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 sug-
gested 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 thru
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,

$$
\begin{aligned}
X &= \{x = (x_1, x_2, \ldots, x_m) \mid x_i \geq 0, \sum_{i=1}^{m} x_i = 1\}, \\
Y &= \{y = (y_1, y_2, \ldots, y_n) \mid y_j \geq 0, \sum_{j=1}^{n} y_j = 1\},
\end{aligned}
\tag{2.1}
$$

and their pure strategies will be represented by

$$
\begin{aligned}
\epsilon &= (\epsilon_1, \epsilon_2, \ldots, \epsilon_m) \text{ a basis vector in } X, \\
\delta &= (\delta_1, \delta_2, \ldots, \delta_n) \text{ a basis vector in } Y.
\end{aligned}
\tag{2.2}
$$

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,

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} = \begin{bmatrix} A^1 & A^2 & \cdots & A^n \end{bmatrix}.
\tag{2.3}
$$

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

$$
\max_{\epsilon}\{\epsilon^T A^j - \min_{r} \epsilon^T A^r\} > 0,
\tag{2.4}
$$

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

$$
\epsilon^T A^j \geq \epsilon^T A y \quad \forall \epsilon \quad \text{ only if } y_j = 1
\tag{2.5}
$$

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

$$R(\epsilon, y) = \mathbb{E}_y(\epsilon^T A \delta) = \epsilon^T A \mathbb{E}_y(\delta) = \epsilon^T A y.$$

For given y, the risk function is representable as the vector, $s = Ay$.
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,

$$\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_m) \text{ in } \{A^1, A^2, \ldots, A^n\}. \tag{2.6}$$

,

II's randomized strategies are m-vectors in the convex hull of the columns of A,

$$s = (s_1, s_2, \ldots, s_m) \text{ in } S = \{Ay : y \in Y\} \tag{2.7}$$

,

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

$$R(\epsilon, s) = \mathbb{E}_s(\epsilon^T \sigma) = \epsilon^T \mathbb{E}_s(\sigma) = \epsilon^T s \tag{2.8}$$

.

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:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

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:

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

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

$$\epsilon = (e_1, e_2) \text{ where } e_1 = (1, 0), e_2 = (0, 1),$$
$$\delta = (\delta_1, \delta_2) \text{ where } \delta_1 = (1, 0), \delta_2 = (0, 1).$$

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

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

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:

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

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:

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

2.2.1 Example of a Trivial Matrix A

A trivial matrix A might look like this:

$$A_{\text{trivial}} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

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:

$$A_{\text{non-trivial}} = \begin{bmatrix} 1 & 1 & 3 \\ 9 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix}$$

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.

2.3 Bayes Risk

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

$$\min_s x^T s$$

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, ϕ and s, on X, ϕ will be termed the **Bayes envelope**, and s the **Bayes response**. It is convenient to extend their definitions to the whole of m-space by

$$s : \mathbb{R}^m \to \mathbb{R}^n, \quad s(w) = \operatorname*{argmin}_{s \in S = \{Ay \mid y \in Y\}} w^T s$$

$$\phi(w) = w^T s(w) = \min_{s \in \text{conv}\{A^1, A^2, \dots, A^n\}} w^T s. \tag{2.9}$$

> **The Bayes envelope, as the minimum value of a finite class of linear functions, is concave and piecewise linear.**

Finite Class of Linear Functions:

The term $x^T s$ where s is a vector in the convex hull of the columns $A^1, A^2, \ldots, A^n$ of matrix A, represents a linear function in terms of s. Each column vector A^j represents a linear function $x^T A^j$. If s is an element of the convex hull of these columns, it can be expressed as a linear combination:

$$s = \sum_{j=1}^{n} \lambda_j A^j,$$

where $\lambda_j \geq 0$ and $\sum_{j=1}^{n} \lambda_j = 1$. Thus, $x^T s$ can be rewritten as:

$$x^T s = x^T \left(\sum_{j=1}^{n} \lambda_j A^j \right) = \sum_{j=1}^{n} \lambda_j (x^T A^j),$$

where $x^T A^j$ are linear functions, and s represents a linear combination of these functions.

Concavity and Piecewise Linearity:

- **Concavity:** The function $\phi(x)$ is concave because it is defined as the minimum of a collection of linear functions $\{x^T A^j\}$ (concave functions). The minimum of linear functions is a concave function, as the intersection of convex sets (hypograph of each linear component) is also convex.

- **Piecewise Linearity:** Each function $x^T A^j$ is linear in x. The function $\phi(x)$, being the minimum over these linear functions, is piecewise linear. It directly adopts the linearity of $x^T A^j$ in regions where a particular $x^T A^j$ provides the minimum value. Where the minimum shifts from one linear function to another due to changes in x, the graph of $\phi(x)$ will exhibit a 'kink' or a change in slope, reflecting its piecewise linear nature.

It holds that

$$(w - w')^T s(w) = \phi(w) - w'^T s(w) \leq \phi(w) - \phi(w') \leq (w - w')^T s(w') \tag{2.10}$$

Although any bayes response will be discontinuous at every point of possible ambiguity, the Bayes response possesses a local weak continuity at each w,

$$w^T s(w') - \phi(w) = w^T (s(w') - s(w)) \leq (w - w')^T (s(w') - s(w)) \tag{2.11}$$

proof:

$$\begin{aligned}
\nabla_w \phi(w) &= \nabla_w (w^T s(w)) = s(w) \\
\phi(w') &\leq \phi(w) + s(w)^T (w' - w) \quad \text{(by concavity)} \\
-\phi(w) &\leq -\phi(w') + s(w)^T (w' - w) \\
w^T s(w') - \phi(w) &\leq w^T s(w') - \phi(w') + s(w)^T (w' - w) \\
&= w^T s(w') - w'^T s(w') + s(w)^T (w' - w) \\
&= w^T s(w') - w'^T s(w') + w^T (-s(w)) - w'^T (-s(w)) \\
&= (w - w')^T (s(w') - s(w))
\end{aligned}$$

Introducing the several norms for m-vectors,

$$\|v\|_1 = \sum_{1}^{m} |v_i| \quad \|v\|_\infty = \max_i |v_i|, \tag{2.12}$$

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

5

$$|B| = \sup_{w,w'} |s(w') - s(w)| = \max_{\sigma} \max_{\epsilon} [\epsilon\sigma - \phi(\epsilon)] = \max_{j} \max_{i} [A_i^j - \min_{r} A_i^r] \qquad (2.13)$$

the inequality (2.11) may conveniently be weakened to

$$0 \le w^T s(w') - \phi(w) \le \|w - w'\| |s(w') - s(w)| \le \|w - w'\| |B|. \qquad (2.14)$$

Letting

$$W_j = \{w \mid wA^j = \phi(w)\}$$

ϕ 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.
Let $d(w)$ be the minimum of these distances (or $+\infty$ if not otherwise defined):

$$d(w) = \begin{cases} \min_j \min_{w' \in W_j} \|w - w'\| & \text{if } \exists j : w \notin W_j \\ \infty & \text{o.w} \end{cases}$$

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,

$$w^T s(w') - \phi(w) > 0 \text{ only if } \sum_{i=1}^{m} (w_i' - w_i)^2 \ge d^2(w). \qquad (2.15)$$

The bound (2.11) is thus improved to zero for all $\|w - w'\|$ sufficiently small, quite non-uniformly in w.

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

$$R(s) := \max_{\epsilon} \epsilon^T s$$

the value of G is given by

$$R = \min_s R(s) = \min_s \max_{\epsilon} \epsilon^T s.$$

and by the min-max theorem,

$$R = \max_{x \in X} \min_{\sigma} x^T \sigma = \max_{x \in X} \phi(x) \qquad (2.16)$$

In games against Nature, the Bayes envelope is considered a worthy defensive goal for II since it is usually felt that I's move is in no way influenced by II's choice of strategy. As a consequence, a strategy s is evaluated for each x in X in terms of the "regret", the additional expected risk above $\phi(x)$ which it incurs,

$$D(x, s) = x^T s - \phi(x). \qquad (2.17)$$

$D(x, s)$ is clearly non-negative and its maximum with respect to x is frequently used to establish a complete ordering on S. It follows from the concavity of ϕ that $\mathbb{E}_x[\epsilon^T s - \phi(\epsilon)] \geq x^T s - \phi(x)$ and hence that equality holds in

$$\max_\epsilon D(\epsilon, s) \leq \max_x D(x, s).$$

Because of the assumption (2.4),

$$\begin{aligned} D(s) &= \max_x D(x, s) = \max_\epsilon D(\epsilon, s) = \max_\epsilon[\epsilon^T s - \phi(\epsilon)] > 0, \\ D &= \min_s D(s) > 0 \end{aligned} \tag{2.18}$$

and, since

$$\max_\epsilon[\epsilon^T s - \phi(\epsilon)] = \max_x \mathbb{E}_x[x^T s - \phi(\epsilon)] = \max_x[x^T s - \mathbb{E}_x\phi(\epsilon)],$$

the min-max theorem yields the representation,

$$D = \min_s \max_x[x^T s - \mathbb{E}_x\phi(\epsilon)] = \max_x \min_s[x^T s - \mathbb{E}_x\phi(\epsilon)] = \max_x[\phi(x) - \mathbb{E}_x\phi(\epsilon)]. \tag{2.19}$$

The restriction of the regret function, $D(\epsilon, s)$, is associated with the degenerate (pointwise) partition of I's pure strategies. Developments in connection with the sequence games of Section 3 suggest that, in certain structured games, it may be of interest to consider modifications of regret associated with more general partitions of I's pure strategies. Letting π be a function of ϵ whose values partition I's pure strategies, T a subset of S, define the T-envelope of maximum risks on the partition π,

$$\phi^T(\pi) = \inf_{s \in T} \max_{\epsilon \in \pi(\epsilon)} \epsilon^T s, \tag{2.20}$$

and define the regret of a strategy s, relative to the use of T with the knowledge of $\pi(\epsilon)$, to be $\epsilon^T s - \phi^T(\pi(\epsilon))$.

3 The Sequence Game, G^N

G^N will denote a game consisting of N successive plays of G where the choices of pure strategies in the component games are completely unrestricted, II's choice in each component is allowed to depend on I's choices in previous components, and the loss is taken to be the sum (or average) of the losses of the component games.

The component G will be considered purely from the point of view of the normal form and the terminological accumulation will be reduced by referring to pure strategies in the component G as moves. (The results of the paper will find their most immediate application to games where I's pure strategies are moves in the ordinary sense.)

Denoting non-randomized move sequences in G^N by the direct product of the m-vectors representing moves in the component games and introducing the abbreviations

$$
\begin{aligned}
\underline{\epsilon}^r &= \times_{k=1}^r \epsilon^k, \qquad \sigma^r = \times_{k=1}^r \sigma^k \quad r = 1, 2, \ldots, N \\
\underline{\epsilon} &= \underline{\epsilon}^N, \qquad\qquad\qquad \underline{\sigma} = \underline{\sigma}^N,
\end{aligned}
\tag{3.1}
$$

the total loss of any play in G^N will be denoted by the function

$$
R^N(\underline{\epsilon}, \underline{\sigma}) = \sum_{k=1}^{N} (\epsilon^k)^T \sigma^k.
\tag{3.2}
$$

In the sequence game, the unsymmetric treatment of the players will be continued. It will not be necessary to introduce strategies for I since the class of results to be obtained will hold uniformly in the move sequences of I. The present and following sections are exceptional only in that it is sometimes technically convenient to consider certain linear functions of II's sequence risk which would be readily interpretable in terms of strategies for I.

The determination of G^N will be completed (and II's class of strategies implicitly circumscribed) by specifying the extent of II's knowledge of N. Two alternative situations are envisaged:

- In the weak sequence game, N is known to II at each component;

- In the strong sequence game, N is completely unknown to II.

The principal results of later sections will involve strong strategies and hence will apply to either form of G^N. In this connection it should be noted that the selection by II of a move sequence, $\times_1^N \sigma^k$, in the strong sequence game can only be accomplished by the selection of a move sequence, $\times_1^\infty \sigma^k$, in G^∞. Such selection is implicit in all results involving strong strategies.

Non-randomized sequence strategies for II will be denoted by

$$
\underline{\sigma}(\underline{\epsilon}) = \times_1^N \sigma^k(\underline{\epsilon}^{k-1})
$$

and their risk given by the natural extension of (2). Because of the linearity of the loss in the component moves, the sequence risks attainable by arbitrary randomization are attained by the class of randomized sequence strategies induced by product p-measures,

$$
\underline{y}(\epsilon) = \times_1^N y^k(\epsilon^{k-1}),
$$

on the possible $\underline{\sigma}$. Letting

$$
s^k(\underline{\epsilon}^{k-1}) = \mathbb{E}_{\underline{y}(\underline{\epsilon})} \sigma^k = \mathbb{E}_{y^k(\underline{\epsilon}^{k-1})} \sigma^k
\tag{3.3}
$$

such strategies will be denoted by

$$
\underline{s}(\underline{\epsilon}) = \times_1^N s^k(\underline{\epsilon}^{k-1})
$$

and their risk by

8

$$R^N(\underline{\epsilon}, \underline{s}) = \mathbb{E}_{\underline{y}(\underline{\epsilon})} \sum_{k=1}^{N} (\epsilon^k)^T \sigma^k = \sum_{k=1}^{N} (\epsilon^k)^T s^k(\underline{\epsilon}^{k-1}). \tag{3.4}$$

The orderings induced by maximum risk and maximum regret in G^N will be considered briefly. Letting

$$\underline{x} = \times_1^N x^j$$

be a product p-measure on $\underline{\epsilon}$,

$$\mathbb{E}_{\underline{x}} \epsilon^k s^k(\underline{\epsilon}^{k-1}) = x^k \mathbb{E}_{\underline{x}} s^k(\underline{\epsilon}^{k-1}) \geq \varphi(x^k),$$

and from this it follows that

$$\max_{\underline{\epsilon}} R^N(\underline{\epsilon}, \underline{s}) \geq \max_{\underline{x}} \sum_{k=1}^{N} \varphi(x^k) = N \max_{\underline{x}} \varphi(x) = NR.$$

9

4 MINIMAX Modified Regret Orderings in G^N

5   The $\underline{\mathbf{H}}, \mu$ Class of Recursive Responses

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

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