adversary's commodities without, however, allowing any of his stocks to vanish.

This is a recursive game, as we can see by defining player 1's payoff on trial k to be either the game with the resources remaining after the game on trial k, provided neither player has lost all of any one of his commodities, or 1 if any of player 2's stocks go to zero, or 0 if none of 2's stocks go to zero and at least one of 1's do. (Observe that the convention has been made that player 1 "wins" whenever both players simultaneously exhaust a commodity.) The play terminates when either a 0 or a 1 payoff occurs. To guard against infinite play, Blackwell requires that at least one commodity be diminished in each engagement and that no resupply ever occurs. Stated formally, for every (i, j) pair,

$$\sum_{r=1}^{R} \alpha_r(i, j) + \sum_{s=1}^{S} \beta_s(i, j) > 0$$

and

$$\alpha_r(i, j) \geqslant 0, \qquad \text{for every } r,$$

and

$$\beta_s(i, j) \geqslant 0, \qquad \text{for every } s.$$

In sum, a multicomponent attrition game is described by two complexes of information: the initial resources $(\mathbf{a}^{(0)}, \mathbf{b}^{(0)})$, and the attrition matrix $[(\boldsymbol{\alpha}(i, j), \boldsymbol{\beta}(i, j)]$, which we shall abbreviate simply as $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Each entry of the matrix $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is an $(R + S)$-tuple, the first R components of the (i, j) entry $[\alpha_1(i, j), \cdots, \alpha_R(i, j)]$, being designated by $\boldsymbol{\alpha}(i, j)$, and the last S components, $[\beta_1(i, j), \cdots, \beta_S(i, j)]$, by $\boldsymbol{\beta}(i, j)$. Since the payoffs have been chosen to be 0 and 1, the value of the game is merely the probability, which we denote by $P[\boldsymbol{\alpha}, \boldsymbol{\beta}; \mathbf{a}^{(0)}, \mathbf{b}^{(0)}]$, that player 1 wins when both players use their optimal strategies. Since multicomponent attrition games are special cases of recursive games, we know that, when P is treated as a function of $(\mathbf{a}^{(0)}, \mathbf{b}^{(0)})$, with $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ held constant, it must satisfy the basic functional equation of stochastic and recursive games. Of course, even in special instances, that equation is monstrous, and Blackwell does not attempt to solve it as such. Rather, he investigates the asymptotic behavior of P as the resources $(\mathbf{a}^{(0)}, \mathbf{b}^{(0)})$ are increased indefinitely subject to the condition that their relative sizes are fixed. For example, one can look for the set of $(\mathbf{a}^{(0)}, \mathbf{b}^{(0)})$ pairs such that

$$\lim_{t \to \infty} P[\boldsymbol{\alpha}, \boldsymbol{\beta}; t\mathbf{a}^{(0)}, t\mathbf{b}^{(0)}] = 1,$$

where, of course, $t\mathbf{a}^{(0)} = (ta_1^{(0)}, ta_2^{(0)}, \cdots, ta_R^{(0)})$ and similarly for $t\mathbf{b}^{(0)}$. For the women and cats versus men and mice example, Blackwell shows that

$$\lim_{t \to \infty} P\left\{\begin{bmatrix} (0, 0; 1, 0) & (1, 0; 0, 0) \\ (0, 1; 0, 0) & (0, 0; 0, 1) \end{bmatrix}, (ta_1^{(0)}, ta_2^{(0)}; tb_1^{(0)}, tb_2^{(0)})\right\} = 1,$$

provided that

$$a_1^{(0)}a_2^{(0)} > b_1^{(0)}b_2^{(0)}.$$

Note, for example, that (0, 0; 1, 0), which is the (1, 1) entry of matrix $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, has the interpretation: team 1 loses zero women and zero cats and team 2 loses one man and zero mice.

A8.6 APPROACHABILITY-EXCLUDABILITY THEORY AND COMPOUND DECISION PROBLEMS

The asymptotic theory of multicomponent attrition games is based on Blackwell's [1956 *a*] analogue of the minimax theorem for games with

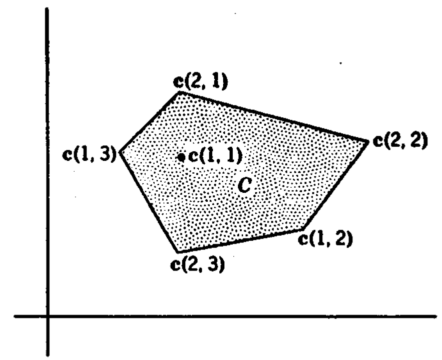

Fig. 2

vector payoffs. In such games, the players have m and n pure strategies as usual, but the payoff corresponding to the (i, j) strategy pair is a Q-tuple (or vector in Q-space) of the form $\mathbf{c}(i, j) = [c_1(i, j), c_2(i, j), \cdots, c_Q(i, j)]$. The multicomponent attrition games are of this form with $Q = R + S$ and $\mathbf{c}(i, j)$ equal to the attrition payoff, but, as we shall see below, quite different interpretations of vector games also exist.

Let us denote by C the convex hull of the set of points (in Q-space) $\mathbf{c}(i, j)$, where i and j vary over their domains. For example, if $Q = 2$, $m = 2$, and $n = 3$, then a typical region C is shown in Fig. 2. Blackwell raises this question: *If such a game is repeated in time, can player* 1 *force the average payoff to approach a preassigned closed subset T of C? Equally well, when can player* 2 *exclude the average payoff from T?*

The following notation will be useful. Let $\mathbf{x} = (x_1, x_2, \cdots, x_m)$ be one of player 1's mixed strategies on a component game; then if player 2 uses pure strategy j, the expected payoff will be

$$\mathbf{c}(\mathbf{x}, j) = \sum_i x_i \mathbf{c}(i, j).$$

Thus, his expected payoff when he uses $\mathbf{x}$ will lie in the smallest convex set containing the n points $\mathbf{c}(\mathbf{x}, j)$, $j = 1, 2, \cdots, n$; we denote this set by $C(\mathbf{x}, \cdot)$. Exactly parallel notation $[\mathbf{y}, \mathbf{c}(i, \mathbf{y})$, and $C(\cdot, \mathbf{y})]$ is introduced for player 2. Finally, the average payoff for k trials is denoted

$$\bar{\mathbf{c}}^{(k)} = [\mathbf{c}(i_1, j_1) + \mathbf{c}(i_2, j_2) + \cdots + \mathbf{c}(i_k, j_k)]/k,$$

where (i_h, j_h) denotes the strategy pair chosen on trial h.

We observe that a sufficient condition for T to be excludable by player 2 is the existence of a strategy $\mathbf{y}^{(0)}$ such that $C(\cdot, \mathbf{y}^{(0)})$ is disjoint from T, for if $\mathbf{y}^{(0)}$ is used at each trial the average payoff will approach $C(\cdot, \mathbf{y}^{(0)})$, and so not T. Blackwell shows, in essence, that this is a necessary condition too. To be more precise: *any convex set T is either approachable by* 1 *or excludable by* 2, *and the latter is equivalent to the existence or a $\mathbf{y}^{(0)}$ such that T and $C(\cdot, \mathbf{y}^{(0)})$ are disjoint.* Furthermore, he displays a strategy for player 1 which will force the average payoff to approach T whenever such a strategy exists.

The idea is simple. If at trial k, the average payoff $\bar{\mathbf{c}}^{(k)}$ is already in T, select any $\mathbf{x}$ on trial $k + 1$. If, however, $\bar{\mathbf{c}}^{(k)}$ and T are disjoint, choose $\mathbf{x}$ so that $C(\mathbf{x}, \cdot)$ and T lie on the same side of the supporting hyperplane of T which both passes through the point $\mathbf{c}'$ of T that is closest to $\bar{\mathbf{c}}^{(k)}$ and which is perpendicular to the line joining these two points. (See Fig. 3.) Such an $\mathbf{x}$ can be shown to exist if and only if the convex set T is not excludable by 2. (Roughly the idea is this: Suppose 1 tries to get an expected payoff which lies as far below the separating hyperplane as possible. Player 2 cannot guarantee that 1 will not get a point on or below this hyperplane since $C(\cdot, \mathbf{y})$ intersects T for all $\mathbf{y}$. Now we invoke the usual form of the minimax theorem to conclude that 1 can therefore guarantee a point on or below the hyperplane.) Since the expected payoff $\mathbf{c}^*$ on trial $k + 1$ will, of course, be in $C(\mathbf{x}, \cdot)$, let us, for heuristic reasons, simplify the argument by supposing that the actual payoff on trial $k + 1$ is the point $\mathbf{c}^*$ in $C(\mathbf{x}, \cdot)$; then the average payoff $\bar{\mathbf{c}}^{(k+1)}$ will lie on the line joining $\bar{\mathbf{c}}^{(k)}$ and $\mathbf{c}^*$. If k is large, $\bar{\mathbf{c}}^{(k+1)}$ will be much nearer to $\bar{\mathbf{c}}^{(k)}$ than to $\mathbf{c}^*$, and so it will be nearer to $\mathbf{c}'$ than $\bar{\mathbf{c}}^{(k)}$ is. This suggests that in time the average payoff will approach T. As yet, however, the argument is not tight for we have been dealing with expected values at a given trial, whereas the approachability theorem asserts something about the time sequence $\{\bar{\mathbf{c}}^{(k)}\}$ being true with probability 1. This gap is bridged by a probability existence theorem that we will not discuss except to remark that it is similar in spirit to the martingale theorem which arose in the section on recursive games.

Two points about approachability-excludability theory need clarification: why is it related to the study of multicomponent attrition games, and in what sense is it an analogue of the minimax theory? The first seems to be a problem since we know that multicomponent games are recursive games, whereas the present theory is not cast in that form. But recall that Blackwell confined himself to questions about ruin probabilities when the initial resources are held in fixed proportion and increased without bound. It is thus plausible that each player's ability to control the limiting behavior of the time average of the attrition payoffs will govern the outcome, and in fact it does.

Next, let us turn to the sense in which the theory generalizes the minimax theorem. Suppose that the payoffs $\mathbf{c}(i, j)$ are actually real numbers,

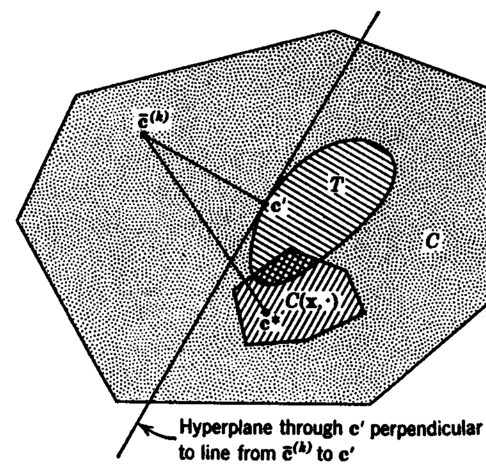

FIG. 3

i.e., $Q = 1$, and that they are interpreted as 1's payoffs. If we let a denote the minimum and b the maximum of these mn numbers, the set C is simply the interval of the real line from a to b inclusive. If v denotes the value of the game, player 1 can approach the interval $[v, b]$ and player 2 can approach the interval $[a, v]$. Or in more familiar words, using the law of large numbers, the expected value v of a two-person zero-sum game can be given a frequency interpretation as the limiting value of a temporal average.

Earlier we promised a second and important interpretation of the approachability-excludability theory, and it is now time to fulfill it. Let us suppose that a two-person game is to be repeated and that player 1 is solely interested in his long-term average payoff. He can certainly secure a limiting average at least equal to the maximum value of the component game by playing maximin at each stage. But, as we have pointed out previously, it has long been recognized that such a strategy is not very realistic in any of the following cases:

i. In a zero-sum game when player 2 is not a conscious minimaxer.
ii. In a non-zero-sum game.
iii. When player 2 is "nature" in the usual decision problem under uncertainty—the statistical inference problem.

Robbins [1951] has emphasized that when a (statistical) decision problem is repeated in time, e.g., when a stream of individuals must be classified by their individual test responses, the statistician can often do as well asymptotically with no prior information as when he knows the exact limiting proportion of times player 2 uses each strategy. To be more specific, suppose 1's payoffs are a_{ij} and that *a priori* he knows that the proportion of the time player 2 will use strategy j, $j = 1, 2, \cdots, n$, is y_j^*. He can, therefore, achieve the limiting average return

$$\rho(\mathbf{y}^*) = \max_i \left(\sum_j a_{ij} y_j^* \right)$$

by playing that strategy i which maximizes the right-hand expression on each trial. Hannan [1957] shows that asymptotically player 1 can do as well as $\rho(\mathbf{y}^*)$ *without knowing* $\mathbf{y}^*$ *beforehand* provided that he bases his choice at each trial on his knowledge of 2's previous choices and on chance. (Actually, he need only consider 2's empirical mixed strategy over the preceding moves.)

Blackwell [1956 *b*] shows that this can be concluded from approachability-excludability theory. He chooses $Q = n + 1$ and defines

$$\mathbf{c}(i, j) = (0, 0, \cdots, 0, 1, 0, \cdots, 0, a_{ij}),$$

where the 1 appears in the jth position and a_{ij} is the (i, j) payoff of the given game to player 1. This definition may seem strange, but it is less so when one observes that the first n components of $\bar{\mathbf{c}}^{(k)}$ equal player 2's empirical mixed strategy over the first k trials and the last component is 1's average payoff during those trials. Now, let T be the set of all $(n + 1)$-tuples whose first n components represent a probability vector, call it $\mathbf{y}$, and whose last component, c_{n+1}, is at least equal to $\rho(\mathbf{y})$, i.e.,

$$T = \{\text{the set of all } (c_1, c_2, \cdots, c_n, c_{n+1}) \text{ such that } c_j \geqslant 0,$$
$$\text{for } j = 1, 2, \cdots, n, \quad \sum_{j=1}^{n} c_j = 1,$$
$$\text{and } c_{n+1} \geqslant \sum_{j=1}^{n} a_{ij} c_j, \quad \text{for } i = 1, 2, \cdots, m\}.$$

The result is proved if we can show that T is approachable by 1, for, if it is approachable, then with any limiting distribution $\mathbf{y}^*$ player 1 receives a limiting average value of at least $\rho(\mathbf{y}^*)$. Note that we do not necessarily assume that the empirical mixed strategy over the first k trials, $\mathbf{y}^{(k)}$, approaches a limit as $k \to \infty$. When the limit does not exist, the result is interpreted roughly as meaning that the average payoff for large k will be close to $\rho(\mathbf{y}^{(k)})$.

The approachability of T follows from the observation that, for each $\mathbf{y}$, the set $C(\cdot, \mathbf{y})$ just touches T. This we can see as follows: If $\mathbf{y} = (y_1, y_2, \cdots, y_n)$, then $C(\cdot, \mathbf{y})$ is the set of $(n + 1)$-tuples $(y_1, y_2, \cdots, y_n, c_{n+1})$, where $\min_i \sum_j a_{ij} y_j \leqslant c_{n+1} \leqslant \max_i \sum_j a_{ij} y_j$, so it intersects T at the point $[y_1, y_2, \cdots, y_n, \rho(\mathbf{y})]$.

The choice of a strategy which leads the average payoff to approach T is far more subtle than it may seem. For example, player 1's "obvious" strategy of playing optimal on trial $k + 1$ against 2's empirical mixed strategy calculated over the first k trials need not force the average payoff to converge to T. Remember that player 2 may not employ the limiting mixed strategy $\mathbf{y}^*$ at every (or indeed, any) of the trials.

Besides this asymptotic result, Hannan [1957] also has a great deal to say about the rates of convergence for certain reasonable classes of player 1's strategies. Other papers which extend the pioneering work of Robbins [1951] on compound statistical decision problems are Hannan and Robbins [1955], Laderman [1955], and Johns [1956].

A8.7 DIVIDEND POLICY AND ECONOMIC RUIN GAMES

Most of the games we have encountered in this appendix meet the following very general description: a known stochastic process is under way, but at periodic intervals two players, perhaps opposing, can exert some influence on the process. Shubik [1957] has pointed out that corporate dividend policy can be looked upon in this way, and he has begun to examine games suggested by this interpretation.

The simplest case is the degenerate single corporation game in which its assets fluctuate from period to period according to a simple chance mechanism. For example, if the capital accumulation is Z units (units in terms of thousands or tens of thousands of dollars) in one period, we might assume that in the next period it becomes $Z + 1$ with probability p, or $Z - 1$ with probability $q = 1 - p$. The corporation is ruined if at any period its capital drops below zero. Clearly, its chance of being ruined within a specified time period is less the greater the capital at the beginning of that period, but, on the other hand, money in the corporate