Countably Additive Gambling

with a General Expected Reward

by

William D. Sudderth
University of Minnesota

Technical Report No. 352

July, 1979

* Research supported by National Science Foundation Grant MCS77-28424.

Abstract

Optimal strategies and the optimal return function are characterized for a Borel gambling problem in which the utility of a strategy is the expectation under the strategy of a general, measurable function g defined on the space of all infinite histories. These results are based on a previous paper with Lester Dubins where g was assumed to be shift-invariant.

1. Introduction. Let Γ be a Borel gambling house defined on the standard space $(F,\mathcal{B})$. (For definitions not given here, see [1] and [2].) Suppose the utility of a history $h = (f_1, f_2, \ldots)$ is $g(h)$ where g is a Borel measurable function from $H = F \times F \times \ldots$ to the extended real line $\overline{R}$, and suppose also that the utility of an analytically measurable strategy σ is

(1.1) $$\sigma g = \int g(h)\, d\sigma(h) .$$

At each $f \in F$, the optimal return $W_0(f)$ is defined by

$$W_0(f) = \sup \{\sigma g\colon \sigma \in \Gamma^\infty(f)\}$$

where $\Gamma^\infty(f)$ is the collection of all analytically measurable strategies available at f. The gambler with fortune f seeks a strategy $\sigma \in \Gamma^\infty(f)$ which is optimal in the sense that $\sigma g = W_0(f)$.

Consider the problem after play has proceeded for n days and the gambler has experienced the partial history $p_n = p_n(h) = (f_1, \ldots, f_n)$. A familiar formula for conditional expectations can be written as

(1.2) $$\sigma g = \int \sigma[p_n](gp_n)\, d\mu(p_n)$$

where μ is the distribution of p_n under σ, $\sigma[p]$ is the conditional strategy given the partial history p, gp is the function from H to $\overline{R}$ defined by $(gp)(h) = g(ph)$, and ph is the history consisting of the elements of p followed by those of h. Thus, in order for σ to be optimal at f, $\sigma[p_n](gp_n)$ must, almost surely under μ, be equal to $W_{p_n}(f_n)$ where, for each f and p,

(1.3) $$W_p(f) = \sup \{\sigma(gp)\colon \sigma \in \Gamma^\infty(f)\} .$$

In the special case where g is invariant in the sense that $gp = g$ for all partial histories p, each W_p is equal to W_0. In general, however, one must introduce the family $W = \{W_p : p \in F^*\}$ where F^* is set of all partial histories including the empty partial history which is denoted 0.

Optimal strategies and the optimal return function W_0 were characterized for invariant problems in [2]. Here a similar characterization of optimal strategies and the family W will be given for the general case. In fact, after a statement of results in section 2, it will be shown in section 3 that every problem (Γ, g) is equivalent, in a sense, to a certain invariant problem (Γ', g'). Then the general results can be derived from those for the invariant case. The theorems are illustrated in section 4 by a simple application to dynamic programming problems.

2. Formulation of results. To each family $Q = \{Q_p : p \in F^*\}$ of functions from F to $\bar{R}$ - in particular, to the family W - is associated a family $Q^* = \{Q_p^* : p \in F^*\}$ of functions from H to $\bar{R}$ defined by

$$Q_p^*(h) = \lim_n \sup Q_{pp_n}(f_n)$$

where $h = (f_1, f_2, \ldots)$ and $p_n = p_n(h) = (f_1, \ldots, f_n)$. The family Q is said to Γ-dominate g if, for every $p \in F^*$ and every analytically measurable σ available in Γ for which $\sigma(gp)$ is finite, the integral σQ_p^* exists and

$$\sigma Q_p^* \geq \sigma(gp) \ . \tag{2.1}$$

For each σ, let $T(\sigma)$ be the collection of all Borel measurable stopping times t defined on H which are finite with σ-probability one. For every stopping time t, $h = (f_1, f_2, \ldots)$, and $p \in F^*$, let

$p_t(h) = (f_1,\ldots,f_{t(h)})$ and set

$$Q_{pp_t}(h) = Q_{pp_t(h)}(f_{t(h)}) .$$

The family Q is Γ-excessive if, for all $f \in F$, $\sigma \in \Gamma^{\infty}(f)$, $t \in T(\sigma)$, and $p \in F^*$, σQ_{pp_t} exists and

(2.2) $$\sigma Q_{pp_t} \leq Q_p(f) .$$

Given two families $Q = \{Q_p\}$ and $R = \{R_p\}$, say that Q is smaller than R if $Q_p(f) \leq R_p(f)$ for all f and p.

Assume that, for all f, p, and $\sigma \in \Gamma^{\infty}(f)$, the integral $\sigma(gp)$ is well-defined so that the family W is also well-defined. Assume also that each W_p assumes only finite real values.

Theorem 1. $W = \{W_p\}$ is the smallest analytically measurable family which is Γ-excessive and Γ-dominates g.

Assume next that $f \in F$, $\sigma \in \Gamma^{\infty}(f)$, and σg is finite.

Theorem 2. For σ to be optimal at f it is necessary and sufficient that $\sigma g = \sigma W_0{}^*$ and any (all) of the following three conditions be satisfied.

(a) $\sigma W_0{}^* = W_0(f)$.

(b) $W_0(f), W_{p_1}(f_1), W_{p_2}(f_2),\ldots$ is a uniformly integrable martingale under σ.

(c) $W_0(f), W_{p_1}(f_1), W_{p_2}(f_2),\ldots$ is an L_1-bounded martingale under σ which satisfies

$$\sigma W_{p_t} \geq W_0(f)$$

for all $t \in T(\sigma)$.

The proofs of these theorems are based on a reduction to the invariant case which is presented in the next section.

3. Reduction to the invariant case. To the problem (Γ, g) defined on the fortune space F will now be associated an invariant problem (Γ', g') on F' which will be seen to be equivalent in many respects.

First, let

$$F' = F \times F^* .$$

Now $F^* = \bigcup_{n=0}^{\infty} F^{(n)}$ where $F^{(n)}$ is the set of all partial histories of length n. ($F^{(0)}$ is the singleton containing the empty partial history 0.) Because the countable union of standard spaces is standard and the product of standard spaces is standard, the space F' is standard.

To define Γ' associate to each $(f,p) \in F'$ and $\gamma \in \Gamma(f)$ the probability measure γ^p defined on the Borel subsets of F' by the formula

$$\int_{F'} \varphi(f') d\gamma^p(f') = \int_F \varphi(f_1, pf_1) d\gamma(f_1)$$

for φ a bounded, Borel function from F' to R. In other words, γ^p is the distribution of (f_1, pf_1) if γ is the distribution of f_1. Set

$$\Gamma'(f,p) = \{\gamma^p : \gamma \in \Gamma(f)\} .$$

It is easy to verify that Γ' is Borel because Γ is.

To complete the definition of the gambling problem, let $h' = (f_1', f_2', \ldots) \in H'$ where $f_i' = (f_i, p_i)$ for $i = 1,2,\ldots,$ and set

$$g'(h') = \lim_n \sup(gp_n)(f_{n+1}, f_{n+2}, \ldots).$$

It is trivial to verify that g' is shift-invariant on H'.

There is a natural correspondence between strategies available at f in Γ and those available at (f,p) in Γ'. First define a correspondence between histories as follows: for $h = (f_1, f_2, \ldots) \in H$ and $p \in F^*$, let $h^p \in H'$ be the history

$$h^p = ((f_1, pp_1), (f_2, pp_2), \ldots)$$

where

$$p_n = p_n(h) = (f_1, \ldots, f_n)$$

for $n = 1, 2, \ldots$. Then for each strategy σ on H and each $p \in F^*$, let σ^p be a strategy with initial gamble

$$(\sigma^p)_0 = (\sigma_0)^p$$

and such that, for every $p' = ((f_1, pp_1), \ldots, (f_n, pp_n))$,

$$(\sigma^p)_n(p') = (\sigma_n(p_n))^{pp_n} .$$

If σ is thought of as the distribution of the random element $h \in H$, then σ^p corresponds to the distribution of $h^p \in H'$. In particular, σ^p assigns full measure to the subset H^p of H' where $H^p = \{h^p : h \in H\}$. Moreover, for every h^p,

$$g'(h^p) = (gp)(h)$$

and, consequently,

$$\int g' \, d\sigma^p = \int (gp) d\sigma . \tag{3.1}$$

In the special case when p is the empty partial history, (3.1) becomes

$$\int g' \, d\sigma^0 = \int g d\sigma . \tag{3.2}$$

It is easy to check that, for each (f,p), $(\Gamma')^{\infty}(f,p)$ is the collection of probability measures $\{\sigma^{p}:\sigma \in \Gamma^{\infty}(f)\}$. So, by (3.1) and (1.3),

$$W'(f,p) = W_p(f)$$

where W' is the optimal return function for the problem (Γ',g'). In particular,

$$W'(f,0) = W_0(f). \tag{3.3}$$

It follows from (3.2) and (3.3) that a strategy $\sigma \in \Gamma^{\infty}(f)$ is optimal at f if and only if $\sigma^{0} \in (\Gamma')^{\infty}(f,0)$ is optimal at $(f,0)$.

The proof of Theorem 1 is now straightforward from Theorem 1 of [2].

Proof of Theorem 1. To each family $Q = \{Q_p : p \in F^{*}\}$ of functions on F corresponds the function Q' on F' defined by $Q'(f,p) = Q_p(f)$. Furthermore, it is easy to verify that Q is analytically measurable, Γ-excessive, Γ-dominates g, and less than the family R precisely when Q' is analytically measurable, Γ'-excessive, Γ'-dominates g' and less than the function R'. By [2, Theorem 1], W' is the smallest analytically measurable function which is Γ'-excessive and Γ'-dominates g'. Theorem 1 is now clear. ∎

The proof of Theorem 2 is a similar translation of Theorem 2 of [2] to the present setting. Additional information about optimal and also ε-optimal strategies can be obtained by translating the results given in [4] for invariant problems to the general setting.

4. An application to dynamic programming. In the setting of dynamic programming problems, the results of this note specialize to yield results closely related to the work of Ulrich Rieder in [3]. To get

the basic idea, suppose that, for $h = (f_1, f_2, \ldots)$,

$$(4.1) \qquad g(h) = \sum_{n=1}^{\infty} r(f_n)$$

where r is a Borel function from F to R. Define

$$r(p) = \sum_{n=1}^{k} r(f_n)$$

for each $p = (f_1, \ldots, f_p) \in F^*$ so that

$$(gp)(h) = r(p) + g(h)$$

and, by (1.3),

$$W_p(f) = r(p) + W_0(f).$$

(It is assumed, as before, that σg exists for all available σ.) Thus the family $\{W_p\}$ is completely determined by W_0 and to characterize it among families $Q = \{Q_p\}$, it is only necessary to consider those Q which also satisfy

$$(4.2) \qquad Q_p(f) = r(p) + Q_0(f)$$

for all f and p. For such a Q, the condition that it be Γ-excessive simplifies to state that, for all $f \in F$, $\sigma \in \Gamma^\infty(f)$, and $t \in T(\sigma)$

$$(4.3) \qquad \sigma[r(p_t) + Q_0(f_t)] \leq Q_0(f).$$

Similarly, for a Q satisfying (4.2),

$$Q_p^*(h) = (gp)(h) + \limsup_n Q_0(f_n)$$

for each h for which the series in (4.1) is well-defined. Thus the condition that Q Γ-dominates g says here only that, for each σ available, (4.4) $$\sigma(\limsup_n Q_0(f_n)) \leq 0.$$

Theorem 1 now specializes to the present setting to say that the optimal return function W_0 is the least analytically measurable function satisfying (4.3) and (4.4).

It is equally straightforward to obtain the specialization of Theorem 2 to this setting.

References

1. Dubins, Lester E. and Savage, Loenard J. (1976). Inequalities for Stochastic Processes: How to Gamble If You Must. Dover, New York.

2. Dubins, Lester E. and Sudderth, William D. (1977). Countably additive gambling and optimal stopping. Z. Wahrscheinlichkeitstheorie, 41, 59-72.

3. Rieder, Ulrich (1976). On optimal policies and martingales in dynamic programming. J. Applied Prob. 13, 507-518.

4. Sudderth, W. (1972). On the Dubins and Savage characterization of optimal strategies. Ann. Math. Statist. 43, 498-507.