

Approximation to Bayes Risk in Repeated Play

Contributions to the Theory of Games, Vol. III, Annals of Mathematics
Studies No. 39

James Hannan

Princeton 1957

1 Introduction

A finitary form of the (static) decision problem with numerical utilities may be described briefly as follows: exactly one of n possible decisions is required to be made in a context in which the decision-maker knows only that one of m possible states of nature obtains and, for each decision and each state, the inutility a_{ij} of decision j when state i obtains.

The problem of choosing a decision which will in some sense minimize inutility has been resolved by many principles of solution [12], [13], and [16]. These principles have in each case been suggested and to some extent supported by the additional assumptions associated with some class of realizations. In particular the classical Bayes principle introduces the assumption that the state index i is subject to a known (a priori) probability distribution. Under this assumption, which is considered much too restrictive for many realizations, a decision which minimizes expected inutility is a very satisfactory solution of the problem.

The present paper is concerned with a sequence of N decision problems, which are formally alike except for the fact that the state of nature may vary arbitrarily from problem to problem. Decisions are required to be made successively and it is assumed that they may be allowed to depend on the e.d. (empirical distribution) of the states of nature across the previous problems in the sequence. This total lack of assumptions regarding the behavior of the state sequence is a feature distinguishing the present structure from many considerations of multistage processes.

In a certain sense the opportunity for minimizing the average inutility of the set of N decisions depends on the e.d. of the N states involved. If this e.d. were known before any decisions had to be made, this knowledge would enable the choice of a decision Bayes with respect to this distribution. The repeated use of such a decision would reduce the average inutility across problems to the minimum expected inutility on a single problem where the Bayes assumption holds and the probability distribution on the state index is the same as the e.d. of the N states in the sequence problem (see (4.10)).

Another hypothetical situation, somewhat more suggestive for the sequence problem, is that in which successive decisions are permitted to depend on the successive e.d. of all states through their respective presents. The use of decisions Bayes with respect to these distributions reduces average inutility to not more than the Bayes single-problem inutility with respect to the N -state e.d. (see (6.4) ff.).

The most important conclusion of this paper is that the knowledge of the successive e.d. of past states makes it constructively possible to do almost as well at reducing the average inutility across problems as in the case where N and the distribution of the N states are known in advance (or even as in the case of the preceding paragraph). The sequence of decisions exhibited

which attain this performance is a sequence of randomized (not necessarily properly) decisions whose expectations are the values of a sequence of smoothed versions of the Bayes response at the successive e.d. of previous states.

The idea of using the Bayes inutility against the N -state e.d. as a goal for the performance of a set of decisions (distinguished by the term compound decision problem and with stochastic information on the N -state e.d. replacing the knowledge of past e.d. in the sequence problem) was enunciated in [14]. The program outlined in [14] for the rigorous investigation of compound decision problems was initiated in [6] and these papers exerted a strong influence on the sequence development.

The inadequacy of min-max solutions has long been noted in connection with statistical decision problems and [7], [8], [11], [14], [18], [19], are particularly relevant. The fact that this inadequacy is always present in certain compound decision problems was noted in [14] and [6]. An example of a compound problem in which the min-max solution fails to be a direct product is exhibited in [17].

2 The Component Game, G

The main purpose of this section is to introduce the notational framework of a single finite game, later to be used as the generic component of a sequence of games. The single game terminology can also be applied to the normal form of sequence games and for this purpose some extensions of the concept of "regret" (the loss of [16]) are introduced. These may induce interesting orderings in games with sufficient structure and, in particular, will do so for the sequence games of Section 3.

Let G be a finite two-person general game in which players I and II have, respectively, m and n strategies. Their spaces of randomized strategies will be denoted by X and Y ,

$$\begin{aligned} X &= \{x = (x_1, x_2, \dots, x_m) \mid x_i \geq 0, \sum_{i=1}^m x_i = 1\}, \\ Y &= \{y = (y_1, y_2, \dots, y_n) \mid y_j \geq 0, \sum_{j=1}^n y_j = 1\}, \end{aligned} \tag{1}$$

and their pure strategies will be represented by

$$\epsilon = (e_1, e_2, \dots, e_m) \text{ a basis vector in } X, \delta = (\delta_1, \delta_2, \dots, \delta_n) \text{ a basis vector in } Y. \tag{2}$$

In accord with the dominant decision theoretic orientation, the game will be consistently viewed from the position of player II. Nothing will be assumed about player I's motivation and the game will be defined only up to player II's inutility which will be described by a loss matrix A . The elements of A will be denoted by a_{ij} , the rows by A_i and the columns by A^j ,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} = [A^1 \quad A^2 \quad \cdots \quad A^n]. \tag{3}$$

To exclude trivial games, it will be assumed that A has no dominant column for each j ,

$$\max_{\epsilon} \{\epsilon A^j - \min_r \epsilon A^r\} > 0, \tag{4}$$

2.1 Binary Game Matrix Analysis

Consider a two-player game where Player I and Player II each have two strategies. The loss matrix A , which represents the cost to Player II when specific combinations of strategies are chosen, is given by:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (5)$$

Player I's strategy choices are represented as a vector x in the space X , and Player II's strategy choices as a vector y in the space Y :

$$\begin{aligned} X &= \{x = (x_1, x_2) \mid x_i \geq 0, x_1 + x_2 = 1\}, \\ Y &= \{y = (y_1, y_2) \mid y_j \geq 0, y_1 + y_2 = 1\}. \end{aligned} \quad (6)$$

Let us denote the pure strategies by ϵ for Player I and δ for Player II:

$$\begin{aligned} \epsilon &= (e_1, e_2) \text{ where } e_1 = (1, 0), e_2 = (0, 1), \\ \delta &= (\delta_1, \delta_2) \text{ where } \delta_1 = (1, 0), \delta_2 = (0, 1). \end{aligned} \quad (7)$$

The expected cost for Player II, when players use strategies x and y respectively, can be calculated by:

$$\text{Cost}_{II}(x, y) = x^T A y \quad (8)$$

To elaborate, if Player I uses a mixed strategy $x = (x_1, x_2)$ and Player II uses a mixed strategy $y = (y_1, y_2)$, the cost for Player II is computed as:

$$\text{Cost}_{II}(x, y) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1 y_2 + x_2 y_1 \quad (9)$$

This function shows the interdependence of strategies. For instance: - If Player I and Player II choose identical strategies ($x_1 = y_1$ and $x_2 = y_2$), the cost is minimized for Player II. - Conversely, if Player I chooses a strategy opposite to that of Player II ($x_1 = y_2$ and $x_2 = y_1$), the cost is maximized for Player II, making it a classical zero-sum setting typical of adversarial games like Matching Pennies.

2.2 Examples of Trivial and Non-Trivial Game Matrices

Given the loss matrix A for a game between two players with the matrices defined as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} = [A^1 \quad A^2 \quad \cdots \quad A^n]. \quad (10)$$

2.2.1 Example of a Trivial Matrix A

A trivial matrix A might look like this:

$$A_{\text{trivial}} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} \quad (11)$$

In this matrix, all columns are identical up to a scalar multiple, which means no matter which column (strategy) player II chooses, the outcome will be predictably proportional to their choice. Here, each column is dominated by any other column (as they are identical). Thus, it offers no strategic diversity, making the game trivial as no genuine strategic decision affects the outcome.

2.2.2 Example of a Non-Trivial Matrix A

A non-trivial matrix A , which provides a more complex and strategic interaction, might look like this:

$$A_{\text{non-trivial}} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad (12)$$

This matrix has distinct elements with no rows or columns being scalar multiples of each other. It provides a range of outcomes based on the strategic choices of the players. Each choice by player II (selecting a column) results in different losses for player I depending on the row (strategy) they choose. This matrix creates a scenario where the strategy truly influences the game's outcome, providing a dynamic and non-trivial game environment.

and to avoid notation distinguishing a subset of non-weakly-dominated columns, that A has no dominated or duplicated column, for each j ,

$$e^j A \geq e^y A \text{ for all } e \text{ only if } y_j = 1. \quad (13)$$

The expectation of the loss when II uses a randomized strategy y will be called the risk,

$$R(e, y) = \sum_y e A \delta = e A y = e \cdot A y.$$

For given y , the risk function is representable as the vector, $s = A y$, and the mapping from Y to $S = A y$ furnishes a convenient canonical form for G . From the point of view of risk, G is identical with the game in which II's pure strategies are m -vectors in the set of columns of A ,

$$\sigma = (\sigma_1, \dots, \sigma_m) \text{ in } [A^1, A^2, \dots, A^n],$$

$$s = (s_1, \dots, s_m) \text{ in } S = [A y \mid y \text{ in } Y],$$

and the risk of s in S is given by the scalar product

$$R(e, s) = \sum_y e \sigma = e \cdot y \sigma = e \cdot s.$$

For each x in X the minimum of the expectation of risk,

$$\min_{s \in S} x s,$$

is attained for s in the convex hull of the set of minimizing σ . This minimum will be called the Bayes risk against x and any minimizing s will be called a Bayes strategy against x . Considered as functions, β and s , on X , β will be termed the Bayes envelope, s a Bayes response. It is convenient to extend their definition to the whole of m -space by

$$\beta(w) = \min_s w s = \min_w w \sigma,$$

$$s^* \text{ any function to } S \text{ such that } w s(w) = \beta(w) \text{ for each } w.$$

Some continuity properties of the Bayes envelope and the risks of Bayes strategies are immediate consequences of their definitions. These properties are little utilized in static game theory and their importance here stems from their direct use and analogical value in connection with the sequence games to be introduced in Section 3.

The Bayes envelope is known to be concave and continuous in more general games ([2] Theorem 2.27). Here β , as the minimum of a finite class of linear functions, is concave and piecewise linear, and

$$(w - w') s(w) \leq \beta(w) - \beta(w') \leq (w - w') s(w').$$

Although any Bayes response will be discontinuous at every point of possible ambiguity, each Bayes response possesses a local weak continuity at each w , $ws(w') - \beta(w) = w(s(w') - s(w)) \leq (w - w')s(w') - s(w)$.

Introducing the several norms for m -vectors,

$$\|v\| = \sum_{i=1}^m |v_i|, \quad |v| = \max_{i=1}^m |v_i|,$$

and the uniform bound on the variation of any Bayes response,

$$|\beta| = \sup_{w, w'} |s(w') - s(w)| = \max_{\epsilon} \max_{\sigma} |\epsilon \cdot \beta - \beta(\epsilon)| = \max_{\max} \max_{j=1}^m \left| A^j - \min_r A^r \right|,$$

the inequality may conveniently be weakened to

$$ws(w') - \beta(w) \leq \|w - w'\| |s(w') - s(w)| \leq \|w - w'\| |\beta|.$$

Letting $W_j = \{w \mid wA^j = \beta(w)\}$, β is linear on each W_j and each W_j is convex and closed. Each W_j not containing a given w has a positive distance from it and, letting $d(w)$ be the minimum of these distances (or $+\infty$ if not otherwise defined), it follows that each σ which is Bayes with respect to a w in the open neighborhood

$$\sum_{i=1}^m (w'_i - w_i)^2 < d^2(w)$$

(and hence also each $s(w')$) is necessarily Bayes with respect to w ,

$$ws(w') - \beta(w) > 0 \text{ only if } \sum_{i=1}^m (w'_i - w_i)^2 \geq d^2(w).$$

In zero-sum games against an intelligent opponent, attention has been concentrated on the ordering of Π 's strategies induced by their maximum risk. Letting $R(s)$ denote

$$\max_{\epsilon} s_{\epsilon},$$

the value of G is given by

$$R = \min_{s \in S} R(s),$$

and, by the min-max theorem,

$$R = \max_{\min_x} \sigma = \max_x \beta(x).$$

3 The Sequence Game, G^N

4 MINIMAX Modified Regret Orderings in G^N

5 The \underline{H}, μ Class of Recursive Responses

6 Modified Regret of Recursive Strategies $\underline{s}^*(\underline{E})$

7 Choise of \underline{H}, μ in $\underline{s}^*(\underline{E})$