U. S. AIR FORCE

PROJECT RAND

RESEARCH MEMORANDUM

COMPARISON OF RECONNAISSANCES

David Blackwell

RM-241

15 September 1949

Assigned to _____

This is a working paper. It may be expanded, modified, or withdrawn at any time. The views, conclusions, and recommendations expressed herein do not necessarily reflect the official views or policies of the United States Air Force.

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

COMPARISON OF RECONNAISSANCES

David Blackwell

Summary.

If in a two-person zero-sum game a chance move is introduced, giving one player partial information about the opponent's pure strategy, the value of the game to the player receiving the information is in general increased. Bohnenblust, Shapley, and Sherman (RM-208) have called one reconnaissance better than another if in every game the increase in value by introducing the first reconnaissance is greater than or equal to the increase by introducing the second. For the special case in which the opponent has two moves, we give here a representation of every reconnaissance better than a given reconnaissance. The representation shows that one reconnaissance is better than another only if the information contained in the second can be duplicated from the information contained in the first.

The observation of the outcome of a chance move whose distribution depends on an unknown action of the enemy has been called by Bohnenblust, Sherman, and Shapley (RM-208) a reconnaissance (= experiment in statistical terminology). More precisely, in case the enemy has only two possible actions (= dichotomy), a reconnaissance r is a pair u_1, u_2 of probability measures on a Borel field of subsets of a space X. For a given reconnaissance r, there are two non-negative functions $p_1(x)$, $p_2(x)$ such that

$u_i(S) = \int_S p_i(x)dm$, where $m = \frac{u_1 + u_2}{2}$. The probability measure with cdf $M_r(a) = m\left\{\frac{p_1(x)}{p_1(x) + p_2(x)} < a\right\}$ on the unit interval will be called the measure associated with the reconnaissance r, and the convex function $c(y) = \int_0^y M(x)dx$ will be called the convex function associated with the measure M. We shall denote both the measure and its cdf by M. Bohnenblust, Sherman, and Shapley have proved

Theorem 1: $\int \phi(x)dM_1(x) \geq \int \phi(x)dM_2(x)$ for all convex functions $\phi(x)$ is equivalent to $c_1(y) \geq c_2(y)$ for all y, with equality at $y = 1$.

We shall say that a reconnaissance r is better than a reconnaissance s ($r \supset s$) if the equivalent conditions of the theorem are satisfied.

By a transformation T we mean a class of probability measures $P_y(A)$, $0 \leq y \leq 1$ defined in the unit interval, such that for fixed A, $P_y(A)$ is a Borel function of y, and for each y, $\int x\, dP_y(x) = y$.

If M is any probability measure on the unit interval, we define

$$TM(A) = \int P_y(A)dM(y).$$

Then TM is again a probability measure on the unit interval.

The following theorem has been noted in another context by Doob:

Theorem 2: For all M, T we have $TM \supset M$.

Proof: For any convex function $\phi(x)$ on (0, 1), $\int \phi(x)dTM(x) = \int\left[\int \phi(y)dP_x(y)\right]dM(x)$.

In fact this formula holds for any TM-integrable function $\phi(x)$; it is equivalent to the definition of TM when ϕ is the characteristic function of an event A, and the general result follows by approximation. If ϕ is convex, $\int \phi(y) dP_x(y) \geq \phi\left[\int y dP_x(y)\right] = \phi(x)$, so that $\int \phi(x) dTM(x) \geq \int \phi(x) dM(x)$.

Conversely we shall prove that for any two measures M, N with $M \supset N$ there is a T with $M = TN$. We first prove

Theorem 3: Let $T_1, T_2, \cdots$ be any sequence of transformations. There is a transformation T such that for every measure M, $T_n T_{n-1} \cdots T_1 M \longrightarrow TM$ at every point of continuity of TM.

Proof: Let Ω be the infinite product space of sequences $\omega = \{x_i\}$, $0 \leq x_i \leq 1$. For each x_0 there is a measure $Q_{x_0}(S)$ defined on Borel sets of Ω such that for any function $f(x_0, \cdots, x_n)$,

$$(1) \quad \int f(\omega) dP_{x_0}(\omega) = \int\int \cdots \int f(x_0, \cdots, x_n) dP^n_{x_{n-1}}(x_n \cdots dP^1_{x_0}(x_1),$$

where $P^n_y(A)$ is the transformation T_n.

It follows from (1) that

$$(2) \qquad E(x_n | x_0, \cdots, x_m) = x_m \quad \text{for all } n \geq m,$$

where $E(x|y)$ denotes, as usual, the conditional expectation of x given y. Chance variables $\{x_n\}$ with property (2) are known as martingales and have been investigated by Doob, who has shown that if $\{x_n\}$ is a martingale and if the x_n's are uniformly integrable (trivial here, since $0 \leq x_n \leq 1$), there is a chance variable x such that $x_n \longrightarrow x$

almost everywhere, and $E(x|x_n) = x_n$. Denote the distribution of x by P_{x_0}. Since $E(x|x_0) = x_0$, $\int y dP_{x_0}(y) = x_0$. The collection of measures P_{x_0}, $0 \leq x_0 \leq 1$, then defines a transformation T. If M is any measure over x_0, the measure $Q(S) = \int Q_{x_0}(S) dM(x_0)$ is a measure on Ω. It is easily shown by approximation that for any Q-integrable function f,

$$(3) \qquad \int f(\omega) dQ(\omega) = \int \left[\int f(\omega) dQ_{x_0}(\omega) \right] dM(x_0).$$

Letting f be the characteristic function of an x_n set in (3) and using (1) shows that the distribution of x_n is $T_n \cdots T_1 M$, letting f be the characteristic function of an x set shows that the distribution of x is TM; since $x_n \longrightarrow x$ almost everywhere (Q measure), $T_n \cdots T_1 M \longrightarrow TM$ at every point of continuity of TM.

Let a, b be any subinterval of (0, 1); we consider the transformation T(a, b) with

$$P_x = \frac{b - x}{b - a} I_a + \frac{x - a}{b - a} I_b \quad \text{for } a \leq x \leq b$$

$$P_x = I_x \quad \text{for } x \text{ outside } a, b,$$

where I_x denotes the measure concentrated at x.

Theorem 4: For any measure M, the convex curve $c_{T(a,b)M}$ is obtained from c_M by replacing the part of c_M between the points $\left(a, c_M(a)\right)$, $\left(b, c_M(b)\right)$ by the line segment joining these points.

Proof: For brevity, write $T = T(a, b)$. We have $TM(S) = M(S)$ for S disjoint from (a, b), so that $TM(x) = M(x)$ for x outside (a, b). Then $c_{TM}(y) = \int_0^y TM(x)dx = \int_0^y M(x)dx = c_M(y)$ for $y \leq a$. Now $c_{TM(1)} = c_{TM}(y) + \int_y^1 TM(x)dx \geq c_M(y) + \int_y^1 M(x)dx = c_M(1)$ for $y \geq b$, since $TM \supset M$ implies $c_{TM}(y) \geq c_M(y)$ for all y, and $TM(y) = M(y)$ for $y > b$. Since $c_{TM}(1) = c_M(1)$, $c_{TM}(y) = c_M(y)$ for $y \geq b$. Finally, TM is constant over the open interval a, b, so that c_{TM} is linear there.

Theorem 5: If $M \supset N$, there is a T with $TN = M$.

Proof: $c_M(y)$, $c_N(y)$ are convex curves on 0, 1, with $c_M(y) \geq c_N(y)$ and equality at $y = 0, 1$. Drawing a tangent to c_M at any point and applying the transformation $T(a, b)$ to N, where a, b are the x-coordinates of the points where the tangent intersects c_N yields a new curve c_{N_1} between c_N and c_M. It is clear that applying the same procedure to c_{N_1} and continuing in this way we can, by drawing the tangents on a set of points dense in c_M, obtain a sequence of transformations $T_1, T_2, \cdots$ such that, if $W_n = T_n T_{n-1} \cdots T_1$, $c_{W_n N} \longrightarrow c_M$. There is by Theorem 3 a transformation T such that $W_n N \longrightarrow TN$ at every point of continuity of TN. It follows that $c_{W_n N}(y) = \int_0^y W_n N(x)dx \longrightarrow \int_0^y TN(x)dx$ for all y. Thus $c_{TN}(y) = c_M(y)$ for all y. Differentiation yields $TN(y) = M(y)$ at all points where both are continuous; since both are right continuous, $TN = M$.

Statistical meaning of Theorem 5: The statistical meaning of Theorem 5 is roughly this: one experiment (reconnaissance) is better in our sense than another experiment only if, given the result of r, we can

duplicate s without removing our eyes from the random number table; in statistical language, there is a (random) function of r whose distribution under either hypothesis is the same as that of s. We shall give the details for the case where each reconnaissance has only a finite number of possible outcomes.

Suppose, then,

for r, $X = (1, 2, \cdots, m)$; $u_1 = (p_1, \cdots, p_m)$, $u_2 = (P_1, \cdots, P_m)$;

for s, $Y = (1, 2, \cdots, n)$; $v_1 = (q_1, \cdots, q_n)$, $v_2 = (Q_1, \cdots, Q_n)$.

Then

$$M_r \text{ assigns probability } \frac{p_i + P_i}{2} \text{ to } \frac{p_i}{p_i + P_i}$$

$$M_s \text{ assigns probability } \frac{q_j + Q_j}{2} \text{ to } \frac{q_j}{q_j + Q_j}.$$

If r is better than s, Theorem 5 asserts the existence of a matrix $A = a_{ji}$ such that

$$(4) \qquad \begin{cases} a_{ji} \geq 0, \ \sum_{i=1}^{m} a_{ji} = 1, \\ \sum_{i=1}^{m} a_{ji} \frac{p_i}{p_i + P_i} = \frac{q_j}{q_j + Q_j}, \\ \sum_{j=1}^{n} \frac{q_j + Q_j}{2} a_{ji} = \frac{p_i + P_i}{2}. \end{cases}$$

Define $t_{ij} = \frac{(q_j + Q_j)}{p_i + P_i} a_{ji}$. Then $t_{ij} \geq 0$, and $\sum_{j=1}^{n} t_{ij} = 1$.

We now describe how, given the result of reconnaissance r, a result equivalent to reconnaissance s can be constructed. When r results in x = i, let z = j with probability t_{ij}. We must show that when x has distribution p(P), z has distribution q(Q).

Suppose x has distribution p: $P\{x = i\} = p_i$. Then

$$P\{z = j\} = \sum_{i=1}^{m} p_i t_{ij} = \left(\sum_{i=1}^{m} \frac{p_i a_{ji}}{p_i + P_i}\right)(q_j + Q_j) = q_j.$$

If x has distribution P: $P\{x = i\} = P_i$, then

$$P\{z = j\} = \sum_{i=1}^{m} P_i t_{ij} = \sum_{i=1}^{m} \frac{P_i (q_j + Q_j) a_{ji}}{p_i + P_i} =$$

$$\left[\sum_{i=1}^{m} \left(1 - \frac{p_i}{p_i + P_i}\right) a_{ji}\right](q_j + Q_j) = \left(1 - \frac{q_j}{q_j + Q_j}\right)(q_j + Q_j) = Q_j.$$

The above computations could have been avoided by appealing to the known fact that a Markoff process is also a Markoff process going backwards, and are essentially a derivation of a consequence of this fact in the finite case.

DB:rb