stichting
mathematisch
centrum

AFDELING MATHEMATISCHE STATISTIEK SW 25/73 DECEMBER

A. HORDIJK & K.M. Van HEE

A SEQUENTIAL SAMPLING PROBLEM SOLVED BY OPTIMAL STOPPING

Prepublication

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS (MOS) subject classification scheme (1970): 60G40, 62L10, 62L15, 62C10, 90C40

A sequential sampling problem solved by optimal stopping*)

K.M. van Hee†) & A. Hordijk

Abstract

This paper investigates a sequential decision problem with a simple hypothesis against a simple alternative, concerning the probability of success of a sequence of independent Bernoulli trials.
It is shown that the minimax decision rule is equivalent to the optimal stopping rule of a Markov chain. An explicit solution of the stopping problem is given. Moreover, a finite algorithm for computing the optimal stopping time is provided.

*) This paper is not for review; it is meant for publication.

†) Technische Hogeschool Eindhoven, Onderafdeling der Wiskunde, Eindhoven, Netherlands.

1. *Introduction*

We consider a sequential decision problem *) with a simple hypothesis against a simple alternative, concerning the probability of success, t, of a sequence of independent Bernoulli trials. The hypotheses are H_0: $t = \theta$ and H_1: $t = 1 - \theta$, with $0 \le \theta \le 1$ and θ is known. Each trial costs an amount c. When we stop sampling we must decide. If we make a correct decision we receive a reward r, otherwise we loose ℓ $(r \ge \ell \ge 0, c > 0)$.

We shall show that the minimax decision rule for this problem is equivalent to the optimal stopping rule of a Markov chain. We provide an explicit solution for the stopping problem.

2. *Sequential decision problem*

Suppose that the variable t is a random variable (represented by $\underline{t}$) with distribution defined by

$$P(\{\underline{t}=\theta\}) = 1 - P(\{\underline{t}=1-\theta\}) = \pi, \qquad 0 < \pi < 1.$$

We introduce a sequence of identically distributed random variables $\underline{x}_1, \underline{x}_2, \underline{x}_3, \ldots$ with $\underline{x}_i = 1$ if the i-th trial is a success and $\underline{x}_i = 0$ otherwise.

The random variables are, given $\underline{t}$, independently distributed.

The posterior probability of $\underline{t}$ given $\underline{x}_1 = x_1, \ldots, \underline{x}_n = x_n$ is

$$(1) \qquad y_n := P(\underline{t}=\theta \mid \underline{x}_1=x_1, \ldots, \underline{x}_n=x_n) =$$

$$= \frac{\pi \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}}{\pi \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} + (1-\pi)\prod_{i=1}^{n}(1-\theta)^{x_i}\theta^{1-x_i}} \quad \text{for } n \in \{1,2,3,\ldots\}$$

where $x_i = 0$ or 1. Note that

---

*) This problem is a simplification of a question posed by a psychologist, investigating human decision processes.

(2) $$y_n = \frac{\theta^{x_n}(1-\theta)^{1-x_n} y_{n-1}}{\theta^{x_n}(1-\theta)^{1-x_n} y_{n-1} + \theta^{1-x_n}(1-\theta)^{x_n}(1-y_{n-1})} \quad \text{for } n \in \{1,2,3,\ldots\}.$$

We define recursively the sequence of random variables $\underline{y}_0, \underline{y}_1, \underline{y}_2, \ldots$ by $P(\{\underline{y}_0=\pi\}) = 1$

(3) $$\underline{y}_n = \frac{\theta^{\underline{x}_n}(1-\theta)^{1-\underline{x}_n} \underline{y}_{n-1}}{\theta^{\underline{x}_n}(1-\theta)^{1-\underline{x}_n} \underline{y}_{n-1} + \theta^{(1-\underline{x}_n)}(1-\theta)^{\underline{x}_n}(1-\underline{y}_{n-1})} \quad \text{for } n \in \{1,2,3,\ldots\}.$$

We shall formulate now the *sequential decision problem:*

$\Theta = \{\theta, 1-\theta\}$ *is the parameter set,* $A = \Theta$ *the action space.* L *is the loss function defined on* $\Theta \times A$ *by*

$$L(\theta,\theta) = L(1-\theta,1-\theta) = -r \textit{ and } L(1-\theta,\theta) = L(\theta,1-\theta) = \ell.$$

ψ, *a stopping rule, is defined by* $\psi = (\psi_0, \psi_1(x_1), \psi_2(x_1,x_2), \ldots)$ *(where* x_i *is a realization of* $\underline{x}_i$, $i = 1,2,3,\ldots$*) with* $\psi_i(x_1,\ldots,x_i) = 0$ *or* 1 *and if* $\psi_i(x_1,\ldots x_i) = 1$ *then* $\psi_n(x_1,\ldots x_n) = 1$ *for all* $n \geq i$.
The random variable $\underline{\tau}$ *defined on the sample space of* $\underline{x}_1, \underline{x}_2, \underline{x}_3, \ldots$ *by* $\underline{\tau} = n$ *if* $n = \min\{i \mid \psi_i(\underline{x}_1,\ldots,\underline{x}_i) = 1\}$ *is called a stopping time.*
δ *is an action rule if* $\delta = (\delta_0, \delta_1(x_1), \delta_2(x_1,x_2), \ldots)$ *where* δ_i *is a function with values in* A.

A sequential decision rule is a pair (ψ,δ).

The risk, with respect to the prior distribution represented by π *of a sequential decision rule* (ψ,δ) *is:*

(4) $$r(\pi,(\psi,\delta)) = \pi \cdot E_{\theta,\psi}\{L(\theta,\delta_{\underline{\tau}}(\underline{x}_1,\ldots,\underline{x}_{\underline{\tau}})) + \underline{\tau}c\} + $$
$$+ (1-\pi)E_{1-\theta,\psi}\{L(1-\theta,\delta_{\underline{\tau}}(\underline{x}_1,\ldots\underline{x}_{\underline{\tau}})) + \underline{\tau}c\}.$$

The Bayes risk with respect to π *is defined as the infimum over all* (ψ,δ). *The pair* (ψ_0,δ_0) *for which the risk equals the infimum is called the Bayes rule with respect to* π.

There exists a one-one correspondence between the random variables $\underline{y}_n$ and $\underline{z}_n$, hence the sequence $\underline{z}_1, \underline{z}_2, \underline{z}_3, \ldots$ forms also a Markov chain. We use the following notation

$$(11) \qquad \beta = \frac{1-\pi}{\pi} \quad \text{and} \quad \alpha = \frac{\theta}{1-\theta} .$$

Without loss of generality we shall suppose $0 < \alpha < 1$. In the cases where $\alpha = 0$ or $\alpha = 1$ the optimal decision rules are trivial. If $\alpha = 0$ we are certain after one trial, if $\alpha = 1$ the hypotheses are identically.

The transition probabilities of the Markov chain $\underline{z}_1, \underline{z}_2, \underline{z}_3, \ldots$ are easily derived from (6) by the transformation (use (9), (10) and (11))

$$(12) \qquad y = \frac{1}{1+\beta\alpha^{-z}} \qquad \text{for } z \in \mathbb{Z}$$

where $\mathbb{Z}$ is the set of integers. The *transition probabilities* are

$$(13) \qquad p^{\beta}_{z,z+1} := P(\underline{z}_{n+1}=z+1 \mid \underline{z}_n=z) = \frac{1}{\alpha+1}\,\frac{\alpha^{z+1}+\beta}{\alpha^{z}+\beta} \qquad \text{for } z \in \mathbb{Z}$$

$$p^{\beta}_{z,z-1} := P(\underline{z}_{n+1}=z-1 \mid \underline{z}_n=z) = \frac{\alpha}{\alpha+1}\,\frac{\alpha^{z-1}+\beta}{\alpha^{z}+\beta} \qquad \text{for } z \in \mathbb{Z}.$$

Note that

$$p^{\beta}_{z,z+1} + p^{\beta}_{z,z-1} = 1.$$

For each state $y \in [0,1]$ of the state space of the chain $\underline{y}_0, \underline{y}_1, \underline{y}_2, \ldots$ we define a *reward*

$$(14) \qquad s(y) = \max\{ry - \ell(1-y), r(1-y) - \ell y\}$$

Obviously

$$s(y) = \begin{cases} ry - \ell(1-y) & \text{if } y \geq \frac{1}{2} \\ r(1-y) - \ell y & \text{otherwise.} \end{cases}$$

For the transformed Markov chain $\underline{z}_1, \underline{z}_2, \underline{z}_3, \ldots$ we find

$$(15) \qquad s^{\beta}(z) = \begin{cases} (r+\ell)\dfrac{1}{1+\beta\alpha^{-z}} - \ell & \text{for } z \le {}^{\alpha}\log\beta,\ z \in \mathbb{Z}, \\ \\ r-(r+\ell)\dfrac{1}{1+\beta\alpha^{-z}} & \text{for } z > {}^{\alpha}\log\beta,\ z \in \mathbb{Z}. \end{cases}$$

In the case $\beta = 1$ we omit the superscript β. (We may distinguish $s(y)$ and $s(z)$ by the domains of the functions.) We state an important property of $s(z)$:

$$(16) \qquad s(z) = p_{z,z+1}s(z+1) + p_{z,z-1}s(z-1), \quad z \in \mathbb{Z}, \quad z \neq 0.$$

Relation (16) says that $s(z)$ is harmonic, except in $z = 0$ (cf.[3]).

We are searching for a stopping time $\underline{\tau}_0$ such that for all integers z

$$(17) \qquad v^{\beta}(z) := E_z[s^{\beta}(\underline{z}_{\underline{\tau}_0}) - \underline{\tau}_0 c] = \sup_{\underline{\tau}} E_z[s^{\beta}(\underline{z}_{\underline{\tau}}) - \underline{\tau}c],$$

where the subscript z means that we start the chain in state z. Of course we are only interested in $v^{\beta}(0)$.

The determination of $\underline{\tau}_0$ and $v^{\beta}(z)$ is known as an *optimal stopping problem in a Markov chain.* Note that it is equivalent to the sequential decision problem of section 2.

Without proof we state some results from the theory of optimal stopping (see [1],[4]). We formulate the properties for the chain $\underline{z}_1, \underline{z}_2, \underline{z}_3, \ldots$

A. *if* $\sup_z s^{\beta}(z) < \infty$ *and if* $c > 0$, *which is true in our case, then there exists an optimal stopping time* $\underline{\tau}_0$ *satisfying* (17).

B. $\underline{\tau}_0$ *is the entry time in the set* Γ, *where*

$$(18) \qquad \Gamma = \{z \mid v^{\beta}(z) = s^{\beta}(z),\ z \in \mathbb{Z}\}$$

C. $v^{\beta}(z)$ *is the minimal element in the set of solutions from the functional equation*

(19) $$w(z) = \max\{s^\beta(z), - c + p^\beta_{z,z+1}\ w(z+1) + p^\beta_{z,z-1}\ w(z-1)\}.$$

D. *We define recursively*

(20) $$v^\beta_0(z) = s^\beta(z)$$

$$v^\beta_n(z) = \max\{s^\beta(z), - c + p^\beta_{z,z+1}\ v^\beta_{n-1}(z+1) + p^\beta_{z,z-1}\ v^\beta(z-1)\}$$

The functions $v^\beta_n(z)$ *are nondecreasing and*

(21) $$\lim_{n\to\infty} v^\beta_n(z) = v^\beta(z)$$

The approximation of $v^\beta(z)$ is called the *method of successive approximations*.

For $\pi = \frac{1}{2}$, or equivalently $\beta = 1$, the solution of the optimal stopping problem shall be given explicitly in section 5.1. In section 4 we provide an algorithm to compute $v^1(0)$ and the set Γ in this case. In section 5.3 we prove that the prior distribution represented by $\pi = \frac{1}{2}$ is *least favourable* and that the decision rule derived for $\pi = \frac{1}{2}$ is a *minimax rule*.

4. *The algorithm for computing the optimal stopping time*

Of course we could use the method of successive approximations, defined in (20), to approximate $v(z)$ by $v_n(z)$, $z \in \mathbb{Z}$. Because the optimal stopping time $\underline{\tau}_0$, defined in (17), is the entry time of Γ, defined in (18), we have to be sure that also $\Gamma_n = \{z \mid v_n(z) = s(z),\ z \in \mathbb{Z}\}$ is equal to Γ. In general it may happen that the method of successive approximations requires an infinite number of iterations to provide this. Even if it is guaranteed that the method of successive approximations provides the optimal stopping set Γ in a finite number of iterations, one needs a criterion which says from which n on $\Gamma_n = \Gamma$. Our algorithm leads in a finite number of simple steps to the determination of $v(z)$ and Γ. In section 5.2 we show that the number of steps is less than $\frac{1}{2}(\frac{r+\ell}{c})$.

We know that $v(z)$ satisfies the functional equation (19). In section 5.1 it is proved that $v(z)$ has the following properties

(22) $v(z)$ *is symmetric around* $z = 0$.

There exists an integer $k > 0$ *such that*

$$(23) \qquad v(z) = \begin{cases} s(z) & \text{for } |z| \geq k \\ -c + p_{z,z+1}\, v(z+1) + p_{z,z-1}\, v(z-1) & \text{for } |z| < k. \end{cases}$$

We shall prove that *for each* k *there is only one function (on the integers)* f^k *which satisfies* (22) *and* (23). And we also prove that *there exists only one function which satisfies* (19), (22) *and* (23) *simultaneously*.

Our method will search in the class of all functions $\{f^k\}_{k=1}^{\infty}$ who satisfy (22) and (23), for some k, such that f^k also satisfies (19). It will be shown that *to check if the function* f^k *satisfies* (19) *it is only necessary to inspect this function in the point* $k-1$.
Indeed, if

$$s(k-1) < f^k(k-1) \text{ and } s(k) \geq -c + p_{k,k+1}\, s(k+1) + p_{k,k-1}\, f^k(k-1)$$

then f^k satisfies (19). Also a simple recursive relation to compute $f^k(k-1)$ shall be derived.

For readers, familiar with Markovian decision processes we note that our method is similar to *Howard's policy-iteration algorithm* (see [1]) in the sense that this algorithm, when started with the policy that prescribes stopping everywhere, will find, sequentially, policies of the type "stop if $|z| \geq k$" for $k = 1,2,3,\ldots$. However, our method does not need the value determination at each iteration. As is well-known the value determination step in Howard's algorithm is rather time consuming. The only thing we have to do is to check a simple recursive relation.

We shall formulate now our algorithm:

1) if $\alpha \geq \frac{r+\ell-2c}{r+\ell+2c}$

then stopping immediately is optimal and the expected return is $v(0) = \frac{1}{2}(r-\ell)$,

otherwise:

2) compute for $k = 1,2,3,\ldots$ the numbers b_k, defined by

$$b_0 = 1, \quad b_k = \frac{1+p_{k,k-1}b_{k-1}}{p_{k,k+1}};$$

check if

$$b_k \geq \frac{s(k+1)-s(k)}{c} - \frac{1}{p_{k+1,k}};$$

the first k which satisfies this inequality gives the decision points: $\pm(k+1)$

3) the expected return $v(0)$ is found by the iteration procedure
$v(k+1) = s(k+1)$,
$v(n) = v(n+1) - b_n c, \quad n = k, k-1,\ldots,0.$

Note that $p_{k,k+1}$, $p_{k,k-1}$ and $s(k)$ are defined in (13) and (15) respectively.

Some numerical results

Only the positive decision points are mentioned,

$r = 100$, $\ell = 5$, $c = 1$.

α	expected return	decision point
0	99.00	1
1/8	95.95	3
2/8	93.53	3
3/8	89.50	4
4/8	83.23	4
5/8	73.35	4
6/8	60.30	3
7/8	50.40	2
$\frac{103}{107} \leq \alpha \leq 1$	47.50	0

$r = 1000$, $\ell = 500$, $c = 1$.

α	expected return	decision point	α	expected return	decision point
0	999	1	11/16	920	12
1/16	996	3	12/16	879	14
2/16	994	4	13/16	803	16
3/16	992	5	14/16	652	17
4/16	990	5	15/16	389	12
5/16	987	6	31/32	285	6
6/16	983	7	63/64	258	3
7/16	977	7	127/128	251	1
8/16	970	8	255/256	250.46	1
9/16	959	9	$\frac{1498}{1502} \leq \alpha \leq 1$	250.00	0
10/16	944	11			

5. *Appendix*

In section 5.1 we give an explicit solution for the function $v(z)$ defined in (17) and for the entry time τ_0 of the set $\Gamma = \{z|s(z) = v(z)\}$, in the case that the prior distribution is represented by $\pi = \frac{1}{2}$. In section 5.2 we provide a justification of the algorithm to compute $v(z)$. In section 5.3 we show that the prior distribution $\pi = \frac{1}{2}$ is least favourable and that our decision rule is a minimax rule.

5.1. *The explicit solution*

In this section we take $\pi = \frac{1}{2}$, so that $\beta = 1$.

Lemma 1. *The function* $v(z)$ *is symmetric around* 0 *for* $z \in \mathbb{Z}$.

Proof. By induction we prove $v_n(z) = v_n(-z)$. From (15) with $\beta = 1$ we see that $s(z) = s(-z)$. So $v_0(z) = v_0(-z)$. Suppose $v_{m-1}(z) = v_{m-1}(-z)$ for $z \in \mathbb{Z}$. From (13) with $\beta = 1$ we see that $p_{z,z+1} = p_{-z,-z-1}$ and $p_{z,z-1} = p_{-z,-z+1}$. Hence

$$v_m(z) = \max\{s(z), -c + p_{z,z+1}\, v_{m-1}(z+1) + p_{z,z-1}\, v_{m-1}(z-1)\} =$$

$$= \max\{s(-z), -c + p_{-z,-z-1}\, v_{m-1}(-z-1) + p_{-z-z+1}\, v_{m-1}(-z+1)\} = v_m(-z).$$

By taking the limit for $m \to \infty$ and using (21) we find that $v(z) = v(-z)$. □

Lemma 2. *The set* $\{z|z \in \mathbb{Z}, v(z) > s(z)\}$ *is the set of integer points of a symmetric interval around* 0.

Proof. We first prove this property for $v_n(z)$. Indeed, $v_0(z) = s(z)$ and $v_1(0) = \max\{s(0), -c + s(1)\}$, hence $v_1(0)$ may be greater than $s(0)$. From (16) we see that $v_1(z) = \max\{s(z), -c + s(z)\}$, $z \neq 0$; therefore the statement is proved for $n = 1$. Suppose it is proved for $n = m - 1$. We may conclude from (21) that $\{z|z \in \mathbb{Z}, v_{m-1}(z) > s(z)\} \subset \{z|z \in \mathbb{Z}, v_m(z) > s(z)\}$. If, for $z \geq 1$, $v_{m-1}(z-1) = s(z-1)$ and $v_{m-1}(z+1) = s(z+1)$ it follows again from (16) that $v_m(z) = s(z)$. Hence there exists a k_m such that

$v_m(z) > s(z)$ if $|z| < k_m$ and $v_m(z) = s(z)$ if $|z| \geq k_m$. Because $v_m(k_m) = v_m(-k_m)$ the statement is proved for $n = m$.
Note that $\{z | z \in \mathbb{Z}, v(z) > s(z)\} = \bigcup_m \{z | z \in \mathbb{Z}, v_m(z) > s(z)\}$ which proves the lemma. □

Corollary. We define,

(24) $$k = \min\{z | z > 0, z \in \mathbb{Z}, v(z) = s(z)\}$$

then the stopping rule becomes: stop sampling as soon as one of the points $+k$ or $-k$ are reached. Note that $k < \infty$. Indeed, when $k = \infty$ then the sampling costs are infinite and this is certainly not optimal. From the functional equation it follows that the function $v(z)$ satisfies

(25) $$v(z) = \begin{cases} -c + p_{z,z+1}\, v(z+1) + p_{z,z-1}\, v(z-1) & \text{for } |z| < k \\ s(z) & \text{for } |z| \geq k. \end{cases}$$

Lemma 3. *If the function* $w(z)$ *is a solution of the functional equation* (19) *and* $w(z)$ *satisfies for some* $\ell > 0$

$$w(z) = \begin{cases} -c + p_{z,z+1}\, w(z+1) + p_{z,z-1}\, w(z-1) & \text{for } |z| < \ell \\ s(z) & \text{for } |z| \geq \ell \end{cases}$$

then $v(z) = w(z)$.

Proof. Assume $k > \ell$. Since $v(z) = -c + p_{z,z-1}\, v(z-1) + p_{z,z+1}\, v(z+1)$ and $w(z) = -c + p_{z,z-1}\, w(z-1) + p_{z,z+1}\, w(z+1)$ for $|z| < \ell$ and since $v(z) \geq s(z) = w(z)$ for $|z| \geq \ell$ we see that $u(z) := v(z) - w(z)$ satisfies

$$u(z) = p_{z,z-1}\, u(z-1) + p_{z,z+1}\, u(z+1) \qquad \text{for } |z| < \ell,$$
$$u(\ell) \geq 0 \text{ and } u(-\ell) \geq 0.$$

Let Q be the restriction of the matrix of transition probabilities of our

Markov chain to the rows and columns with numbers $-\ell+1,-\ell+2,\dots,0,\dots,\ell-1$. Then we may write the above equation in vector notation

$$u = Qu + d, \text{ with } d(-\ell+1) = p_{-\ell+1,\ell}u(-\ell) \text{ and } d(\ell-1) = p_{\ell-1,\ell}u(\ell)$$

and $d(z) = 0$ for $|z| < \ell - 1$. Hence $(I-Q)u = d$. Because $Q^n \to 0$ for $n \to \infty$ there exists an inverse of $I - Q$ which has only positieve entries, so that $u = (I-Q)^{-1}d \geq 0$ which implies

$$v(z) \geq w(z) \text{ for } z \in \mathbb{Z}.$$

On the other hand for $|z| < k$ we have

$$w(z) \geq -c + p_{z,z+1}\, w(z+1) + p_{z,z-1}\, w(z-1)$$

and

$$v(z) = -c + p_{z,z+1}\, v(z+1) + p_{z,z-1}\, v(z-1)$$

so that

$$u(z) \leq p_{z,z+1}\, u(z+1) + p_{z,z-1}\, u(z-1).$$

Since the Markov chain with transition probabilities as defined in (13) (with $\beta=1$) for $|z| < k$ and absorbing barriers in $\pm k$ is an absorbing Markov chain and, moreover, $u(-k) = u(k) = 0$, it follows from the above inequality that $u(z) \leq 0$ hence $v(z) \leq w(z)$. Therefore $w(z) = v(z)$ for $z \in \mathbb{Z}$. For $k < \ell$ the proof proceeds in a similar way. □

Lemma 4. *Let* k *be defined as in* (24), *then*

$$(26) \qquad s(k-1) < v(k-1) \leq s(k-1) + \frac{c}{p_{k,k-1}}$$

Proof. The first inequality is immediate from (19) and (24). Since v is a solution of (19) we have $v(k) = s(k) \geq -c + p_{k,k+1}\, s(k+1) + p_{k,k-1}\, v(k-1)$. Hence, using (16),

$$v(k-1) \le \{s(k) + c\} \frac{1}{p_{k,k-1}} - s(k+1) \frac{p_{k,k+1}}{p_{k,k-1}} = s(k-1) + \frac{c}{p_{k,k-1}} . \square$$

We shall introduce a class of functions $\{f^n(z) | n = 1,2,3,\dots, z \in \mathbb{Z}\}$ and we prove that $v(z)$ is the unique element of this class which satisfies (26) when $k \ge 1$.

Define recursively a sequence $\{b_n | n = 0,1,2,\dots\}$ by

$$(27) \qquad \begin{cases} b_0 = 1 \\ b_n = \dfrac{1+p_{n,n-1}b_{n-1}}{p_{n,n+1}} \end{cases}$$

For arbitrary positive integer we define a function $f^n(z)$ for $z \in \mathbb{Z}$ by

$$(28) \qquad \begin{cases} f^n(z) = s(z) & \text{for } z \ge n \\ f^n(z) = f^n(z+1) - b_z c & \text{for } z = n-1, n-2, \dots, 0 \text{ and} \\ f^n(z) = f^n(-z) & \text{for } z < 0. \end{cases}$$

<u>Lemma 5</u>. *The function* $f^n(z)$ *satisfies*

$$(29) \qquad f^n(z) = -c + p_{z,z+1} f^n(z+1) + p_{z,z-1} f^n(z-1) \quad \textit{for } |z| < n$$

<u>Proof</u>. From (27) we see that $p_{z,z+1}b_z = 1 + p_{z,z-1}b_{z-1}$.
By (28) we may state that

$$p_{z,z+1} f^n(z) = p_{z,z+1} f^n(z+1) - p_{z,z+1}b_z c, \qquad 0 \le z < n$$

hence

$$p_{z,z+1} f^n(z) = p_{z,z+1} f^n(z+1) - c - p_{z,z-1}b_{z-1}c.$$

Also from (28) it follows that

$$p_{z,z-1}\, f^n(z-1) = p_{z,z-1}\, f^n(z) - p_{z,z-1} b_{z-1} c, \qquad 1 \le z < n+1$$

so that

$$p_{z,z+1}\, f^n(z) + p_{z,z-1}\, f^n(z) = - c + p_{z,z+1}\, f^n(z+1) + $$
$$+ p_{z,z-1}\, f^n(z-1)$$

from which the statement follows for $1 \le z < n$. For $z = 0$ we see by (28) and (27) that $f^n(0) = f^n(1) - c$. Because $f^n(1) = f^n(-1)$ the statement is also proved in this case. □

Remark. Note that

$$(30) \qquad f^n(n-1) = s(n) - b_{n-1} c.$$

If $s(n-1) = s(n) - b_{n-1} c$, then $f^{n-1}(n-1) = s(n-1) = f^n(n-1)$ from which it follows by induction that $f^n(z) = f^{n-1}(z)$ for all z.

In lemma 6 we show that if $f^n(z)$ satisfies (26) then $f^n(z)$ is a majorant of $s(z)$.

Lemma 6. *If* $f^n(n-1) \ge s(n-1)$, *then* $f^n(z) \ge s(z)$ *for all* $z \in \mathbb{Z}$.

Proof. Since $f^n(z) = s(z)$ for $|z| \ge n$ and f and s are symmetric around 0 it remains to show that

$$\theta(z) := f^n(z) - s(z)$$

is nonnegative for $0 \le z \le n-2$. Let

$$t_i = \frac{1}{p_{n-i+1,n-i}}\,, \text{ then } 1 - t_i = - \frac{p_{n-i+1,n-i+2}}{p_{n-i+1,n-i}}\,.$$

From (16) we have

$$s(n-i) = t_i\, s(n-i+1) + (1-t_i)\, s(n-i+2), \qquad \text{for } i \le n,$$

and from (29) we see for $i \geq 2$

$$f^n(n-i) = t_i\{f^n(n-i+1) + c\} + (1-t_i)f^n(n-i+2).$$

Subtracting $s(n-i)$ from $f^n(n-i)$ gives

$$\theta(n-i) = t_i\theta(n-i+1) + (1-t_i)\theta(n-i+2) + t_i c. \tag{31}$$

With $\theta(n) = 0$ and $\theta(n-1) \geq 0$ by assumption we have $\theta(n-1) \geq \theta(n)$. Assume $\theta(n-i+1) \geq \theta(n-i+2)$ then, since $1-t_i<0$, it follows

$$(1-t_i)\theta(n-i+2) \geq (1-t_i)\theta(n-i+1).$$

Substituting this in (31) gives

$$\theta(n-i) \geq t_i\theta(n-i+1) + (1-t_i)\theta(n-i+1) + t_i c \geq \theta(n-i+1)\ .$$

Hence by induction is $\theta(z)$ decreasing in z for $z \geq 0$. With $\theta(n-1) \geq \theta(n) = 0$ we find $\theta(z) \geq 0$ for all $z \in \mathbb{Z}$. □

<u>Theorem 1</u>. 1) *If* $\alpha \geq \frac{r+\ell-2c}{r+\ell+2c}$ *then* $v(z) = s(z)$ *and immediately stopping is optimal.*

2) *Otherwise, let* n *be the unique natural number with*

$$\frac{s(n) - s(n-1)}{c} - \frac{1}{p_{n,n-1}} \leq b_{n-1} < \frac{s(n) - s(n-1)}{c}\ . \tag{32}$$

Then $f^n(z) = v(z)$, *and hence* $n = k$ *with* k *defined as in* (24).

<u>Proof of 1)</u>. $s(z)$ satisfies (16), hence

$$s(z) > - c + p_{z,z+1}\ s(z+1) + p_{z,z-1}\ s(z-1), \qquad z \neq 0.$$

$s(0) = \frac{1}{2}(r-\ell)$ and from (15) $s(1) = r-(r+\ell)\ \frac{\alpha}{\alpha+1}$. Therefore $s(0) \geq - c + s(1)$ if $\alpha \geq \frac{r+\ell-2c}{r+\ell+2c}$. Hence if $\alpha \geq \frac{r+\ell-2c}{r+\ell+2c}$ then $s(z)$ satisfies (19) for all z, and it is also the minimal solution of (19) hence $v(z) = s(z)$ for all $z \in \mathbb{Z}$

(this is also a consequence of lemma 3).

Proof of 2). Since $s(0) < -c + s(1)$ the function $s(z)$ does not satisfy the functional equation (19) and hence $k \geq 1$.

We first prove that $f^n(z)$ is a solution of (19). From $b_{n-1} < (s(n) - s(n-1))c^{-1}$ it follows that $s(n-1) < s(n) - b_{n-1}c$. According to lemma 6 and relation (28) we obtain $f^n(z) \geq s(z)$ for $|z| < n$. From lemma 5 and from (16) we have

$$f^n(z) \geq -c + p_{z,z+1}\, f^n(z+1) + p_{z,z-1}\, f^n(z-1), \tag{33}$$

for $|z| < n$ and $|z| > n$. From the first of the inequalities (32) and relation (28) we find

$$f^n(n-1) \leq s(n-1) + \frac{c}{p_{n,n-1}}\ .$$

Hence

$$f^n(n) = s(n) = p_{n,n+1}s(n+1) + p_{n,n-1}s(n-1) \geq$$

$$\geq -c + p_{n,n+1}s(n+1) + p_{n,n-1}f^n(n-1).$$

Consequently (33) holds for all $z \in \mathbb{Z}$. Hence $f^n(z)$ is a solution of (19). According to lemma 3 we have $f^n(z) = v(z)$ and in view of lemmas 3 and 4 there is exactly one natural number with property (32). $\square$

Remark. According to the lemmas 3 and 4 and theorem 1 we have that for n, defined in (32), $f^n(z)$ satisfies

$$s(n-1) < f^n(n-1) \leq s(n-1) + \frac{c}{p_{n,n-1}}\ . \tag{34}$$

and that $f^n(z)$ is the only f-function as defined in (28) with this property.

5.2. *Justification of the algorithm*

In this section we suppose again that $\beta = 1$. We have seen, that to find the value k it is only necessary to compute the numbers b_n, $n = 1,2,3,\ldots$ and to check the inequality

$$\frac{s(n) - s(n-1)}{c} - \frac{1}{p_{n,n-1}} \le b_{n-1} < \frac{s(n) - s(n-1)}{c} ,$$

which is equivalent to:

$$s(n-1) < f^n(n-1) \le s(n-1) + \frac{c}{p_{n,n-1}} .$$

We shall prove in lemma 7 that it is only necessary to check the left-hand side inequality of (32), or equivalently the right-hand side of (34). The first $n \ge 1$ for which this inequality holds is k. This gives, with the check if $\alpha < \frac{r+\ell-2c}{r+\ell+2c}$, the algorithm. Part 3) of the algorithm follows from theorem 1 ($f^n(z) = v(z)$) and relation (28).

Lemma 7. *If* $f^n(n-1) > s(n-1) + \frac{c}{p_{n,n-1}}$, *then* $f^{n+1}(n) > s(n)$ *for* $n = 1,2,3,\ldots$

Proof. From $f^n(n-1) > s(n-1) + \frac{c}{p_{n,n-1}}$ and $f^n(n-1) = s(n) - b_{n-1}c$ it follows that

$$s(n-1) + b_{n-1}c < s(n) - \frac{c}{p_{n,n-1}} \tag{35}$$

From (28) we have

$$f^{n+1}(n) = s(n+1) - b_n c = s(n+1) - \frac{(1+p_{n,n-1}b_{n-1})c}{p_{n,n+1}}$$

Hence, with (16)

$$p_{n,n+1}\, f^{n+1}(n) = s(n) - p_{n,n-1}s(n-1) - (c + p_{n,n-1}b_{n-1}c) =$$

$$= s(n) - p_{n,n-1}(s(n-1) + b_{n-1}) - c.$$

With (35) we find

$$p_{n,n+1}\, f^{n+1}(n) > s(n) - p_{n,n-1}\, s(n) = p_{n,n+1}\, s(n).$$

Therefore $f^{n+1}(n) > s(n)$. □

Corollary. Lemma 7 is equivalent to the following assertion:

$$\text{if } b_{n-1} < \frac{s(n)-s(n-1)}{c} - \frac{1}{p_{n,n-1}} \quad \text{then} \quad b_n < \frac{s(n+1)-s(n)}{c}.$$

The following lemma gives an upperbound for k and so for the number of iterations of the algorithm.

Lemma 8. *Let* $\tilde{n}$ *be the least* n *such that* $n > \frac{1}{2}(r+\ell)/c$, *then* $k < \tilde{n}$.

Proof. Consider the stopping time $\underline{\tau} \equiv 0$. Evidently when we start in $z = 0$,

$$E_0\, s(\underline{z}_{\underline{\tau}}) - \underline{\tau}c = s(0) = (r-\ell)/2.$$

All stopping times $\underline{\tau}$ with $P\{\underline{\tau} \geq \tilde{n}\} = 1$ have an expected return not larger than $r-\tilde{n}c$. Because $r-\tilde{n}c < (r-\ell)/2$, these stopping times have an expected return less than the stopping time $\underline{\tau} \equiv 0$, hence they are not optimal.

Suppose now that $k \geq \tilde{n}$. Then the optimal stopping time $\underline{\tau}$ when starting in $z = 0$, requires at least k steps, so that $P\{\underline{\tau} \geq k\} = 1$. Hence it is not optimal, which is a contradiction. Therefore $k < \tilde{n}$. □

5.3. *The minimax decision rule*

In this section we shall show that the prior distribution characterized by $\pi = \frac{1}{2}$ is least favourable.

We return to the original Markov chain $\underline{y}_0, \underline{y}_1, \underline{y}_2, \ldots$ defined in (3), with state space $[0,1]$.

In the same way as in (17) we define

$$w(y) := \sup_{\underline{\tau}} E_y[s(\underline{y}_{\underline{\tau}} - \underline{\tau}c] .$$

Obviously from (11) and (12) $v^{\beta}(0) = w(\frac{1}{1+\beta}) = w(\pi)$. Moreover, $w(\frac{1}{2})$ equals the expression defined in (5) for the optimal stopping time.

Similar to (20) we define recursively (recall (7) and (8))

$$w_0(y) = s(y)$$

$$w_n(y) = \max\{s(y), -c + \sum_{x=0,1} p_y(x)\, w_{n-1}(g_y(x))\}.$$

The properties (19) and (21) become for the chain $\underline{y}_0, \underline{y}_1, \underline{y}_2, \ldots$

$$w(z) = \max\{s(y), -c + \sum_{x=0,1} p_y(x)\, w(g_y(x))\}$$

and

$$w(z) = \lim_{n\to\infty} w_n(y).$$

Lemma 9. *The function* $w(y)$ *on* $[0,1]$ *is symmetric around* $y = \frac{1}{2}$ *and* $w(y)$ *is convex.*

Proof. Note that $s(y+\frac{1}{2}) = s(\frac{1}{2}-y)$ for $0 < y \le 1$ and that $p_{\frac{1}{2}+y}(x) = p_{\frac{1}{2}-y}(1-x)$ for $x = 0$ or 1. In the same way as in lemma 1 we may prove that $w_n(y)$ is symmetric around $\frac{1}{2}$ and therefore $w(y)$ also. Note that $s(y)$ is convex on $[0,1]$, therefore $w_0(y)$ is convex. Suppose now that $w_{n-1}(y)$ is convex for some n. Then $\sum_{x=0,1} p_y(x)\, w_{n-1}(g_y(x))$ is also convex (for a proof see [3]) and we may conclude that $w_n(y) = \max\{s(y), -c + \sum_{x=0,1} p_y(x) \cdot w_{n-1}(g_y(x))\}$ is convex. Therefore $w(y) = \lim_{n\to\infty} w_n(y)$ is convex. □

Corollary. $w(\frac{1}{2}) = \min_{y\in[0,1]} w(y)$.

We shall use the following version of the minimax theorem (see [2]): *If for a sequential decision problem the parameter set* Θ *is finite and the risk set* S *is bounded from below then*

$$\inf_{(\psi,\delta)} \sup_{\pi} r(\pi,(\psi,\delta)) = \sup_{\pi} \inf_{(\psi,\delta)} r(\pi,(\psi,\delta)),$$

where $r(\pi,(\psi,\delta))$ is defined in (4). The infimum has to be taken over all prior distributions and the supremum over all sequential decision rules. $R(\theta,(\psi,\delta))$ is the risk with respect to the prior distribution which gives probability 1 to θ.

A sequential decision rule (ψ_0,δ_0) *is called minimax if*

$$\sup_{\theta\in\Theta} R(\theta,(\psi_0,\delta_0)) = \inf_{(\psi,\delta)} \sup_{\theta\in\Theta} R(\theta,(\psi,\delta)).$$

In our case the risk set S is defined by

$$S = \{R(\theta,(\psi,\delta)),\ R(1-\theta,(\psi,\delta)) \mid (\psi,\delta) \text{ a sequential decision rule}\}.$$

Because

$$R(\theta,(\psi,\delta)) \geq -r \text{ and } R(1-\theta,(\psi,\delta)) \geq -r \qquad \text{for all } (\psi,\delta)$$

the risk set S is bounded from below and we may apply the minimax theorem.

Theorem 2. *The prior distribution represented by* $\pi = \frac{1}{2}$ *is least favourable and the sequential decision rule* (ψ_0,δ_0) *with* ψ_0 *the entry time of the set* $\{z \mid z \in \mathbb{Z},\ v(z) = s(z)\}$ *and* δ_0 *chooses* H_1 *if* $z > 0$ *and* H_0 *if* $z \leq 0$ *is a minimax rule.*

Proof. From (5) and (14) it follows that

$$-w(\pi) = \inf_{(\psi,\delta)} r(\pi,(\psi,\delta)).$$

According to lemma 9 and section 2 we have

$$-w(\tfrac{1}{2}) = \sup_{\pi} \inf_{(\psi,\delta)} r(\pi,(\psi,\delta)) = r(\tfrac{1}{2},(\psi_0,\delta_0)).$$

Therefore the prior distribution represented by $\pi = \frac{1}{2}$ is least favourable. Because $P_\theta\{\text{deciding } H_0\} = P_{1-\theta}\{\text{deciding } H_1\}$ for the decision rule (ψ_0,δ_0)

we have $R(\theta,(\psi_0,\delta_0)) = R(1-\theta,(\psi_0,\delta_0))$. Hence $r(\pi,(\psi_0,\delta_0)) = \sup_{\theta\in\Theta} R(\theta,(\psi_0,\delta_0))$ for all π. Note that

$$\sup_{\pi} r(\pi,(\psi,\delta)) = \sup_{\theta\in\Theta} R(\theta,(\psi,\delta)).$$

Applying the minimax theorem, we obtain

$$\sup_{\theta\in\Theta} R(\theta,(\psi_0,\delta_0)) = r(\tfrac{1}{2},(\psi_0,\delta_0)) = \sup_{\pi} \inf_{(\psi,\delta)} r(\pi,(\psi,\delta)) =$$
$$= \inf_{(\psi,\delta)} \sup_{\theta\in\Theta} R(\theta,(\psi,\delta))$$

which proves that (ψ_0,δ_0) is a minimax rule. □

References

[1] Derman, C., *Finite state Markovian decision processes*, Academic Press, New York, 1970.

[2] Ferguson, T.S., *Mathematical statistics. A decision theoretic approach*, Academic Press, New York, 1967.

[3] Hordijk, A. & K.M. Van Hee, *A Bayes process*, Mathematical Centre Report SW 23, Amsterdam, 1973.

[4] Hordijk, A., R. Potharst & J.Th. Runnenburg, *Optimal stopping of Markov chains*, MC Syllabus 19, Mathematical Centre, Amsterdam, 1973 (in Dutch).

[5] Ross, S., *Probability models with optimization applications*, Holden Day, New York, 1969.