The Annals of Mathematical Statistics
1968, Vol. 39, No. 6, 1932–1945

ON PSEUDO-GAMES[1]

By Alfredo Baños[2]

San Fernando Valley State College

1. Introduction and summary. In the definition of a two-person zero-sum game given by Von Neumann and Morgenstern it is assumed that both players know the rules of the game (e.g., the game tree, the information sets as well as the distributions of the ensuing payoffs for given strategy choices, etc.). We use the term *pseudo-game* to denote the case where at least one player does not have complete information.[3]

In this paper we restrict our attention to those pseudo-games in which player I, say, is only aware of his set of pure strategy choices (assumed to contain m elements: $2 \leqq m < \infty$) and not of player II's strategy choices (assumed to have uniformly bounded second moments). Player II is assumed to have complete information. More precisely, we shall study pseudo-games G that have the format given below:

Let $A = \{a_1, \cdots, a_m\}$ denote the pure strategy choices of player I. Denote by A^* the set of probability distributions p over A (player I's mixed strategy choices). We sometimes write p in the form $(p(1), \cdots, p(m))$, $\sum_{j=1}^{m} p(j) = 1$, and $p(j) \geqq 0$, with the interpretation that when player I uses p he will play a_j with probability $p(j)$. Any element of A^* that assigns mass 1 to some $a \varepsilon A$ will be simply denoted by a.

Let B denote the set (not necessarily finite) of pure strategies for player II. Let $\mathcal{B}$ be a fixed σ-field of subsets of B and denote by B^* the set of all probability distributions q over $\mathcal{B}$ (player II's mixed strategies). We assume that $\mathcal{B}$ contains all single point sets of B, so that B^* contains all finite probability distributions over B. We postulate that we are given for each pair (a, b) in the product space $A \times B$ a distribution $P_{(a,b)}$ on the real line which represents the distribution of the loss incurred by player I (or gain by player II) if $a \varepsilon A$ is the strategy choice of I and $b \varepsilon B$ is the strategy choice of II.

Contrary to the usual practice, the payoff for given pure strategy choices is thus allowed to be random. We do this in order that our main results may be proved in greater generality. An example of a pseudo-game with random payoffs is given in Section 2.

Received 22 March 1967; revised 9 February 1968.

[1] The initial version of this paper was the author's doctoral dissertation at UCLA. The research was supported in part by the National Science Foundation under Grant GP-5224.

[2] Now at the University of Connecticut.

[3] "Complete information," a knowledge of the rules of the game is to be distinguished from "perfect information," the latter requiring that each branch point of the game tree be an information set, or equivalently that at each stage, both players have complete knowledge of all past moves of the game. See Von Neumann and Morgenstern [8].

The distributions $P_{(a,b)}$ are assumed to have uniformly bounded second moments. For each $a \varepsilon A$, and fixed Borel set C, $P_{(a,\cdot)}(C)$ is assumed to be $\mathcal{B}$-measurable. For each pair $(a, b) \varepsilon A \times B$, let $X_{(a,b)}$ be a random variable having $P_{(a,b)}$ as its distribution. Suppose that players I and II are using strategies p and q, respectively. They can determine the payoff of the pseudo-game by first selecting an $a \varepsilon A$ and a $b \varepsilon B$ according to the distributions p and q, respectively, and then treating an observed value of $X_{(a,b)}$ as the payoff.

For every pair of strategies (p, q) that players I and II may use we define the expected value of the payoff $R(p, q)$ by means of the equation[4]:

$$(1.1) \qquad R(p, q) = \textstyle\sum_{j=1}^{m} p(j) \int [\int x \, dP_{(a_j,b)}(x)] \, dq(b).$$

We are assuming that player I is only aware of the set A, while player II has complete information. However, by assuming instead that player I is also aware of the set B as well as the distributions $P_{(a,b)}$, $(a, b) \varepsilon A \times B$, we can associate with every such pseudo-game G a game with complete information G'. Such concepts as "value" and "minimax strategy" do not carry over to pseudo-games. However by the minimax theorem, since A is assumed finite, every such game G' will have a value v_G and player I will have a minimax strategy p':

$$(1.2) \quad v_G = \sup_{q\varepsilon B^*} R(p', q) = \inf_{p\varepsilon A^*} \sup_{q \varepsilon B^*} R(p, q) = \sup_{q\varepsilon B^*} \inf_{p \varepsilon A^*} R(p, q).$$

Suppose now that players I and II are playing a sequence of identical pseudo-games of the type we have been describing; i.e., they play one game, observe their losses and play the same game again (with possibly different strategy choices), continuing in this manner ad-infinitum. We shall refer to the individual games that make up the sequence as the *subgames* of the sequence. When playing such a sequence of pseudo-games a strategy for player I would be a rule P that would tell him for every j, as a function of his past plays (mixed strategy choices) and losses what mixed strategy to play during the jth subgame; a strategy for player II would be a rule Q that would tell him for every j, as a function of his own, and his opponent's past plays and losses, what mixed strategy q to play during the jth subgame of the sequence. We are thus allowing player II to know what plays player I has made, but we are not granting I the same favor.

Among the rules P available to player I we define a special class of rules to be called rules constant on intervals. If x is any real number let $[x]$ denote the largest integer that is less than or equal to x. For every $\alpha > 1$, let $\Pi(\alpha) = (I_1(\alpha), I_2(\alpha), \cdots, I_n(\alpha), \cdots)$ denote the partition on the set I of positive integers defined by the equations:

$$(1.3) \quad I_n(\alpha) = \{(\textstyle\sum_{k=1}^{n-1} [k^\alpha]) + 1, (\sum_{k=1}^{n-1} [k^\alpha]) + 2, \cdots, \sum_{k=1}^{n} [k^\alpha]\}; \qquad n = 1, 2, 3, \cdots.$$

For example, $I_1(2) = \{1\}$, $I_2(2) = \{2, 3, 4, 5\}$, $I_3(2) = \{6, 7, \cdots, 14\}$; etc. We

[4] Note that the functions $\int x^k \, dP_{(a_j,\cdot)}(x)$, $k = 1, 2$ and $j = 1, \cdots, m$, can be expressed as the limits of finite sums of $\mathcal{B}$-measurable functions and are therefore also $\mathcal{B}$-measurable.

shall refer to $I_n(\alpha)$ as the nth interval of the partition $\Pi(\alpha)$. Note that the cardinality of $I_n(\alpha)$ is $[n^\alpha]$. Let us suppose that player I is using some rule P that assigns, with probability 1, the same mixed strategy to the ith subgame as it does to the jth subgame whenever i and j belong to the same interval $I_n(\alpha)$, $n = 1, 2, 3, \cdots$. In this case we say that P is *constant on intervals*. Thus if we say that player I is to play a certain strategy p during the nth interval of a partition $\Pi(\alpha)$, we mean that he is to play p during every subgame whose index belongs to $I_n(\alpha)$. The particular strategy that player I uses in the nth interval (a random variable depending on plays and losses occurring prior to the nth interval) will be denoted by p_n.

For $j = 1, 2, 3, \cdots, N, \cdots$ let X_j represent the loss incurred by player I during the jth subgame. Note that the sequence $\{X_n\}$ is a discrete stochastic process whose index set is the set I of positive integers and whose law of evolution is determined by the distributions $P_{(a,b)}$ and by the rules P and Q that the players use. The first objective of this paper is to prove:

THEOREM. *Suppose players* I *and* II *are playing a sequence of identical pseudo-games G satisfying* (i) *and* (ii):

(i) *Player* I *has $m \geqq 2$ pure strategy choices.*

(ii) *The distributions $P_{(a,b)}$ have uniformly bounded second moments and for each $a \varepsilon A$ and every Borel set C, $P_{(a,\cdot)}(C)$ is $\mathcal{B}$-measurable. Then there exists a class of rules $\{P\}_m$ for player* I *such that for all rules Q that player* II *may use we have*:

$$P \,\varepsilon\, \{P\}_m \Rightarrow \Pr\left(\limsup_{N\to\infty} N^{-1} \textstyle\sum_{j=1}^{N} X_j \leqq v_G \mid P, Q\right) = 1.$$

We will show, that is, that the player with incomplete information can do as well asymptotically as he could if he had complete information.

The members of $\{P\}_m$ will all be constant on intervals. Our second objective will be to seek a strong convergence rate for $N^{-1} \sum_{j=1}^{N} X_j$. In the course of achieving this goal we will show that a good partition is obtained by setting α equal to $(m + 2)/m$.

2. Examples. A good poker player gains information about an opponent's strategies by observing his eccentricities: his hesitations, his apparent nervousness or calm, the way he holds his cards, etc. Because a player may not be aware of his eccentricities poker is, from this viewpoint, an example of a pseudo-game.

The following is a more concrete example: Consider first a game of matching pennies: players I and II's possible plays being H or T (head or tail). Suppose player I pays player II one unit if the sides of the coins match, and incurs no loss otherwise. A strategy for player I would be a number π $(0 \leqq \pi \leqq 1)$, with the interpretation that when he uses π he will play H with probability π. A minimax strategy for player I would be $\pi = \frac{1}{2}$ and the value of the game is $\frac{1}{2}$. Suppose now that player II is a very perceptive opponent and is gaining information from player I's eccentricities. More precisely, let us suppose that player I initiates the game by playing either H or T. After I's play, Nature (a third player who operates as player II's spy) will play either θ_1 or θ_2. We assume that $P(\theta_1 \mid H)$

$= P(\theta_2 \mid H) = \frac{1}{2}$ and that $P(\theta_1 \mid T) = 1 = 1 - P(\theta_2 \mid T)$; but this information as well as Nature's actual play in any particular instance will be known only by player II. Player II observes Nature's play and proceeds to play either H or T. As before, we let π denote a mixed strategy choice for player I. A strategy for player II would now be a pair of numbers $(p, q)(0 \leqq p \leqq 1$ and $0 \leqq q \leqq 1)$, with the interpretation that when he uses (p, q) he will play H with probability $p(q)$ if he observes $\theta_1(\theta_2)$. (p, q) is known as a test in statistical parlance. Some of the payoffs for given pure strategy choices are random. Thus if player II uses $(0, 1)$ and player I plays H, the payoff will be 1 with probability $\frac{1}{2}$ and 0 with probability $\frac{1}{2}$. For this example the risk function $R(\pi, (p, q))$ for player I (his expected loss) can be computed as follows:

$$\begin{aligned} R(\pi, (p, q)) &= \Pr((H, H)|(\pi, (p, q))) + \Pr((T, T)|(\pi, (p, q))) \\ &= \pi[P(\theta_1 \mid H)p + P(\theta_2 \mid H)q] \\ &\quad + (1 - \pi)[P(\theta_1 \mid T)(1 - p) + P(\theta_2 \mid T)(1 - q)] \\ &= \tfrac{1}{2}\pi(p + q) + (1 - \pi)(1 - p). \end{aligned} \tag{2.1}$$

Suppose, only for the moment, that player I is also aware of Nature's sample space and of the probabilities $P(\theta_i \mid H)$ and $P(\theta_i \mid T)$ for $i = 1, 2$. Under this added condition, our example becomes a game in the Von Neumann-Morgenstern sense. The value of the game is $v = \frac{2}{3}$ and $\pi = \frac{2}{3}$ is a minimax strategy for player I.

Player I, by observing his losses over such a sequence of pseudo-games might begin to suspect that he is divulging information to his opponent in one way or another. He knows that $\pi = \frac{1}{2}$ is a minimax strategy in the ordinary game of matching pennies, and this fact might lead him to believe that $\pi = \frac{1}{2}$ would still be a reasonable strategy in the more general case that we have been examining. However, if in this example, player I uses $\pi = \frac{1}{2}$, then player II can use the pure strategy $(0, 1)$ and we will have:

$$R(\tfrac{1}{2}, (0, 1)) = \tfrac{1}{2}\cdot\tfrac{1}{2}\cdot 1 + \tfrac{1}{2}\cdot 1 = \tfrac{3}{4} > \tfrac{2}{3}. \tag{2.2}$$

Player I will begin to suspect that he is playing a pseudo-game when it appears to him that although he is playing what he thinks is his minimax strategy, $\pi = \frac{1}{2}$, he is losing more than half the time. The question is what can he do about it? It is clear (to player II) that player I's average loss can be kept to at least $\frac{2}{3}$. The results of this paper show that player I has a strategy that keeps the limiting average loss to at most $\frac{2}{3}$.

3. The class of rules $\{P\}_m$. We begin by giving some preliminary definitions. Let $S_k(k = 1, 2, 3, \cdots)$ be the set of distributions on A that satisfy for every $j(j = 1, \cdots, m)$ the condition that $p(j) = i/(2^k m)$ for some $i \varepsilon \{0, 1, \cdots, 2^k m\}$. The number of such distributions is:

$$\Phi_k = \binom{2^k m + m - 1}{m - 1}. \tag{3.1}$$

Now let α and β be two fixed constants. Define the sequence $n_0, n_1, n_2, \cdots, n_k, \cdots$ by means of the equations:

$$(3.2) \qquad n_0 = 0, \qquad n_k = n_{k-1} + [\Phi_k^{1/\beta}] \quad \text{for} \quad k = 1, 2, 3, \cdots.$$

The members of $\{P\}_m$ are determined by two parameters: α and β. An arbitrary member of $\{P\}_m$ will be denoted as $P(\alpha, \beta)$. We impose two conditions on the parameters:

$$(3.3) \qquad \text{(i)} \quad \alpha > 1, \qquad \text{(ii)} \quad 0 < \beta < 1.$$

Restriction (ii) insures that $n_k - n_{k-1}$ is never smaller than Φ_k. Condition (i) is necessary for the proof that is presented of Lemmas 5.2 and 5.3.

A rule $P(\alpha, \beta)$ will be constant on the intervals of the partition $\Pi(\alpha)$. For each interval, player I records the mixed strategy he uses for that interval as well as the average loss incurred by him in that interval. For every interval n, p_n (the strategy used in the nth interval) is determined by $P(\alpha, \beta)$ in the following manner:

Begin by ordering the elements of S_k ($k = 1, 2, 3, \cdots$) in any manner and call the jth member of the sequence $p_j^{(k)}$. The members of S_k are called the *available probabilities for the kth stage* (the intervals numbered $n_{k-1} + 1$ through n_k). Begin play as follows: During interval j ($j = 1, 2, \cdots, \Phi_1$) play $p_j^{(1)}$. During those intervals numbered $\Phi_1 + 1$ through n_1 play any of the available probabilities for the first stage whose greatest recorded average loss incurred in any interval is a minimum. In general, during interval i ($i = n_{k-1} + 1, \cdots, n_{k-1} + \Phi_k$) play $p_{i-n_{k-1}}^{(k)}$. During those intervals numbered $n_{k-1} + \Phi_k + 1$ through n_k, play any of the available probabilities for the kth stage whose greatest recorded average loss incurred in any interval *after interval* n_{k-1} is a minimum.

An example may clarify the ideas. Suppose $m = 2$, $\alpha = 2$, and $\beta = \frac{1}{3}$. Then $\Phi_1 = 5$ and $n_1 = 125$. $S_1 = \{(1, 0), (0, 1), (\frac{1}{4}, \frac{3}{4}), (\frac{3}{4}, \frac{1}{4}), (\frac{1}{2}, \frac{1}{2})\}$. Let $p_i^{(1)}$ ($i = 1, \cdots, 5$) be the ith member of the ordered sequence given above. During the interval i ($i = 1, \cdots, 5$) player I is to play $p_i^{(1)}$. Suppose that the average losses incurred by him in each of those intervals were $7^{\frac{1}{2}}$, 4.8, 2, 3, and 5/2, respectively.

According to our rule during the 6th interval he would play $p_3^{(1)} = (\frac{1}{4}, \frac{3}{4})$. Suppose that he does this and that the average loss incurred in that interval is -8. For the 7th interval the indicated choice is again $(\frac{1}{4}, \frac{3}{4})$. Suppose he plays $p_3^{(1)}$ in the 7th interval and that the average loss incurred in that interval is 5/2. For the 8th interval he can choose either $p_3^{(1)}$ or $p_5^{(1)}$, but suppose that he plays $p_3^{(1)}$ again and that the average loss incurred in that interval is 10^{100}. For the 9th interval the indicated choice is now $p_5^{(1)}$ or $(\frac{1}{2}, \frac{1}{2})$. And so on, until the 125th interval. Then the whole process begins again; this time using Φ_2 and n_2. He repeats this procedure again and again, ad-infinitum.

4. The method of proof. Define the distance d of any two distributions p_1 and p_2 in A^* to be:

$$(4.1) \qquad d(p_1, p_2) = [\sum_{j=1}^{m} (p_1(j) - p_2(j))^2]^{\frac{1}{2}}.$$

Note that under this relation A^* is a metric space. For $p \,\varepsilon\, A^*$, define $F(p)$ by means of the equation:

$$F(p) = \sup_{q \varepsilon B^*} R(p, q). \tag{4.2}$$

The proof of the theorem presented in the introduction depends on the following results:

LEMMA 4.1. *For all p_1 and p_2 in A^* we have $|F(p_1) - F(p_2)| \leqq c\, d(p_1, p_2)$ where $c = m^{\frac{1}{2}} \max_{1 \leqq j \leqq m} \sup_{q \varepsilon B^*} |R(a_j, q)|$.*

PROOF. Since the distributions $P_{(a,b)}$ are assumed to have uniformly bounded second moments $m^{\frac{1}{2}} \max_{1 \leqq j \leqq m} \sup_{q \varepsilon B^*} |R(a_j, q)|$ is finite and we are able to prove that the function F is Lipschitzian. Assume without loss of generality that $F(p_1) \geqq F(p_2)$. According to (4.2) for every $\epsilon > 0$ there exists a $q(\epsilon) \,\varepsilon\, B^*$ such that $F(p_1) < R(p_1, q(\epsilon)) + \epsilon$. Therefore:

$$\begin{aligned} |F(p_1) - F(p_2)| &= F(p_1) - F(p_2) < |R(p_1, q(\epsilon)) - R(p_2, q(\epsilon))| + \epsilon \\ &= |\textstyle\sum_{j=1}^{m} R(a_j, q(\epsilon))(p_1(j) - p_2(j))| + \epsilon \\ &\leqq \max_{1 \leqq j \leqq m} \sup_{q \varepsilon B^*} |R(a_j, q)| \textstyle\sum_{i=1}^{m} |p_1(i) - p_2(i)| + \epsilon \\ &\leqq \max_{1 \leqq j \leqq m} \sup_{q \varepsilon B^*} |R(a_j, q)| m^{\frac{1}{2}} d(p_1, p_2) + \epsilon. \end{aligned} \tag{4.3}$$

Since this relation is valid for all $\epsilon > 0$ the theorem is proved.

LEMMA 4.2 *For every $p \,\varepsilon\, A^*$ and every $k \geqq 1$ there exists a distribution $g_p^{(k)} \,\varepsilon\, S_k$ such that $d(p, g_p^{(k)}) < 2^{-k}$.*

PROOF. The proof is straightforward but tedious and will therefore be omitted. As no uniqueness is implied by the lemma, for fixed $p \,\varepsilon\, A^*$ we will let $g_p^{(k)}$ simply denote an arbitrary element of S_k satisfying $d(p, g_p^{(k)}) < 2^{-k}$.

Although the proof of the main results that is to follow depends on many concepts and is necessarily laborious, the basic idea, which we now outline, is quite simple and intuitive. Consider play within the kth stage only. From (1.2) and (4.2) there exists a $p' \,\varepsilon\, A^*$ satisfying $v_G = F(p') = \inf_{p \varepsilon A^*} F(p)$. Combining Lemmas 4.1 and 4.2 player I knows that for some available probability $g_{p'}^{(k)}$, the average loss incurred in any interval in which $g_{p'}^{(k)}$ is used will not exceed v_G by more than $c2^{-k} + \epsilon_k$ (where ϵ_k is the kth member of a sequence of random variables satisfying $P(\limsup_{k \to \infty} \epsilon_k > 0) = 0$, as will be shown by Lemma 5.2). Now player II can trick him into using an inferior strategy by allowing him some early wins and then reap the profits temporarily. But the key point is that player I may use a strategy p'' other than $g_{p'}^{(k)}$ only so long as the average loss incurred in any interval in which p'' is used is less than or equal to $v_G + c2^{-k} + \epsilon_k$. Since the number of available probabilities for the kth stage is Φ_k, it is clear that player I will not incur an average loss exceeding v_G by more than $c2^{-k} + \epsilon_k$ on more than $\Phi_k - 1$ intervals. To convince oneself of the plausibility of the theorem presented in the introduction, one has only to note that $\Phi_k = o(n_k - n_{k-1})$, the number of intervals comprising the kth stage.

5. The proof. Throughout this section we assume that players I and II are playing a sequence of identical pseudo-games G having the format described in the introduction and that the value of the corresponding game G' with complete information is v_G. We assume that player I is using a rule $P(\alpha, \beta) \varepsilon \{P\}_m$, for some fixed $m \geqq 2$, as described in Section 3. Except for the discussion of Theorem 5.2, α and β are assumed fixed.

Define the constant M as follows:

$$M = \sup_{(p,q)\varepsilon A^* \times B^*} \textstyle\sum_{j=1}^{m} p(j) \int [\int x^2 \, dP_{(a_j,b)}(x)] \, dq(b). \tag{5.1}$$

For $p \varepsilon A^*$, define $G(p)$ as follows:

$$G(p) = \inf_{q\varepsilon B^*} R(p, q). \tag{5.2}$$

For $n = 1, 2, 3, \cdots$, and $j = 1, \cdots, [n^\alpha]$, let $Y_{n,j}$ denote the loss incurred by player I during the jth subgame of the nth interval, so that $\sum_{j=1}^{[n^\alpha]} Y_{n,j}$ denotes the total loss incurred by player I in the nth interval. For arbitrary n, the random variables $Y_{n,j}$ are not in general independent, since even if player I is using a rule that is constant on intervals, there is no guarantee that player II is behaving in a similar manner. However, because of the way M, $F(p)$, and $G(p)$ were defined, the following important inequalities are immediately forthcoming for all n and $j = 1, \cdots, [n^\alpha]$:

$$E((Y_{n,j})^2 \mid Y_{n,j-1}, \cdots, Y_{n,1}) \leqq M \qquad \text{a.s.,} \tag{5.3}$$

$$G(p_n) \leqq E(Y_{n,j} \mid Y_{n,j-1}, \cdots, Y_{n,1}, p_n) \leqq F(p_n) \qquad \text{a.s.} \tag{5.4}$$

If $j = 1$, (5.3) and (5.4) are simply to be read as $E((Y_{n,1})^2) \leqq M$ and $G(p_n) \leqq E(Y_{n,1} \mid p_n) \leqq F(p_n)$, respectively.

Lemma 5.1. *For $n = 1, 2, 3, \cdots ; 1 \leqq i \leqq [n^\alpha]$, and all $\gamma > 0$, there is a positive constant M^* satisfying*:

(i) $P(\sum_{j=i}^{[n^\alpha]} Y_{n,j} > ([n^\alpha] - i + 1)F(p_n) + \gamma \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) < M^*([n^\alpha] - i + 1)/\gamma^2$ a.s.

(ii) $P(\sum_{j=i}^{[n^\alpha]} Y_{n,j} < ([n^\alpha] - i + 1)G(p_n) - \gamma \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) < M^*([n^\alpha] - i + 1)/\gamma^2$ a.s.

Proof. By symmetry it is sufficient to prove (i). Note that if M_1 and M_2 are constants satisfying (i) and (ii) respectively, then $M^* = \max[M_1, M_2]$ will satisfy both assertions. Assume without loss of generality that $F(p_n) = 0$, so that (5.4) and (i) are to be read as $G(p_n) \leqq E(Y_{n,j} \mid Y_{n,j-1}, \cdots, Y_{n,1}, p_n) \leqq 0$ a.s. and $P(\sum_{j=i}^{[n^\alpha]} Y_{n,j} > \gamma \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) < M^*([n^\alpha] - i + 1)/\gamma^2$ a.s., respectively. If in addition $i = 1$, then (i) is simply to be read as

$$P(\textstyle\sum_{j=1}^{[n^\alpha]} Y_{n,j} > \gamma \mid p_n) < M^*[n^\alpha]/\gamma^2.$$

M, as defined in (5.1), will play the role of M^* in assertion (i).

If X is a random variable, we define the random variable X^+ as follows:

$$X^+ = X, \quad \text{if } X \geqq 0, \qquad X^+ = 0, \quad \text{otherwise.} \tag{5.5}$$

Note that if X and Y are random variables, then we have:

$$(X + Y)^+ \leqq |X^+ + Y|. \tag{5.6}$$

We have only to prove:

$$P((\textstyle\sum_{j=i}^{[n^\alpha]} Y_{n,j})^+ > \gamma \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) < M([n^\alpha] - i + 1)/\gamma^2 \quad \text{a.s.} \tag{5.7}$$

From a trivial extension of Chebyshev's inequality, it is sufficient to prove:

$$E(((\textstyle\sum_{j=i}^{[n^\alpha]} Y_{n,j})^+)^2 \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) \leqq M([n^\alpha] - i + 1) \quad \text{a.s.} \tag{5.8}$$

Now:

$$E((Y_{n,i}^+)^2 \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) \leqq M \quad \text{a.s.} \qquad \text{by (5.3)} \tag{5.9}$$

Thus (5.8) is automatic if $i = [n^\alpha]$. If $i < [n^\alpha]$, then for all k ($i \leqq k < [n^\alpha]$) we have:

$$\begin{aligned}
&E(((\textstyle\sum_{j=i}^{k+1} Y_{n,j})^+)^2 \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) \\
&\quad \leqq E(((\textstyle\sum_{j=i}^{k} Y_{n,j})^+ + Y_{n,k+1})^2 \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) \quad \text{a.s.} \quad \text{by (5.6)} \\
&\quad = E(E(((\textstyle\sum_{j=i}^{k} Y_{n,j})^+)^2 + 2(\textstyle\sum_{j=i}^{k} Y_{n,j})^+(Y_{n,k+1}) \\
&\qquad\qquad + (Y_{n,k+1})^2) \mid Y_{n,k}, \cdots, Y_{n,i-1}, \cdots, Y_{n,1}, p_n) \quad \text{a.s.} \\
&\quad \leqq E(((\textstyle\sum_{j=i}^{k} Y_{n,j})^+)^2 \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) + M \quad \text{a.s.} \\
&\qquad\qquad\qquad \text{by (5.3) and (5.4).}
\end{aligned} \tag{5.10}$$

Hence, by recursion, we have:

$$E(((\textstyle\sum_{j=i}^{k+1} Y_{n,j})^+)^2 \mid Y_{n,i-1}, \cdots, Y_{n,1}, p_n) \leqq (k - i + 2)M \quad \text{a.s.} \tag{5.11}$$

(5.8) follows by setting $k = [n^\alpha] - 1$.

Lemma 5.2. *If* $0 < \epsilon < (\alpha - 1)/2$, *then*

$$\Pr(\textstyle\sum_{j=1}^{[n^\alpha]} Y_{n,j} > [n^\alpha]F(p_n) + n^{\alpha-\epsilon} \quad \text{i.o.}) = 0.$$

Proof. Set $i = 1$, and $\gamma = n^{\alpha-\epsilon}$ in Lemma 5.1. Since $\alpha > 1$, according to (3.3), $\sum_{n=1}^{\infty} [n^\alpha]/\gamma^2 < \infty$. The lemma follows directly from the Borel-Cantelli lemma [3].

Lemma 5.3. *There exists a constant* C *satisfying*:

$$\Pr(\max_{k \leqq [n^\alpha]} \textstyle\sum_{j=1}^{k} Y_{n,j} > Cn^\alpha \quad \text{i.o.}) = 0.$$

Proof. Let $M' = \sup_{(p,q)\in A^* \times B^*} |R(p, q)|$. Let n and $\gamma > 0$ be fixed. Let n^* denote the first i ($1 \leqq i \leqq [n^\alpha]$) for which $\sum_{j=1}^{n^*} Y_{n,j} > 2(n^\alpha M' + \gamma)$, if such exists. Then:

$$\begin{aligned}
&P(\textstyle\sum_{j=1}^{[n^\alpha]} Y_{n,j} > ([n^\alpha]M' + \gamma)) \\
&\quad \geqq \textstyle\sum_{i=1}^{[n^\alpha]} P(n^* = i)P(\textstyle\sum_{j=i+1}^{[n^\alpha]} Y_{n,j} \geqq -([n^\alpha]M' + \gamma) \mid n^* = i) \\
&\quad \geqq (1 - (M^*[n^\alpha])/\gamma^2)P(\max_{k \leqq [n^\alpha]} \textstyle\sum_{j=1}^{k} Y_{n,j} > 2(n^\alpha M' + \gamma)) \\
&\qquad\qquad \text{for some positive constant } M^* \text{ by Lemma 5.1.}
\end{aligned} \tag{5.12}$$

Let $0 < \epsilon < (\alpha - 1)/2$, and set $\gamma = n^{\alpha-\epsilon}$. Then:

$$(5.13)\quad \begin{aligned} &P(\max_{k \leqq [n^\alpha]} \textstyle\sum_{j=1}^{k} Y_{n,j} > 2(n^\alpha M' + n^{\alpha-\epsilon})) \\ &\qquad = O(P(\textstyle\sum_{j=1}^{[n^\alpha]} Y_{n,j} > [n^\alpha]M' + n^{\alpha-\epsilon})) \quad \text{by (5.12)} \\ &\qquad = O(n^{2\epsilon-\alpha}) \quad \text{by Lemma 5.1.} \end{aligned}$$

By the Borel-Cantelli lemma, since $\alpha > 1$, we have for all ϵ satisfying $0 < \epsilon < (\alpha - 1)/2$:

$$(5.14)\quad P(\max_{k \leqq [n^\alpha]} \textstyle\sum_{j=1}^{k} Y_{n,j} > 2(n^\alpha M' + n^{\alpha-\epsilon}) \quad \text{i.o.}) = 0.$$

Lemma 5.3 is an immediate corollary of (5.14).

Define N_k and L_k as follows:

$$(5.15)\quad N_k = \textstyle\sum_{n=n_{k-1}+1}^{n_k} [n^\alpha],$$

$$(5.16)\quad L_k = \max_{n_{k-1}+1 \leqq l \leqq n_k} \max_{1 \leqq i \leqq [l^\alpha]} (\textstyle\sum_{n=n_{k-1}+1}^{l-1} \sum_{j=1}^{[n^\alpha]} Y_{n,j} + \sum_{j=1}^{i} Y_{l,j}).$$

Note that N_k denotes the total number of subgames comprising the kth stage, and that L_k denotes the maximum loss incurred by player I during the kth stage.

Recall that α, β, and m are assumed fixed. For all $\gamma > 0$ define $\psi(\gamma)$ as follows:

$$(5.17)\quad \psi(\gamma) = \max [1 - \beta/((\alpha + 1)(m - 1)), (\alpha + 3 + 2\gamma)/(2\alpha + 2), (\alpha + \beta)/(\alpha + 1)]$$

LEMMA 5.4. *Assume $v_G = 0$. Then there exists a constant H such that for all $\gamma > 0$ we have:*

$$P(L_k > HN_k^{\psi(\gamma)} \quad \text{i.o.}) = 0.$$

PROOF. One can show that Lemma 5.2 implies for all $\gamma > 0$:

$$(5.18)\quad P(\textstyle\sum_{j=1}^{[n^\alpha]} Y_{n,j}/[n^\alpha] > F(p_n) + n^{(1-\alpha)/2+\gamma} \quad \text{i.o.}) = 0.$$

Since $v_G = 0$ by assumption, (1.2) and (4.2) imply the existence of a $p' \varepsilon A^*$ satisfying $F(p') = 0$. By Lemma 4.2 there exists for every k a strategy $g_{p'}^{(k)} \varepsilon S_k$ (the available probabilities for the kth stage) satisfying $d(p', g_{p'}^{(k)}) < 2^{-k}$. By Lemma 4.1, we have for all k: $F(g_{p'}^{(k)}) < c2^{-k}$ where

$$c = m^{\frac{1}{2}} \max_{1 \leqq j \leqq m} \sup_{q \varepsilon B^*} |R(a_j, q)|.$$

(5.18) implies the existence of a random variable k^*, such that if $k > k^*$, then for all $\gamma > 0$ we have:

$$(5.19)\quad n_{k-1} + 1 \leqq n \leqq n_k \quad \text{and} \quad p_n = g_{p'}^{(k)} \Rightarrow \textstyle\sum_{j=1}^{[n^\alpha]} Y_{n,j}/[n^\alpha] \leqq c2^{-k} + n^{(1-\alpha)/2+\gamma}$$

i.e., if $k > k^*$, then for all $\gamma > 0$, player I will not incur an average loss greater than $c2^{-k} + n^{(1-\alpha)/2+\gamma}$ on any interval within the kth stage in which $g_{p'}^{(k)}$ is used.

For all $\gamma > 0$, define the set $T(k, \gamma)$ as follows:

(5.20) $T(k,\gamma) = \{n \mid n_{k-1} + 1 \leqq n \leqq n_k \cap \sum_{j=1}^{[n^\alpha]} Y_{n,j}/[n^\alpha] > c2^{-k} + n^{(1-\alpha)/2+\gamma}\}$.

(5.15), (5.16), (5.20), and Lemma 5.3 imply the existence of constants c and C such that for all $\gamma > 0$, we have:

(5.21) $P(L_k > N_k(c2^{-k} + n_k^{(1-\alpha)/2+\gamma}) + \sum_{n\varepsilon T(k,\gamma)} \sum_{j=1}^{[n^\alpha]} Y_{n,j} + Cn_k^\alpha \text{ i.o.}) = 0$.

Let k be fixed, and assume $k > k^*$ so that (5.19) is satisfied. Each of the members of S_k is played in one and only one of the intervals numbered $n_{k-1} + 1$ through $n_{k-1} + \Phi_k$. (5.19) implies for all $\gamma > 0$, that in the intervals numbered $n_{k-1} + \Phi_k + 1$ through n_k a strategy $p'' \varepsilon S_k$ other than $g_{p'}^{(k)}$ may be used only if the greatest average loss incurred in any interval in which p'' is used is not greater than $c2^{-k} + n^{(1-\alpha)/2+\gamma}$. Since S_k contains Φ_k elements, for all $\gamma > 0$, $T(k,\gamma)$ will not contain more than $\Phi_k - 1$ elements. (5.21), Lemma 5.3, and the contents of this paragraph imply that for some constants c and C, and all $\gamma > 0$, we have:

(5.22) $$P(L_k > N_k(c2^{-k} + n_k^{(1-\alpha)/2+\gamma}) + \Phi_k C n_k^\alpha \text{ i.o.}) = 0.$$

In order that (5.22) may imply Lemma 5.4, a careful study of several interesting inequalities is required. (3.1) and (3.2) yield:

(5.23) $$n_k = \sum_{j=1}^k [\Phi_k^{1/\beta}] < \sum_{j=1}^k (m(2^j + 1))^{(m-1)/\beta} = O(2^{k(m-1)/\beta})$$

and

(5.24) $$n_k^\beta \geqq \Phi_k \sim (2^k m)^{m-1}/(m-1)!$$

Hence:

(5.25) $$\begin{aligned} n_k^{\alpha+1} &= n_k n_k^\alpha \\ &= O[(2^{k(m-1)/\beta})(2^{\alpha k(m-1)/\beta})] && \text{by (5.23)} \\ &= O([\Phi_k^{1/\beta}][n_{k-1}^\alpha]) && \text{by (5.24)} \\ &= O(N_k) && \text{by (3.2) and (5.15).} \end{aligned}$$

(5.15) also implies:

(5.26) $$N_k = O(n_k^{\alpha+1}).$$

(5.23) and (5.26) yield:

(5.27) $$2^{-k} = O(n_k^{-\beta/(m-1)}) = O(N_k^{-\beta/((m-1)(\alpha+1))}).$$

(5.24) and (5.25) yield:

(5.28) $$n_k = O(N_k^{1/(\alpha+1)})$$

and

(5.29) $$\Phi_k \leqq (n_k^\beta) = O(N_k^{\beta/(\alpha+1)}).$$

Lemma 5.4 follows directly from (5.17), (5.22), (5.27), (5.28) and (5.29).

THEOREM 5.1. *Suppose players* I *and* II *are playing a sequence of identical pseudo-games* G *satisfying* (i) *and* (ii):

(i) *Player* I *has* $m \geqq 2$ *purely strategy choices.*

(ii) *The distributions* $P_{(a,b)}$ *have uniformly bounded second moments and for each* $a \varepsilon A$ *and every Borel set* C, $P_{(a,\cdot)}(C)$ *is* $\mathcal{B}$*-measurable. Then there exists a class of rules* $\{P\}_m$ *for player* I *such that for all rules* Q *that player* II *may use we have*:

$$P \varepsilon \{P\}_m \Rightarrow \Pr\,(\limsup_{N\to\infty} N^{-1} \textstyle\sum_{j=1}^{N} X_j \leqq v_G \mid P, Q) = 1.$$

PROOF. Since α and β satisfy (3.3) and m is $\geqq 2$, (5.17) implies that for sufficiently small $\gamma > 0$ we will have $\psi(\gamma) < 1$. Thus it is sufficient to prove that there exists a constant K such that for all $\gamma > 0$ we have:

$$P(\textstyle\sum_{j=1}^{N} X_j > Nv_G + KN^{\psi(\gamma)} \quad \text{i.o.}) = 0. \tag{5.30}$$

We lose no generality by assuming that $v_G = 0$ and proving instead that for some constant K and all $\gamma > 0$ we have:

$$P(\textstyle\sum_{j=1}^{N} X_j > KN^{\psi(\gamma)} \quad \text{i.o.}) = 0. \tag{5.31}$$

Lemma 5.4 implies the existence of a constant H and a random variable k^* such that for all $\gamma > 0$ we have:

$$k > k^* \Rightarrow L_k \leqq HN_k^{\psi(\gamma)}. \tag{5.32}$$

By (5.32), for all $k > k^*$ and $\gamma > 0$ we have:

$$\textstyle\sum_{j=1}^{k} L_j \leqq \sum_{j=1}^{k^*} L_j + \sum_{j=k^*+1}^{k} HN_j^{\psi(\gamma)}. \tag{5.33}$$

But:

$$P(\textstyle\sum_{j=1}^{k^*} L_j < \infty) = 1. \tag{5.34}$$

For all $\gamma > 0$, (5.33) and (5.34) yield:

$$P(\textstyle\sum_{j=1}^{k} L_j > 2H \sum_{j=1}^{k} N_j^{\psi(\gamma)} \quad \text{i.o.}) = 0. \tag{5.35}$$

Now:

$$\textstyle\sum_{j=1}^{k-1} N_j < N \leqq \sum_{j=1}^{k} N_j \Rightarrow \sum_{j=1}^{N} X_j \leqq \sum_{j=1}^{k} L_j. \tag{5.36}$$

Let $\inf_\gamma \psi(\gamma) = r_0 > 0$. By (5.35) and (5.36), in order to prove (5.31) we need only prove that there exists a constant C such that for all N and r ($r_0 \leqq r < 2r_0$) we will have:

$$\textstyle\sum_{j=1}^{k-1} N_j < N \leqq \sum_{j=1}^{k} N_j \Rightarrow \sum_{j=1}^{k} N_j^{r} < CN^{r}. \tag{5.37}$$

To prove (5.37) it is in turn sufficient to prove the existence of a constant C such that for all k and r ($r_0 \leqq r < 2r_0$) we will have:

$$\textstyle\sum_{j=1}^{k} N_j^{r} < CN_{k-1}^{r}. \tag{5.38}$$

But r_0 determines constants C_1, C_2, and C_3 that satisfy for all k and $r_0 \leqq r < 2r_0$:

$$
\begin{aligned}
\textstyle\sum_{j=1}^{k} N_j^r &< C_1 \textstyle\sum_{j=1}^{k} 2^{rj(m-1)(\alpha+1)/\beta} && \text{by (5.23) and (5.26)} \\
&< C_2(2^{(k-1)(m-1)(\alpha+1)/\beta})^r && \\
&< C_3 N_{k-1}^r && \text{by (5.24) and (5.25).}
\end{aligned} \tag{5.39}
$$

(5.39) completes the proof.

It is interesting to find the optimal choice of α and β within the approximation given by (5.30). We therefore choose α and β to minimize max $[1 - \beta/((\alpha + 1) \cdot (m - 1)), (\alpha + 3)/(2\alpha + 2), (\alpha + \beta)/(\alpha + 1)]$ and prove:

THEOREM 5.2. *For all rules Q that player* II *may use and all $\epsilon > 0$ we have*:

$$\Pr\left(\textstyle\sum_{j=1}^{N} X_j > N v_G + N^{(2m+1)/(2m+2)+\epsilon} \text{ i.o.} \mid P((m + 2)/m, (m - 1)/m), Q\right) = 0.$$

PROOF. Throughout this proof we restrict our attention to values of $m \geqq 2$, $\alpha > 1$, and $0 < \beta < 1$, so that (3.3) is satisfied. To prove Theorem 5.2 it is sufficient to prove:

$$
\begin{aligned}
&\inf_{\alpha,\beta} (\max [1 - \beta/((\alpha + 1)(m - 1)), \\
&\qquad (\alpha + 3)/(2\alpha + 2), (\alpha + \beta)/(\alpha + 1)]) = (2m + 1)/(2m + 2)
\end{aligned} \tag{5.40}
$$

and that the minimum value is obtained for $\alpha = (m + 2)/m > 1$ and $\beta = (m - 1)/m < 1$. We have:

$$
\begin{aligned}
&\inf_{\alpha,\beta} [\max [1 - \beta/((\alpha + 1)(m - 1)), (\alpha + 3)/(2\alpha + 2), \\
&\qquad (\alpha + \beta)/(\alpha + 1)]] = \inf_{\alpha} [\max ((\alpha + 3)/(2\alpha + 2), \\
&\qquad\qquad \inf_{\beta} [\max (1 - \beta/((\alpha + 1)(m - 1)), (\alpha + \beta)/(\alpha + 1))])].
\end{aligned} \tag{5.41}
$$

Now note the following: For fixed β, both $1 - \beta/(\alpha + 1)(m - 1)$ and $(\alpha + \beta)/(\alpha + 1)$ are strictly increasing functions of α. For fixed α, $1 - \beta/((\alpha + 1)(m - 1))$ is a strictly decreasing function of β, while $(\alpha + \beta)/(\alpha + 1)$ is a strictly increasing function of β. Therefore

$$\inf_{\beta} [\max (1 - \beta/((\alpha + 1)(m - 1)), (\alpha + \beta)/(\alpha + 1))]$$

is a strictly increasing function of α. Also $(\alpha + 3)/(2\alpha + 2)$ is a strictly decreasing function of α. This implies that the minimum occurs when the functions are equal, if such occurs.

The reader can verify that for $\alpha_0 = (m + 2)/m$ and $\beta_0 = (m - 1)/m$ we have:

$$
\begin{aligned}
1 - \beta_0/((\alpha_0 + 1)(m - 1)) &= (\alpha_0 + 3)/(2\alpha_0 + 2) \\
&= (\alpha_0 + \beta_0)/(\alpha_0 + 1) = (2m + 1)/(2m + 2).
\end{aligned} \tag{5.42}
$$

The contents of the previous paragraph imply that this solution is unique and that (5.40) follows from (5.41) and (5.42).

6. Remarks.

1. Recall that $\Phi_k = O(2^{k(m-1)})$. The author experimented with rules in which Φ_k satisfied: (a) $\Phi_k = O(k^{m-1})$ or (b) $\Phi_k = O(e^{ab^k})$, $a > 1$ and $b > 1$. In both

cases the main result was obtainable. However the method of proof presented in this paper was not sufficient to produce results as good as those obtained in (5.30) and Theorem 5.2.

2. The method of proof presented required α to be greater than 1. Suppose we remove this restriction and suppose that the distributions $P_{(a,b)}$ are allowed to have $m + 2 + \epsilon$ uniformly bounded moments for some $\epsilon > 0$. The author believes that it is possible to prove the following stronger versions of (5.30) and Theorem 5.2:

(a) There exists a constant H such that for every $\gamma > 0$ we have:

$$\Pr\left(\sum_{j=1}^{N} X_j > Nv_G + HN^{\max[1-\beta/(\alpha+1)(m-1),\,(\alpha+2+\gamma)/2(\alpha+1),\,(\alpha+\beta)/(\alpha+1)]} \text{ i.o.}\right) = 0.$$

(b) For all rules Q that player II may use and all $\epsilon > 0$ we have:

$$\Pr\left(\sum_{j=1}^{N} X_j > Nv_G + N^{(m+1)/(m+2)+\epsilon} \text{ i.o. } \mid P(2/m, (m-1)/m), Q\right) = 0.$$

He will endeavor to prove these results in a later paper.

3. Suppose player II is using a fixed strategy $q_0 \varepsilon B^*$ throughout the sequence of games and player I is using a strategy $P \varepsilon \{P\}_m$, as described in Section 3. Then

$$\Pr\left(\limsup N^{-1} \sum_{i=1}^{N} X_i \leqq \inf_{p \varepsilon A^*} R(p, q_0)\right) = 1.$$

Thus every $P \varepsilon \{P\}_m$ is asymptotically Bayes with respect to q_0. If player I uses $P((m + 2)/m, (m - 1)/m)$, then the convergence rate given in Theorem 5.2 is attained.

4. The argument presented in (5.12) can be extended to prove the following generalization of Skorohod's inequality [5]:

THEOREM. *Let $Z_1, Z_2, Z_3, \cdots$ be a sequence of random variables. Let $S_k =_{\text{def}} \sum_{j=1}^{k} Z_j$. Let n be a fixed positive integer, $\lambda > 0$ a fixed constant, and let n^* be the first integer k such that $|S_k| > 2\lambda$. Suppose that for some $c \leqq 1$ and all $k \leqq n$ we have $P(|S_n - S_k| > \lambda \mid n^* = k) \leqq c$. Then $(1 - c)P(\max_{k \leqq n} |S_k| > 2\lambda) \leqq P(|S_n| > \lambda)$.*

PROOF.

$$\begin{aligned} P(|S_n| > \lambda) &\geqq \sum_{k=1}^{n} P(n^* = k, |S_n - S_k| \leqq \lambda) \\ &= \sum_{k=1}^{n} P(n^* = k) P(|S_n - S_k| \leqq \lambda \mid n^* = k) \\ &\geqq (1 - c)P(\max_{k \leqq n} |S_k| > 2\lambda). \end{aligned}$$

COROLLARY (*Skorohod's Inequality*). *Let $Z_1, Z_2, Z_3, \cdots$ be a sequence of independent random variables and let $S_k =_{\text{def}} \sum_{j=1}^{k} Z_j$. Let n be a fixed positive integer and suppose that for some pair (λ, C) of positive constants and all $k \leqq n$ we have $P(|S_n - S_k| > \lambda) \leqq C < 1$. Then $P(\max_{k \leqq n} |S_k| > 2\lambda) \leqq P(|S_n| > \lambda)/ (1 - C)$.*

5. Convergence problems similar to the one presented here have been studied by Feldman [1], Samuel [4], and Van Ryzin [6], [7]. However in all cases at least a partial knowledge of the payoff function is assumed. Harsanyi [2] has studied

the problem of reducing the analysis of a game with incomplete information G to that of a game with complete information G^* equivalent to G.

7. Acknowledgment. The author would like to thank Professor Thomas S. Ferguson for invaluable advice throughout the preparation of this paper, Professor Charles J. Stone for some very helpful comments, and the referee for suggestions that considerably improved an earlier version.

REFERENCES

[1] FELDMAN, DORIAN (1962). Contributions to the "two-armed bandit" problem. *Ann. Math. Statist.* **33** 847–856.

[2] HARSANYI, J. C. (1965). Games with incomplete information played by "Bayesian" players, I–II. Papers presented at the Fifth Princeton Conference on Game Theory.

[3] LOÈVE, MICHEL (1963). *Probability Theory.* Van Nostrand, Princeton.

[4] SAMUEL, ESTER (1963). Asymptotic solutions of the Sequential Compound Decision problem. *Ann. Math. Statist.* **34** 1079–1097.

[5] SKOROHOD, A. V. (1957). Limit theorems for stochastic processes with independent increments. *Theor. Prob. Appl.* **2** 138–171.

[6] VAN RYZIN, J. (1966). The sequential compound decision problem with $m \times n$ finite loss matrix. *Ann. Math. Statist.* **37** 954–975.

[7] VAN RYZIN, J. (1966). Repetitive play in finite statistical games with unknown distributions. *Ann. Math. Statist.* **37** 976–994.

[8] VON NEUMANN, J. and MORGENSTERN, O. (1953). *Theory of Games and Economic Behavior.* Princeton Univ. Press.