables with density functions given by (1). The optimal stopping rule for the problem of optimal stopping of the Markov chain ($Y_n = \sum_{i=1}^{n} X_i$, $\mathcal{F}_n$) with the gain function (2) is given by:

$$\tau_N^* = \min\{1 \le k \le N : Y_k \ge t_1\}$$

with t_1 being given by the formula (8)

The value $V_N(y)$ of the problem can be calculated for $y < t_1$ with the help of the following recursive equation:

$$V_n(y) = \int_0^{t_1 - y} V_{n-1}(y+x)\lambda \exp(-\lambda x)dx + \frac{1}{\lambda}\exp(\lambda(y-T))(B+K+C\lambda)(T\lambda - \ln(\frac{B+K+C\lambda}{B})) \quad (11)$$

$n=2,\dots,N$,

with the initial condition $V_1(y) = I_1(y)$ given by (7).

We omit the proof of the Proposition because it directly follows Theorem 1 and the formula (11)

The results stated in the Proposition imply that at any moment k the decision maker should continue the observations if he is at any state y which is less than t_1 and should stop otherwise. Such a stopping rule maximizes his expected payoff. The maximum that the decision maker may expect to gain is $V_N(0)$

V. Some Examples and Final Remarks

We can see that the recursive equation (11) involves integrating and one cannot be happy about it. We may notice however, that for any natural number n the functions V_n can be expressed in terms of elementary functions (involving e^x, $\ln x$, x^n) though the calculations are rather arduous, even for small numbers n. Fortunately, one may use a computer with symbolic manipulation software such as *Mathematica*, *Maple, Maxima, Axiom*, etc., to obtain the form of the functions. We applied *Mathematica* 4.0 software to compute the form of the functions V_n, $n=1,\dots,15$. As an example, we present here *Mathematica* output for the function V_5 (when n is greater the formulae for V_n are too large to paste into the text). In the presented below output variable y and constants T, λ, K, B, C have their previous meaning.

Mathematica output for the function V_5:

$$(e^{-T\lambda}(-e^{y\lambda}(K+C\lambda)(120 - 96y\lambda + 36y^2\lambda^2 - 8y^3\lambda^3 + T^4\lambda^4 + y^4\lambda^4 - 4T^3\lambda^3(-2+y\lambda) + 6T^2\lambda^2(6 - 4y\lambda + y^2\lambda^2) - 4T\lambda(-24 + 18y\lambda - 6y^2\lambda^2 - T^3\lambda^3 + y^3\lambda^3)) + B(24e^{T\lambda}(5+y\lambda) - e^{y\lambda}(120 - 96y\lambda + 36y^2\lambda^2 - 8y^3\lambda^3 + T^4\lambda^4 + y^4\lambda^4 - 4T^3\lambda^3(-2+y\lambda) + 6T^2\lambda^2(6 - 4y\lambda + y^2\lambda^2) - 4T\lambda((-24 + 18y\lambda - 6y^2\lambda^2 + y^3\lambda^3))) - 4e^{y\lambda}(B+K+C\lambda)(-24 + 18y\lambda - 6y^2\lambda^2 - T^3\lambda^3 + y^3\lambda^3 + 3T^2\lambda^2(-2+y\lambda) - 3T\lambda(6 - 4y\lambda + y^2\lambda^2)) \ln((B+K+C\lambda)/B) - 6e^{y\lambda}(B+K+C\lambda)(6 - 4y\lambda + T^2\lambda^2 + y^2\lambda^2 - 2T\lambda(-2+y\lambda)) \ln^2((B+K+C\lambda)/B) - 4e^{y\lambda}((B+K+C\lambda)(-2 - T\lambda + y\lambda) \ln^3((B+K+C\lambda)/B) - e^{y\lambda}((B+K+C\lambda) \ln^4((B+K+C\lambda)/B))) / 24\lambda$$

With the help of such formulae we can compute an expected payoff for a decision maker and analyze dependence between the payoff and the parameters characterizing the problem. For example let us consider the case where $T=10$ and $\lambda=B=K=C=1$. Then $t_1=10-\ln(3)$ which means that in the case the decision maker should continue his observations until the sum of already observed values exceeds $10-\ln(3) \approx 8.901$. If the initial state y equals 0 by applying this stopping rule he may expect to win (in average) about $V_5(0) \approx 4.911$ if he has got five steps to the end of observations, about $V_{10}(0) \approx 8.168$ if he has 10 observations ahead, and $V_{15}(0) \approx 8.863$ if he has 15 observations before the end of the process. No other stopping rule can guarantee the decision maker as much.

In our paper we consider the problem of optimal stopping where the decision maker *must* make at least one observation and the initial state for the process is 0 (the decision maker's gain is 0 if he does not make any observation) But we can also consider the situation where he can stop the process without any observation – it would be justified e.g. in the case where t_1 is less than 0. For example if $T=10$ and $B=K=C=1$ it would be the case where $\lambda < \lambda_{min} \approx 0.07289$ or, in other words, where $EX_1 > 1/\lambda_{min} \approx 13.7184$ (i.e. an average length of the step while approaching the limit is greater than 13.7184) In such a case the value of the problem is negative e.g. if $N=5$ and $\lambda = 0.072$ then $V_5(0) = -0.393552$, and it would be better for the decision maker to stop without any observations.

Fig. 1 shows the expected payoff for the decision maker as a function of λ in case where $N=5$.

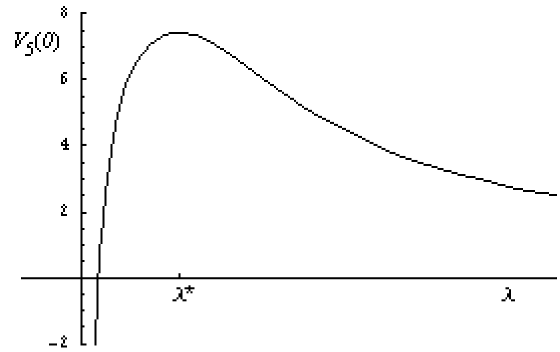

Fig.1. The value $V_5(0)$ as a function of λ when $T=10$ and $B=K=C=1$.

One can verify easily, that in our example the value of λ for which the expected payoff achieves its maximum equals $\lambda^* = 0.411868$. Such knowledge can be important if we have influence on the value of λ (or, equivalently, on the average length of the step while approaching the limit). If so, we would choose the value λ^* to achieve the greatest gain.

References

[1] Y. S. Chow, H. Robbins, D. Siegmund, *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin, Boston, 1971

[2] T. S. Ferguson, *Optimal Stopping and Applications*., unpublished manuscript, 2000, Available: http://www.math.ucla.edu/~tom/Stopping/Contents.html

[3] A. Z. Grzybowski, "On some blackjack type optimal stopping problem", *Scientific Research of the Institute of Mathematics and Computer Science,* Czestochowa University of Technology Publishing, 1(7), Częstochowa 2008

[4] A. N. Shiryayev, *Optimal stopping rules*, Springer-Verlag, New York, 1978