This is to certify that the
dissertation entitled

STRATEGIES OF REPEATED GAMES

presented by

MINGFEI LI

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Applied Mathematics

Dennis Gilliland
Major Professor's Signature

July 18, 2008
Date

MSU is an affirmative-action, equal-opportunity employer

STRATEGIES IN REPEATED GAMES

By

Mingfei Li

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Applied Mathematics

2008

ABSTRACT

STRATEGIES IN REPEATED GAMES

By

Mingfei Li

In games that are repeated, the players have the opportunity to use information on opponents' past moves in selecting a move for the current stage. Strategies for Player II are considered in this thesis. In particular, the Play Against the Past strategy (PAP), the Play Against the Past plus Present strategy (PAP+), the Play Against the Random Past strategy (PARP), and Hannan-type strategies are investigated, especially in the repeated play of the two-person game called matching pennies. The effectiveness of a strategy is measured in terms of difference in average loss and an envelope loss; this difference is called regret. In some cases, exact expressions for regret are derived; more often, asymptotic properties are derived.

The PAP strategy for Player II is not effective against all Player I move sequences. Hannan (1957) used a Bayes response to random perturbations of Player I's empirical distribution of past moves as a strategy and established good asymptotic regret properties uniform in Player I move sequences for the repeated play of a variety of games. Gilliland (2004) and Gilliland and Jung (2006) introduced the PARP strategy where the randomization comes through bootstrap sampling of Player I's past moves and established results for the repeated play of matching pennies.

The PAP, PAP+, PARP and Hannan-type strategies are defined in Chapter 2. The adaptation of PARP to achieve regret results relative to k-extended envelopes is demonstrated in Chapter 3 for matching pennies. Chapter 4 documents cases where strategies published following Hannan's seminal (1957) paper are unrecognized, special cases of his work. PARP is discussed in the context of the expert selection problem in Chapter 5, and regret asymptotics are derived for certain classes of Player I move sequences.

DEDICATION

This dissertation is dedicated to my wonderful parents, Jiequn and Shumay who raised me, supported me, taught me and loved me. You have been with me every step of the way, through good times and bad.

Also, this dissertation is dedicated to my grandparents, aunts and uncles, who have been great sources of motivation and inspiration.

Thank you for all the unconditional love, guidance, and support that you have always given me, helping me to succeed and instilling in me the confidence that I am capable of doing anything I put my mind to. Thank you for everything. I love you!

ACKNOWLEDGMENT

From the formative stages of this dissertation, to the final draft, I owe an immense debt of gratitude to my PhD supervisor, Dr. Dennis Gilliland. His sound advice and careful guidance were invaluable. I have learned so much, and without you, this would not have been possible.

For their efforts, assistance and support, a special thanks as well to the Dr. Zhou, Dr. Yan, Dr. Weil, Dr. Page, Dr. Dass, Dr. Vorro, Dr. Bush, Barbara Miller and all my good friends and students colleagues in Michigan State University.

Finally, I would be remiss without mentioning Dr. Mark Hurwitz, Mr. and Mrs. Mawhorter, whose extreme generosity will be remembered always.

To each of the above, I extend my deepest appreciation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

[1]

[1]Images in this dissertation are presented in color

Chapter 1

Introduction

1.1 Game theory

Game theory is the theory of rational behavior for interactive decision problems. In a game, participants strive to maximize their expected gain by choosing particular courses of action, and each participant's final payoff depends on the profiles of the courses of action chosen by all participants. The interactive situation, specified by the set of participants, the information flow, the possible courses of action of each participant, and the set of all possible payoffs, is called a *game*. Participants, i.e., those who are 'playing' the game, are called the *players*.

In a game, if the goal of each player is to achieve the largest possible individual gain (profit or payoff), the game is called a *noncooperative game*. Games in which the actions of the players are directed to maximized the gains of coalitions without subsequent subdivision of the gain among the players within the coalition are called *cooperative games*. In this thesis, we focus on some noncooperative games.

The basic objects of interest in noncooperative games are players' *strategies*. A player's strategy is a complete plan of action, i.e., the moves to be taken when the game is played; it must be completely specified before the actual play of the game starts, and it prescribes the course of play for each move that a player might be called upon to take, for each possible piece of information that the player may have at each time where he or she might be called upon to act.

In simple form, a two-person game is a triple (A, B, L) where A is the set of moves for Player I, B is the set of moves for Player II, and L is a nonnegative function defined on $A \times B$ with $L(x, y)$ denoting loss to Player II when Player I plays x and Player II plays y. With a $\sigma-$field of subsets defined for A, suitable integrability conditions for L, and A^* denoting the class of probability distributions on the $\sigma-$field, the domain of L is extended to $A^* \times B$ by $L(\pi, y) = \int L(x, y) d\pi(x)$. If the class of probability distributions includes all degenerate probability distributions for the points in A, then (A^*, B, L) formally extends the game (A, B, L) to include randomized strategies for Player I. Under suitable assumptions, the game extends to (A^*, B^*, L) where both players have randomized strategies $\pi \in A^*$, $\gamma \in B^*$. For the extension, the loss function is an expectation (expected loss), but it will still be called the loss function.

Our focus is on moves or strategies for Player II and generally Player I's utility or inutility are not defined. For a zero-sum game, it is understood that Player I's gain is Player II's loss. If Player II uses the distribution γ to generate his/her move, we refer to this as randomization. Here the move y is determined as the realization of a random variable with a probability distribution γ specified by the player.

A minimax strategy for Player II is any move γ_{mM} such that $\max_\pi L(\pi, \gamma_{mM}) = \min_\gamma \max_\pi L(\pi, \gamma)$, the upper value of the game. A maxmin strategy for Player I is any move π_{Mm} such that $\min_\gamma L(\pi_{Mm}, \gamma) = \max_\pi \min_\gamma L(\pi, \gamma)$, the lower value of the game. If the upper value is equal to the lower value, the common value is called the value of the game.

A Bayes rule for Player II versus the distribution π is any γ such that $L(\pi, \gamma) = \min_\gamma L(\pi, \gamma)$. The minimum is denoted as $R(\pi)$ and called minimum Bayes risk. $R(\cdot)$ is called the Bayes envelope for the game. A minimizer exists in the set B of pure moves. Any function σ on A^* with range in B^* and such that $L(\pi, \sigma(\pi)) = R(\pi)$ for all $\pi \in A^*$ is called a Bayes response. Hannan (1957) took σ to be *B-valued.*

Consider a zero-sum game where Player I selects a move from the set $A = \{a_1, a_2, \ldots, a_m\}$ and Player II selects a move from set $B = \{b_1, b_2, \ldots, b_n\}$ with loss $L(a_i, b_j)$ to Player II if I chooses a_i and II chooses b_j. This is called a *finite* $m \times n$ *game.* Here A^* is the probability simplex in R^m, B^* is the probability simplex in R^n and all of the expectations are inner products.

Much of our study concerns the simple 2×2 game where each player selects from $\{0, 1\}$ and the loss function is $L(i, j) = |i - j| = i \cdot (1 - j) + (1 - i) \cdot j$, $i, j = 0, 1$. This is the game of *matching pennies* (*or matching binary bits*) with Player II's objective to match Player I. In matching binary bits, $1 - 2L$ is the gain for Player II while $2L - 1$ is Player I's gain.

Suppose that Player II generates his/her move in the matching pennies with a Bernoulli distribution $B(1, p)$, i.e. $prob(j = 1) = p$ and $prob(j = 0) = 1 - p$. Then Player II's expected loss is seen to be $L(i, p) = i \cdot (1 - p) + (1 - i) \cdot p = |i - p|$, $i \in \{0, 1\}$, $p \in [0, 1]$. Thus the loss function extends to expected loss on the domain $\{0, 1\} \times [0, 1]$ and we will call it simply loss where there is no chance of confusion.

Applying the extended loss to a weather forecast of rain with probability p, the forecaster (Player II) suffers "loss" $1 - p$ if it rains ($i = 1$) and "loss" p if it does not rain ($i = 0$). The choice $p = \frac{1}{2}$ in the minimax choice for II, it minimizes the maximum possible expected "loss", i. e., $p = \frac{1}{2}$ minimizes $(1 - p) \vee p$. Notice that the weather forecaster is not required to actually generate the Bernoulli random variable to serve as his/her move. Rather he/she simply specifies a probability p. If "Nature" flips an unbiased coin to determine whether it rains or not, then the equilibrium "loss" $\frac{1}{2}$ is achieved, the value of the game.

1.2 Repeated play

If a fixed group of players plays a given game repeatedly, we say this is a *repeated game* or is *repeated play*. In another words, a repeated game is the same simultaneous game played repeatedly. The payoffs add across repeated play. In repeated play, rules will specify what information generated in the repeated play is made available to what players and when. We will assume that all players will be fully informed of the rules that govern the game that is being repeated together with the history of moves of all players at all stages of the repeated play. Thus, the player may use strategies. i.e., sequences of functions that map the history of past moves into a move for the current stage.

The repeated play is of two types:

(1) *Finite Horizon* where there is to be a sequence of N plays where N is finite, specified and known to the players in advance. See Hannan's weak sequence game (1957, Sec 3). In finite horizon play, the players' strategies can depend on N. Conceptually, the finite horizon repeated play game is another example of a simultaneous move game where the strategies are finite sequences of recursive functions. Player II strategies are evaluated in terms of the average loss over the N games. We will consider finite horizon repeated play only in Chapter 4, Section 3.

(2) *Infinite Horizon* where there is an infinite sequence of plays and the players know this. See Hannan's strong sequence game (1957, Sec 3). A player II strategy is evaluated in terms of the sequence of average loss over initial segments.

Our study concerns the review of and the development of "good" strategies for Player II in the repeated play of a two-person game. Generally, results uniform in sequences of Player I moves are sought and obtained. With such results, the findings extend to results uniform in Player I strategies and show that the motivation for Player I is irrelevant (Hannan, 1957). The two-person construct is not as restrictive as it seems since the term Player I may be taken to name a coalition or collection of players.

Now consider the repeated play of matching pennies. We let $\underline{a}$ and $\underline{b}$ denote infinite sequences of moves for the respective players and let $\underline{a}_t$ and $\underline{b}_t$ denote initial sequences, $t = 1, 2, \ldots$ A deterministic strategy (*pure strategy*) for Player II has as components recursive functions $\underline{b}_t(\underline{a}_{t-1})$ taking values in $\{0, 1\}$, $t = 2, 3, \ldots$ with $b_1 \in \{0, 1\}$. The associated average (Cesaro) loss to Player II across N plays at the Player I sequence $\underline{a}$ is:

$$CL_N(\underline{a}, \underline{b}) = \sum_{t=1}^{N} L(a_t, b_t(\underline{a}_{t-1}))/N = \sum_{t=1}^{N} |a_t - b_t(\underline{a}_{t-1})|/N$$

As is rather obvious and perhaps first recorded by Cover (1967),

$$\max\{\sum_{t=1}^{N} |a_t - b_t(\underline{a}_{t-1})| \mid \underline{a}_N \in \{0, 1\}^N\} = N$$

for every b_1 and sequence of $\underline{a}_{t-1}$ - measurable functions $\underline{b}_t(\cdot)$, $t = 2, 3, \ldots$. Thus, no deterministic strategy for Player II can produce the uniform convergence of average loss to zero. $\sum_1^N |a_t - b_t|$ is the Hamming distance between the binary sequences $a_1, a_2, \ldots, a_N$ and $b_1, b_2, \ldots b_N$.

A *stochastic strategy (mixed strategy)* for Player II has as components recursive functions $\underline{p}_t(\underline{a}_{t-1})$ taking values in $[0, 1]$, $t = 2, 3, \ldots$ with $p_1 \in [0, 1]$. Identifying 1 with $p = 1$ and 0 with $p = 0$, the stochastic strategies include the deterministic strategies as a subclass. The associated average (Cesaro) loss to Player II across N plays is

$$CL_N(\underline{a}, \underline{p}) = \sum_{t=1}^{N} L(a_t, p_t(\underline{a}_{t-1}))/N = \sum_{t=1}^{N} |a_t - p_t(\underline{a}_{t-1})|/N$$

Note that

$$\max\{\sum_{t=1}^{N} |a_t - p_t(\underline{a}_{t-1})| \mid \underline{a}_N \in \{0, 1\}^N\} \geq N/2.$$

for every p_1 and sequence of $\underline{a}_{t-1}$ - measurable functions $\underline{p}_t(\cdot)$, $t = 2, 3, \ldots$. Thus, no stochastic strategy for Player II can produce the uniform convergence of average loss to zero.

Hannan (1957, Sec 3, (11)) introduced what he called *modified regret* for the evaluation of Player II strategies (using the scale of total loss). We use the term *regret* to denote the difference between the average loss for a strategy and the minimum average loss (*envelope loss*) across a specified set of (often simple) strategies.

We will illustrate regret in the repeated play of matching pennies where Player I selects $a \in \{0, 1\}$ and Player II selects a probability $p \in [0, 1]$ with loss $L(a, p) = |a-p|$ to Player II.

Example 1.1 *The Simple Envelope and Regret for Repeated Play of Matching Pennies.*

Consider the two strategies $\underline{p}^{(0)}$ and $\underline{p}^{(1)}$ where

$$p_1^{(0)} = 0 \quad and \quad p_t^{(0)}(\underline{a}_{t-1}) = 0$$

for $t = 2, 3, \ldots$ (i.e., always play a 0) and

$$p_1^{(1)} = 1 \quad and \quad p_t^{(1)}(\underline{a}_{t-1}) = 1$$

for $t = 2, 3, \ldots$ (i.e., always play a 1). Let $S = \{\underline{p}^{(0)}, \underline{p}^{(1)}\}$. The *simple envelope* is defined as

$$R^{(1)}(\underline{a}_N) = min\{CL_N(\underline{a}, \underline{p}) | \underline{p} \in S\} = \min\{\sum_{t=1}^{N} a_t/N, \quad 1 - \sum_{t=1}^{N} a_t/N\}.$$

Dropping the superscript and letting $g_N = \sum_1^N a_t/N$, $N = 1, 2, \ldots$, the simple envelope evaluated at $\underline{a}_N$ can be written by:

$$R(g_N) = g_N \wedge (1 - g_N).$$

This is the Bayes envelope of the component game evaluated at the empirical probability distribution g_N of $\{a_1, a_2, \ldots, a_N\}$. The regret sequence associated with a strategy $\underline{p}$ relative to the simple envelope is

$$D_N(\underline{a}, \underline{p}) = CL_N(\underline{a}, \underline{p}) - R(g_N), \quad N = 1, 2, \ldots$$

Remarkably, Hannan (1957) and Blackwell (1956ab) independently developed strategies of a very different structure for which $D_N(\underline{a}, \underline{p})$ is $O(N^{-1/2})$ uniformly in $\underline{a}$. In each case the development was for repeated play of general finite games and more. Hannan (1957) worked to get tight bounds and, therefore, good constants in his $O(N^{-1/2})$ demonstrations. Of course, Player II's concern is

$$\limsup_N D_N(\underline{a}, \underline{p}) \leq 0$$

and this limit condition may have first been referred to as *Hannan consistency* (at $\underline{a}$) in Hart and Mas-Colell (2001, p.27). We will also refer to the sufficient condition

$$\lim_N D_N(\underline{a}, \underline{p}) = 0$$

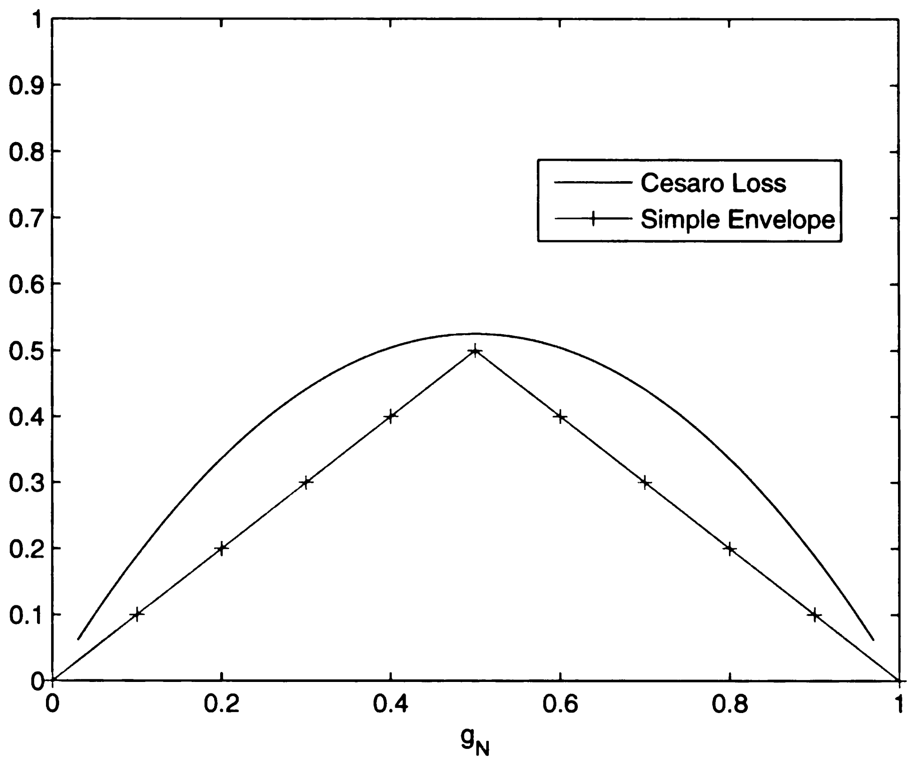
as Hannan consistency (at $\underline{a}$). This thesis has elaborations on various strategies demonstrating Hannan consistency.

Example 1.1 (continued) Figure 1.1 below is a plot of the simple envelope R and the Cesaro loss for a hypothetical strategy with Hannan consistency. Note this interpretation: Player II using the Hannan consistent strategy does almost as well on average through horizon N as if he/she were told in advance what was to be Player I's majority move in the stages $t = 1, 2, \ldots, N$ and he/she simply played that choice in attempting to match Player I. The Hannan consistent strategy $\underline{p}$ has the property

$$\limsup_N \max\{CL_N(\underline{a}, \underline{p}) | \underline{a}_N \in \{0, 1\}^N\} \leq 1/2,$$

i.e., it is *asymptotically subminimax.* In the limit, Player II loses at most one-half the time and does better if g_N stays away from $\frac{1}{2}$.

Figure 1.1: Cesaro Loss vs Bayes Envelope



1.3 Summary of Thesis

In this section we summarize the thesis and the results herein.

Chapter 2 introduces the *play against the past strategy* which we label as the *PAP strategy*. With this strategy, Player II at stage t plays component Bayes versus the empirical distribution of Player I's past moves $\{a_1, a_2, \dots a_{t-1}\}$, $t = 2, 3, \dots$. In matching pennies, this has Player II playing the majority move found in Player I's past moves. We also consider the unrealizable strategy called *play against the past plus present* (PAP+ strategy). In matching pennies, this has Player II playing the majority move found in Player I's past and present moves $\{a_1, a_2, \dots, a_{t-1}, a_t\}$. Hannan (1957) used these strategies and their properties both in motivation and in proofs for Hannan consistent strategies for repeated play. We develop exact expressions for the simple regrets of the PAP and PAP+ strategies in matching pennies. In Chapter 2, we introduce the *play against the random past strategy* (PARP strategy) first considered by Gilliland (2004) and Gilliland and Jung (2006). With this strategy, Player II plays component Bayes versus a random sample drawn with replacement from the past moves $\{a_1, a_2, \dots a_{t-1}\}$. A goal of this thesis work was to demonstrate Hannan consistency for this strategy for the repeated play of the expert selection problem, a goal only partially reached (Chapter 5). In Chapter 2 we define *Hannan-type strategies* for later reference in Chapter 4. Essentially, a Hannan-type strategy plays component Bayes versus either a controlled random perturbation of the empirical distribution of Player I's past moves or the expectation of such. We conclude Chapter 2 by illustrating the need for fresh randomization across stages when implementing a strategy for matching pennies. We base the example on a Hannan-type strategy.

In Chapter 3, we examine extended envelopes for matching pennies. These envelopes are called *k-extended envelopes* and are more stringent than the simple envelope. Whereas the simple envelope is the Bayes envelope of the component game evaluated at the empirical distribution of $\{a_1, a_2, \dots, a_{N-1}, a_N\}$, the 2-extended envelope is a Bayes envelope evaluated at the empirical distribution of pairs $\{(a_1, a_2), (a_2, a_3), \ldots, (a_{N-1}, a_N)\}$. If Player I's moves exhibit Markov structure, for example, a tendency to follow a 0 with a 1, then the 2-extended envelope can be considerably less than the simple envelope. In matching pennies, we develop exact expressions for the 2-extended regrets for the PAP and PAP+ strategies and establish Hannan-consistency for a PARP strategy.

Chapter 4 reports on a literature search to document specific theorems and results published by others after Hannan's (1957) seminal paper, results that are found in or are direct consequences of Hannan (1957) results. Because of the cryptic style and possibly the notations used in Hannan (1957), it is understandable that other researchers failed to recognize the specific results therein. The style and notations makes the documentations rather challenging in some cases. The literature includes Cover (1967), Feder, Merhav and Gutman (1992), Foster and Vohra (1993), Chung (1994), and Cesa-Bianchi and Lugosi (1999). This search was motivated in part by the Gina Kolta (2006) New York Times article *Pity the Scientist Who Discovers the Discovered* in which Hannan is mentioned.

Chapter 5 introduces the *expert selection* problem, which has gotten considerable attention in the game theory and computer science research communities. Here Player II must select from a class of experts and assume whatever loss is incurred by that expert in a specified game. This problem is often cast in terms of a forecasting problem. For example, consider a set of K weather forecasters (experts). Player II must make a weather forecast for tomorrow; rather than do his/her own analysis, Player II examines the records of accuracy for the K experts and selects the forecast of the one who has the best record of past accuracy. As described, this would be a PAP strategy. PAP strategies here and in general are not Hannan consistent on all Player I sequences $\underline{a}$. In repeated games, the set of experts could be a set of strategies. Player II uses the performance record of the strategies to select one to implement in the current stage. Chapter 5 starts by discussing *focus forecasting* (Smith, 1978) which can be described as PAP where the tests of the forecasting strategies in the pool are over recent performance, not the complete past. In practice, this is a criticized methodology since the pool of experts seems to have grown in a rather ad hoc fashion. For example, see Gardner, Anderson-Flether and Wicks (2001). Smith's company (Focus Forecasting.com) continues to serve customers. In Chapter 5, we investigate the use of the PARP strategy in expert selection. We examine the case the pool has only two experts and show the problem to be reducible to a one-dimensional problem. This problem is examined and a class of sequences $\underline{a}$ where PARP is Hannan consistent is identified. We conclude Chapter 5 with empirical tests of the PARP strategy for selecting from competing time series models for prediction.

Chapter 2

The PAP, PAP+, PARP and Hannan-Type Strategies

2.1 Play Against the Past (PAP) and Past plus Present (PAP+)

Play against the past in the repeated play of a two-person game denotes the strategy for Player II in which II at each stage $t = 2, 3, \ldots$ plays component game *Bayes* versus the empirical distribution of I's past moves. The study of this strategy in general settings is undertaken in Hannan (1957) where basic inequalities (Sec 8, (11)) show the possible importance of the study to the construction of good strategies for Player II in repeated games. Gilliland (1972) continues the discussion of play against the past strategies in sequences of statistical decision problems. Play against the past is a one-sided version of what is called *fictitious play* in the repeated play of a two-person, zero-sum game (Robinson, 1951).

Recall that a Bayes rule for Player II versus a prior distribution over the possible moves by Player I is any move that minimizes the expected loss to Player II. For example, in matching pennies, a Bayes rule versus the probability distribution

$Prob(a = 1) = \pi$, $Prob(a = 0) = 1 - \pi$, is any rule where $p = 1$ (Player II plays $b = 1$) if $\pi > \frac{1}{2}$ and $p = 0$ (Player II plays $b = 0$) if $\pi < \frac{1}{2}$. In our study, we will usually take the determination $p = \frac{1}{2}$ when $\pi = \frac{1}{2}$. Formally, the Bayes response we consider in our analyses of matching pennies is denoted by $\sigma(\cdot, \cdot)$, where

$$\sigma(1 - \pi, \pi) := [\pi > \frac{1}{2}] + \frac{1}{2}[\pi = \frac{1}{2}], \quad 0 \leq \pi \leq 1.$$

Here and throughout this thesis, square brackets denote indicator functions. Moreover, it is convenient for future use to extend the domain of the Bayes response to $\sigma(\omega_1, \omega_2) \in [0, \infty)^2 - (0, 0)$ by

$$\sigma(\omega_1, \omega_2) := \sigma(\omega_1/(\omega_1 + \omega_2), \omega_2/(\omega_1 + \omega_2)).$$

The PAP strategy for Player II in matching pennies is denoted by and defined by

$$PAP: \quad pap_1 = \frac{1}{2}, \quad pap_t(\underline{a}_{t-1}) = [g_{t-1} > \frac{1}{2}] + \frac{1}{2}[g_{t-1} = \frac{1}{2}]$$

where recall from Chapter 1 that g_{t-1} denotes the proportion of 1's in the sequence $\underline{a}_{t-1}$. With the PAP strategy for matching pennies, Player II starts with a coin toss (assumed to be a fair coin) and subsequently plays the majority choice in Player I past moves with a coin toss in the event of a tie

Hannan (1957, Sec 8, (11)) also considered the unrealizable strategy for Player II that in the context of matching pennies is

$$PAP+: \quad pap+_1 = \frac{1}{2}, \quad pap+_t(\underline{a}_t) = [g_t > \frac{1}{2}] + \frac{1}{2}[g_t = \frac{1}{2}].$$

This can be thought of as play against the past including present. Note that this strategy has Player II's move at stage t to be the Bayes response versus the empirical distribution of $\{a_1, a_2, \ldots, a_{t-1}, a_t\}$. Hannan (1957) established for the repeated play of a general game that the average loss from PAP+ in no greater than the simple envelope loss and that the average loss from PAP is no less than the simple envelope loss.

The evaluations for PAP and PAP+ are simple and illustrative in the case of matching pennies. In developing a new strategy PARP for matching binary bits (matching pennies) Gilliland and Jung (2006) proved the following proposition in regard to PAP.

Proposition 2.1.1. *In Matching Pennies, the Cesaro loss sequence for the PAP strategy is given by*

$$CL_N(\underline{a}, \underline{pap}) = g_N \wedge (1 - g_N) + 0.5\nu_N/N + 0.5 \cdot [g_N \neq 1/2]/N, \quad N = 1, 2, \ldots$$

where ν_N is the number of g_t visit to 1/2, $t = 1, 2, 3, \ldots, N$

Note that the excess average loss over the simple envelope loss $g_N \wedge (1 - g_N)$ is positive and is maximized at $\underline{a} = (0, 1, 0, 1, \ldots)$ or $(0, 1, 0, 1, \ldots)$ with the maximum being 0.25 asymptotically. Here we have $\lim_N D_N(\underline{a}, \underline{pap}) = 0.25$. That PAP is not Hannan consistent at all in Player I move sequences is a well known result.

We now turn to the the unrealizable strategy PAP+ in matching pennies.

Proposition 2.1.2. *The Cesaro loss sequence for the PAP+ strategy is given by*

$$CL_N(\underline{a}, \underline{pap+}) = g_N \wedge (1 - g_N) - 0.5\nu_N/N,$$

where ν_N is the number of g_t visit to 1/2, $t = 1, 2, 3, \ldots, N$. Furthermore,

$$\max_{\underline{a}}\{CL_N(\underline{a}, \underline{pap+}) | fixed \quad Ng_N\} = g_N \wedge (1 - g_N)$$

and

$$\min_{\underline{a}}\{CL_N(\underline{a}, \underline{pap+})|fixed \quad Ng_N\} = g_N \wedge (1-g_N) - 0.5(greatest \quad integer \quad in \quad N/2)/N.$$

Proof: Let $N > 0$ be fixed and take $a_1 = 1$ without loss of generality. Suppose that g_t returns to $1/2$ at stages $i_1, i_2, \ldots, i_k$, where $1 < i_i < i_2, \ldots, < i_k \leq N$. Consider the first epoch $1 \leq t \leq i_1$. Note that $g_t > 1/2$ on $1 \leq t < i_1$ and $g_{i_1} = 1/2$ so that Player II plays 1 on $1 \leq t < i_1$ and $1/2$ at $t = i_1$. Player I has played $i_1/2$ 0's including the 0 at stage i_1 and $i_1/2$ 1's on epoch $1 \leq t \leq i_1$. Thus, the total loss to Player II on the first epoch is $(i_1/2 - 1) + 1/2 = (i_1/2 - 1/2)$. The total loss across all epochs is $(i_1/2 - 1/2) + (i_2 - i_1)/2 - 1/2 + \cdots + (i_k - i_{k-1})/2 - 1/2 = i_k/2 - k/2$. If $i_k = N$, then $g_N = 1/2$ and the average loss is

$$CL_N(\underline{a}, \underline{pap+}) = g_N - 0.5\nu_N/N = g_N \wedge (1-g_N) - 0.5\nu_N/N$$

where $\nu_N := \{number \quad of \quad g_t \quad visits \quad to \quad 1/2|t = 1, 2, 3, \ldots, N\}$.

Now suppose that $i_k < N$. Let $a_{i_k+1} = 1$ without of loss of generality so that $g_N > 1/2$. Then Player II plays 1 on the $N - i_k$ stages $i_k < t \leq N$. On these stages, Player I plays a total of $(Ng_N - i_k/2)$ 1's and therefore, $(N - i_k) - (Ng_N - i_k/2)$ 0's, which is the total loss for Player II. Thus, the average loss for Player II across all N stages is

$$CL_N(\underline{a}, \underline{pap+}) = (i_k/2 - \nu_N/2)/N + (1 - g_N) - i_k/2N = g_N \wedge (1-g_N) - 0.5\nu_N/N.$$

For the fixed total number Ng_N of 1's in the sequence $\underline{a}_N$, we see that the Cesaro loss is maximized when $\nu_N = 0$ and minimized by alternating 1's and 0's.

Proof is done.□

It follows from Propositions 2.1.1 and 2.1.2 that

$$D_N(\underline{a}, \underline{pap}) = \frac{\nu_N}{2N} + \frac{[g_N \neq 1/2]}{2N}$$

and

$$D_N(\underline{a}, \underline{pap+}) = -\frac{\nu_N}{2N}.$$

Suppose that Player I generates moves $a_1, a_2, \ldots$ as independent, identically distributed $B(1, \pi)$. We show that PAP is Hannan consistent at $\underline{a}$. It suffices to show that $\nu_N/N + [g_N \neq 1/2]/N \to 0$. The second term is bounded by $1/N$ so we need only consider the first term ν_N/N. ν_N is the number of visits of the random walk $S_t = \sum_1^t (2a_i - 1)$ to 0 across $t = 1, 2, \ldots, N$, $N = 1, 2, \ldots$. The strong law of large numbers shows that $S_N/N \to 2\pi - 1$ a.s.. Thus, if $\pi \neq 1/2$, S_N is 0 only finitely often a.s., which implies that $\nu_N/N \to 0$ a.s.. Thus, where PAP is not Hannan consistent at all sequences $\underline{a}$, it is Hannan consistent a.s. if Player I repeatedly and independently generates his/her moves by a coin toss that has probability π of turning up Heads ($a = 1$) provided $\pi \neq 1/2$. Since $g_N \to \pi$ a.s. and $R(\underline{a}_N) = g_N \wedge (1 - g_N) \to \pi \wedge (1 - \pi) \leq 1/2$ a.s. with equality if and only if $\pi = 1/2$, Player II is sure to win more than 50% of the time in the limit if $\pi \neq 1/2$, i.e., the coin is biased. If $\pi = 1/2$ (the coin is unbiased), there is simple expression for $E(\nu_N)$. From Grinstead and Snell (2008, p. 481),

$$E(\nu_{2N}) = \alpha_N - 1$$

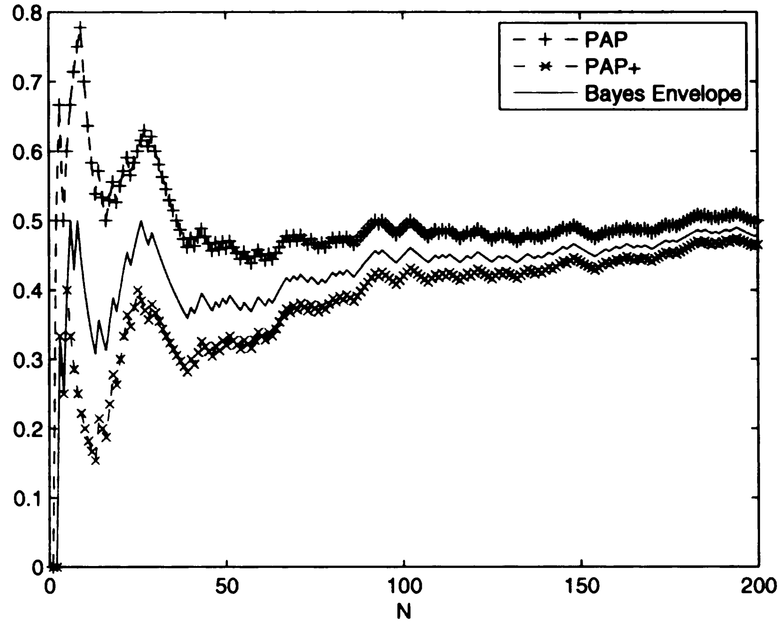
where

$$\alpha_N = \frac{(2N+1)!}{4^N N! N!}$$

will appear again in chapter 4, section 3. Since $\alpha_N \sim \sqrt{4N/\pi}$, $E(\nu_{2N}/2N) \sim 1/\sqrt{\pi N}$ and using $\nu_{2N+1} = \nu_{2N}$ it follows that $E(\nu_N/N) \to 0$ in L_1.

Figure 2.1 shows the result of a simulation where the a_t are i.i.d Bernoulli (1,1/2), $t = 1, 2, \ldots, 100$.

Figure 2.1: PAP vs PAP+ vs Envelope for i.i.d Bernoulli sequence



2.2 Play Against the Random Past (PARP)

Gilliland (2004) announced the result that play against the random past in matching pennies is Hannan consistent with uniform rate $O(N^{-1/2})$. Proof was given in Gilliland and Jung (2006). The PARP strategy for Player II in matching pennies is denoted by and defined by

$$PARP: \quad parp_1 = \frac{1}{2}, \quad parp_t(\underline{a}_{t-1}) = [g^*_{t-1} > \frac{1}{2}] + \frac{1}{2}[g^*_{t-1} = \frac{1}{2}]$$

where g^*_{t-1} is the proportion of 1's in a random sample of size $t-1$ drawn with replacement from Player I's moves $\{a_1, a_2, \ldots, a_{t-1}\}$. It is assumed that the bootstrap samples are independent across the stages t, i.e., that fresh samples are drawn at each stage $t = 2, 3, \ldots$ Study of PARP in matching pennies requires the analysis of the half-binomial probabilities

$$P_{t-1, g_{t-1}} := E([g^*_{t-1} > \frac{1}{2}] + \frac{1}{2}[g^*_{t-1} = \frac{1}{2}]).$$

Gilliland and Jung (2006) show that there exist constants A and B such that

$$|D_N(\underline{a}, \underline{parp})| \leq (A + B\sqrt{N \cdot (g_N \wedge (1 - g_N))} \quad)/N,$$

thus establishing uniform Hannan consistency for PARP with rate $O(N^{-1/2})$.

In Chapter 5 we explore the PARP approach for repeated play of an infinite component game that is motivated by the expert selection problem.

2.3 Hannan-Type Strategies (H)

Hannan-type (1957) strategies overcome the weakness in PAP by playing Bayes responses or the expectations of Bayes responses to properly scaled random perturbations of the empirical distributions G_{t-1}. Specifically, with a component game where

Player I has m moves $\{1, 2, \ldots, m\}$, the empirical probability distribution of Player I's moves through time $t-1$ is the vector $G_{t-1} := (n_1^{t-1}, n_2^{t-1}, \ldots, n_m^{t-1})/(t-1)$ where $n_i^{t-1} = num\{a_j = i | j = 1, 2, \ldots, t-1\}$, $i = 1, 2, \ldots, m$. We define a *Hannan-type strategy* as any Player II strategy that at stage t plays

$$\sigma(G_{t-1} + h_{t-1} Z_{t-1}) \quad , \quad t = 2, 3, \ldots \tag{2.1}$$

or

$$E(\sigma(G_{t-1} + h_{t-1} \cdot Z)) \quad , \quad t = 2, 3, \ldots \tag{2.2}$$

where $\{h_{t-1}\}$ is a sequence of positive real numbers, Z_{t-1} and Z are random vectors take values in $(0, \infty)^m$, and E is expectation over Z. To simplify proofs, Hannan extends the domain of the Bayes response σ from the probability simplex in R^m to all of $[0, \infty)^m$ with σ being positive homogeneous of order 0, that is, $\sigma(cw) = \sigma(w)$ for all $c > 0$, $w \in [0, \infty)^m$. Hannan (1957) for repeated play of a variety of component games, including the finite two-person game and the S-game, imposes conditions on the sequence of constants $\{h_t\}$ and the distribution of Z to achieve uniform Hannan consistency for the strategy (2.2) with rates.

In matching pennies where $m = 2$, we have labeled the pure moves as 0 and 1 and let $g_{t-1} = num\{a_j = 1 | j = 1, 2, \ldots, t-1\}/(t-1)$ denote the empirical proportions of 1's in Player I's initial move sequences. Since $1 - g_{t-1} = num\{a_j = 0 | j = 1, 2, \ldots, t-1\}/(t-1)$, the empirical probability distribution is $G_{t-1} = (1 - g_{t-1}, g_{t-1})$.

Chapter 4 includes a survey of published results that are subsumed by Hannan (1957). It appears that many of the authors were unaware of the specific results contained in Hannan (1957). Since a positive homogeneous (order 0) Bayes response function plays a key role in proofs for Hannan-type strategies, we conclude this section with examples to illustrate σ, its properties and the notations that are used. Hopefully, these examples will help make the proofs in Chapter 4 of connections of Hannan-type strategies to other work understandable.

Example 2.2 *(Matching pennics)*

Here Player I and Player II have $m = n = 2$ pure moves which we have denoted as $\{0, 1\}$. Suppose that Player I selects his/her move with the (prior) probability distribution $\text{Prob}(0) = 1 - \pi$, $\text{Prob}(1) = \pi$. Consider Player II selects his/her move with $\text{Prob}(0) = 1 - p$, $\text{Prob}(1) = p$. The expected loss to Player II is $L(\pi, p) = (1 - \pi)p + \pi(1 - p)$. Any p that minimizes $L(\pi, p)$ is *Bayes versus* π. A *Bayes response* is any function σ on the probability simplex in R^2 such that $p = \sigma(1 - \pi, \pi)$ is Bayes versus π. For each $\pi > 1/2$, $p = 1$ is uniquely Bayes versus π; for each $\pi < 1/2$, $p = 0$ is uniquely Bayes versus π; for $\pi = 1/2$, all $p \in [0, 1]$ are Bayes versus π. To specify a Bayes response one must select a minimizer when the minimizer is not unique. Here is the example given earlier in section 2.1:

$$\sigma(1 - \pi, \pi) = \begin{cases} 0, & 0 \le \pi < \frac{1}{2}, \\ \frac{1}{2}, & \pi = \frac{1}{2}, \\ 1, & \frac{1}{2} < \pi \le 1. \end{cases}$$

When $m = 2$, there is the notational convenience derived from identifying $(1-\pi, \pi)$ by π . However, this identification hides features, including the positive homogeneous property of order 0 imposed by Hannan on the Bayes response. As noted in section 2.1, once a Bayes response is defined on the probability simplex in R^2, the domain is easily extended to all of $[0, \infty)^2 - \{(0, 0)\}$ by $\sigma(w_1, w2) = \sigma(w_1/(w_1 + w_2), w_2/(w_1 + w_2))$ and then to all of $[0, \infty)^2$ by defining $\sigma(0, 0)$ to be any specific move. Then, $\sigma(\cdot, \cdot)$ is a positive homogeneous function of order 0 defined on $[0, \infty)^2$. Then for matching pennies, $\sigma(4, 7) = \sigma(4/11, 7/11) = \sigma(2 \cdot 4, 2 \cdot 7) = 1; \sigma(12, 12) = \sigma(1/2, 1/2) = \sigma(7, 7) = 1/2$. Note, for example, that with $Z = (Z_1, Z_2)$ and h a positive constant, $\sigma(1 - \pi + hZ_1, \pi + hZ_2) = 1$ if and only if $(Z_2 - Z_1) > (1 - 2\pi)/h$. In the Hannan-type strategy (2.2), a random perturbation is used, in particular, (Z_1, Z_2) is a random vector. Thus the expected Bayes response (2.2) is a probability distribution on Player II's pure moves, specifically, $P(1) = Prob(Z_2 - Z_1 > (1 - 2\pi)/h) + 0.5 \cdot Prob(Z_2 - Z_1 =$

$(1-2\pi)/h)$ and $P(0) = Prob(Z_2 - Z_1 < (1-2\pi)/h) + 0.5 \cdot Prob(Z_2 - Z_1 = (1-2\pi)/h)$.

Example 2.3 *(Matching 3-sided pennies)* Here Player I and Player II have $m = n = 3$ pure moves which we denote as $\{1, 2, 3\}$. Consider the Player II loss matrix shown below

		Player II		
		side 1	side 2	side 3
	side 1	0	1	1
Player I	side 2	1	0	1
	side 3	1	1	0

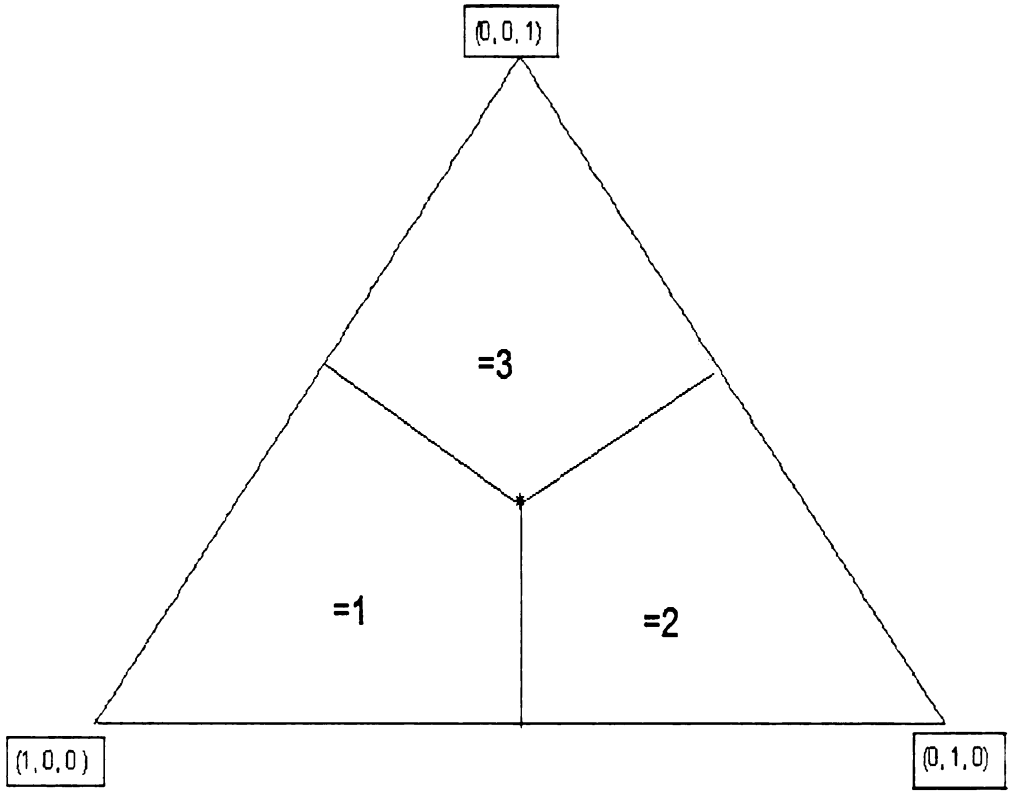
Suppose that Player I selects his/her moves with the (prior) probability distribution $P(1) = \pi_1$, $P(2) = \pi_2$, $P(3) = \pi_3$. A Bayes response for Player II defined on the probability simplex in R^3 must satisfy

$$\sigma(\pi_1, \pi_2, \pi_3) = \begin{cases} 1, & \pi_1 > \pi_2 \vee \pi_3, \\ 2, & \pi_2 > \pi_1 \vee \pi_3, \\ 3, & \pi_3 > \pi_1 \vee \pi_2. \end{cases}$$

These (π_1, π_2, π_3)-sets are the interiors of the convex regions shown in the probability simplex in figure 2.2.

The domain of σ can be extended to the boundaries by any choices of probability distributions supported on the maximizing coordinates. For example, $\sigma(0.25, 0.40, 0.35) = 2 = (0, 1, 0)$ and one could take $\sigma(0.35, 0.35, 0.30) = (1/2, 1/2, 0)$ and $\sigma(1/3, 1/3, 1/3) = (1/3, 1/3, 1/3)$. Note, for example, that if the domain of the function σ is extended to all of $[0, \infty)^3$ as a positive homogeneous function of order 0, then $\sigma(\pi_1 + hZ_1, \pi_2 + hZ_2, \pi_3 + hZ_3) = (1, 0, 0)$ if $(Z_1 - Z_2) > (\pi_2 - \pi_1)/h$ and $(Z_1 - Z_3) > (\pi_3 - \pi_1)/h$ where h is a positive constant. In the Hannan-type strategy (2.2), a random perturbation is used, in particular, $Z = (Z_1, Z_2, Z_3)$. The expected Bayes response (2.2) is then a probability distribution on Player II's pure moves.

Figure 2.2: simplex



2.4 Need for Fresh Randomizations

Consider the situation where Player II's moves are probabilities p (the weatherman example) or, more generally, probability distributions or expectations. Contrast this with the situation where Player II is forced to play the realization of his/her randomization. For example, in matching pennies Player II is required to select a 0 or a 1, albeit, he/she may generate the move with a probability distribution. Because the histories of Player II's past moves are available to Player I, Player II must be concerned about the joint distribution of the random variables he/she generates across the stages of the sequence.

Hannan (1957) did not deal with this issue since his concern was with repeated play where II's moves were probability distributions or expectations and component loss was measured following expectation over the randomization. To be more specific, a single random variable $Z \sim F$ (serving like a dummy variable of integration) is used in describing the Hannan moves $E(\sigma(G_1 + h_1 \cdot Z))$, $E(\sigma(G_2 + h_2 \cdot Z))$, $E(\sigma(G_3 + h_3 \cdot Z))$,... in his theorems.

Suppose that Player II must play a move b_2 at stage $t = 2$ determined by the Bayes response $\sigma(G_1 + h_1 \cdot Z)$, a move b_3 at stage $t = 3$ determined by the Bayes response $\sigma(G_2 + h_2 \cdot Z)$,.... In this case, information on the realization of Z passes to Player I through the the sequence $b_2, b_3, \ldots$. (Player II applying $\sigma(G_1 + h_1 \cdot Z_1)$, $\sigma(G_2 + h_2 \cdot Z_2)$, $\sigma(G_3 + h_3 \cdot Z_3)$,... with iid $Z_i \sim F$ removes this possibility for Player I.) However, we take the matching pennies example to show how Player II can be trapped if employing a Hannan-type strategy based on a single randomization Z.

Recall our matching pennies example and consider the Hannan-type strategy.

$$\sigma(1 - g_{t-1} + h_{t-1}Z_1, g_{t-1} + h_{t-1}Z_2) = [g_{t-1} > \frac{1}{2} + \frac{h_{t-1}(Z_1 - Z_2)}{2}]$$

where g_{t-1} is the proportion of 1's in the sequence of Player I moves from stage 1 to stage t-1. In our example we take the scale factor $h_t = 2/\sqrt{t}$ and let $U = Z_1 - Z_2$ where (Z_1, Z_2) is uniformly distributed in the unit square $[0, 1]^2$. Then

$$b_t = [g_{t-1} > \frac{1}{2} + \frac{U}{\sqrt{t-1}}],\ t = 2, 3, \ldots$$

where $U \in [-1, 1]$.

Consider this strategy for Player I: $a_1 = 0$, $a_2 = 0$, $a_3 = 1$ and

$$a_t = \begin{cases} a_{t-1} & \text{if } b_{t-1} = 1 - a_{t-1}, \\ 1 - a_{t-1} & \text{if } b_{t-1} = a_{t-1}. \end{cases}$$

Thus, Player I continuous to play the same move until he /she observes that Player II has matched his/her move. Assume that Player II generates his/her first move b_1 and the randomization (Z_1, Z_2) independently with $P(b_1 = 0) > 0$ so that $P(b_1 =$

$0, U > 0) > 0$. We will show that on the event $(b_1 = 0, U > 0)$ that Player I wins at least 3/5ths of the time in the limit so that $\underline{\lim}_N D_N(\underline{a}, \underline{b}) \geq \frac{3}{5} - \frac{1}{2} = \frac{1}{10}$. Since the event has positive probability, there are move sequences $\underline{a}$ where the strategy $\underline{b}$ is not Hannan consistent. (If $P(b_1 = 0) = 0$, then $P(b_1 = 1) = 1$ and analysis of the Player I strategy starting with 1 1 0 will lead to a similar conclusion.)

Assume that $U \geq 0$. Let $N_0 + 2$ be the maximum of stage before Player I switches his/her play to begin to play the opposite, i.e

$$g_{N_0} = \frac{N_0 - 2}{N_0} \leq 0.5 + \frac{U}{\sqrt{N_0}}, \quad but \quad g_{N_0+1} = \frac{N_0 - 1}{N_0 + 1} > 0.5 + \frac{U}{\sqrt{N_0 + 1}}.$$

From the inequality of g_{N_0} and $g_{N_0} = \frac{N_0 - 2}{N_0}$, we have

$$0.5 + \frac{U}{\sqrt{N_0}} - \frac{N_0 - 2}{N_0} > 0$$

i.e.,

$$0 < N_0 < 2U^2 + 4 + 2U \cdot \sqrt{U^2 + 4} = (U + \sqrt{U^2 + 4})^2$$

where N_0 is the maximum integer number such that $N_0 \leq 2U^2 + 4 + 2U \cdot \sqrt{U^2 + 4}$. Since $U \in [0, 1]$, the maximum of N_0 is achieved when $U = 1$, and minimum at $U = 0$.

$$\max N_0 = 10 \; when \, U = 1; \quad \min N_0 = 4 \; when \, U = 0.$$

i.e., $4 \leq N_0 \leq 10$, for all $U \in [0, 1]$.

From stage 3 to stage N_0, $N_0 + 1$ and $N_0 + 2$, Player I still plays 1. And Player II plays 0 from stage 1 to stage N_0. Then at stage $N_0 + 1$, Player II observe $g_{N_0+1} = \frac{N_0 - 1}{N_0 + 1} > 0.5 + \frac{U}{\sqrt{N_0+1}}$, so Player II switch his play from 0 to 1 at stage $N_0 + 1$. Then at stage $N_0 + 2$, Player I begins to play 0.

Player II switches his play at stage $N_0 + 2$, then Player I switches his play from 1

Table 2.1: Player I and Player II's dynamic play to $N = N_0 + 3$

Stage	1	2	3	4	5	...	N_0	N_0+1	N_0+2	N_0+3	...
Player I	0	0	1	1	1	...	1	1	1	0	...
Player II	0	0	0	0	0	...	0	0	1	0	...

Table 2.2: Player I and Player II's dynamic play to $N = N_0 + +m_1 + 3$

Stage	Player I	Player II
...	...	...
N_0	1	0
N_0+1	1	0
N_0+2	1	1
N_0+3	0	1
...	...	...
N_0+m_1	0	1
N_0+m_1+1	0	1
N_0+m_1+2	0	0
N_0+m_1+3	1	0
...	...	...

to 0 at stage $N_0 + 3$.

Let m_1 be the number of stages needed for Player II to switch his play back to 1 after stage $N_0 + 2$. i.e.

As Player I and Player II's play are listed above, the total number of 1's in Player I's play from stage 1 to stage $N_0 + m_1$. thus $g_{N_0+m_1} = \frac{N_0}{N_0+m_1}$. And by the definition of m_1 and N_0,

$$\begin{aligned} g_{N_0+m_1} &= \frac{N_0}{N_0+m_1} > 0.5 + \frac{U}{\sqrt{N_0+m_1}} \\ g_{N_0} &= \frac{N_0-2}{N_0} \leq 0.5 + \frac{U}{\sqrt{N_0}}. \end{aligned}$$

$$(0.5 + \frac{U}{\sqrt{N_0}} - \frac{N_0-2}{N_0}) \cdot (0.5 + \frac{U}{\sqrt{N_0+m_1}} - \frac{N_0}{N_0+m_1}) \leq 0.$$

i.e., m_1 is the largest integer such that

$$0.5 + \frac{U}{\sqrt{N_0 + m_1}} - \frac{N_0}{N_0 + m_1} \leq 0.$$

therefore, we have

$$\sqrt{N_0 + m_1} \leq -U + \sqrt{U^2 + 2N_0}$$

i.e.,

$$m_1 \leq 2U^2 + N_0 - 2U \cdot \sqrt{2N_0 + U^2}.$$

Lemma 2.4.1. *With the results about N_0 above, we have the following bounds for m_1, $2 \leq m_1 \leq 4$, for all $U \in [0, 1]$. And m_1 reaches its minimum at $U = 1$, and maximum at $U = 0$. Similarly, we have the same conclusion for m_3, m_5, m_7, ... ,i.e*

$$2 \leq m_j \leq 4, \quad j = 1, 3, 5, 7, \ldots$$

Proof: For m_1, $(U^2 + 4) \leq N_0 \leq (U + \sqrt{U^2 + 4})^2$, N_0 is the maximum integer to satisfy this inequality.

$$\begin{aligned} m_1 &\leq (U - \sqrt{2N_0 + U^2})^2 - (\sqrt{N_0})^2 \\ &= (U - \sqrt{2N_0 + U^2} + \sqrt{N_0}) \cdot (U - \sqrt{2N_0 + U^2} - \sqrt{N_0}) \\ &\leq (2U + \sqrt{U^2 + 4} - \sqrt{2N_0 + U^2}) \cdot (-\sqrt{2N_0 + U^2} + (U - \sqrt{U^2 + 4})) \end{aligned}$$

Obviously $-\sqrt{2N_0 + U^2} + (U - \sqrt{U^2 + 4}) < 0$, with the fact that $m_1 > 0$,

$$\sqrt{2N_0 + U^2} \geq 2U + \sqrt{U^2 + 4}$$

then

$$\begin{aligned} m_1 &\leq 2U^2 + N_0 - 2U \cdot (2U + \sqrt{U^2+4}) \\ &= N_0 - 2U^2 - 2U \cdot \sqrt{U^2+4} \\ &= U^2 + U^2 + 4 + 2U \cdot \sqrt{U^2+4} - 2U^2 - 2U \cdot \sqrt{U^2+4} \\ &= 4. \end{aligned}$$

For the lower bound of m_1, assume $m_1 < 2$, i.e $0 < m_1 \leq 1$, then

$$2U^2 + N_0 - 2U \cdot \sqrt{2N_0 + U^2} - 1 < m_1 \leq 1.$$

Thus,

$$\begin{aligned} (2U^2 + N_0 - 2)^2 &\leq 2N_0 + U^2 \\ N_0^2 + (4U^2 - 6) \cdot N_0 + 4U^4 - 9U^2 + 4 &\leq 0 \\ N_0 &\leq 3 - 2U^2 + \sqrt{5 - 3U^2} \end{aligned}$$

Then, we have $N_0 \leq 2.4$ when $U = 1$. Contradiction with $N_0 = 10$ when $U = 1$ from previous discussion on N_0. Therefore, $m_1 \geq 2$. Together with the first part proof, $2 \leq m_1 \leq 4$. With similar argument, for m_j, $j = 1, 3, 5, 7, ..$, we all have

$$2 \leq m_j \leq 4.$$

(for example, to prove m_3, one just need to replace N_0 by N_2, which equals $N_0 + m_1 + m_2$) Proof is done. □

Let $N_1 = N_0 + m_1$, proportion of 1 from stage 1 to stage N_1 has property that $g_{N_1} \geq 0.5 + \frac{U}{N_1}$. Let m_2 be the number of stages needed to switch play back to 0.

Then, $g_{N_1+m_2} = \frac{N_1-m_1+m_2-2}{N_1+m_2} \leq 0.5 + \frac{U}{\sqrt{N_1+m_2}}$.Therefore,

$$(0.5 + \frac{U}{N_1} - g_{N_1}) \cdot (0.5 + \frac{U}{\sqrt{N_1 + m_2}} - \frac{N_1 - m_1 + m_2 - 2}{N_1 + m_2}) \leq 0$$

and m_2 is the largest integer such that

$$0.5 + \frac{U}{\sqrt{N_1 + m_2}} - \frac{N_1 - m_1 + m_2 - 2}{N_1 + m_2} > 0.$$

Then, solve for m_2, with the property of m_1, $m_1 - N_0 \leq 2U^2 - 2U \cdot \sqrt{2N_0 + U^2}$

$$m_2 \leq 4U^2 + 4 - 2U \cdot \sqrt{2N_0 + U^2} + 2U \cdot \sqrt{U^2 + 2m_1 + 4}.$$

Lemma 2.4.2. *With the discussion of m_1 and N_0, we claim that $3 \leq m_2 \leq 4$. Similarly, for m_4, m_6, m_8,... we also have $3 \leq m_j \leq 4$, $j = 2, 4, 6, 8, \ldots$*

Proof: For m_2, claim that $m_2 \leq 4$.

If $m_2 > 4$, i.e. $m_2 \geq 5$, since m_2 is integer and satisfy

$$4U^2 + 4 - 2U \cdot \sqrt{2N_0 + U^2} + 2U \cdot \sqrt{U^2 + 2m_1 + 4} \geq m_2 \geq 5.$$

Then

$$4U^2 + 2U \cdot \sqrt{U^2 + 2m_1 + 4} \geq 1 + 2U \cdot \sqrt{2N_0 + U^2}$$

i.e when we assume $U \neq 0$,

$$\sqrt{U^2 + 2m_1 + 4} \geq \frac{1}{2U} - 2U + \sqrt{2N_0 + U^2}.$$

By previous discussion, $\sqrt{2N_0 + U^2} > 2U + \sqrt{U^2 + 4}$, thus the right hand side is > 0.

Then,

$$
\begin{aligned}
2m_1 &\geq (\frac{1}{2U} - 2U + \sqrt{2N_0 + U^2})^2 - U^2 - 4 \\
&\geq (\frac{1}{2U} + 2U + \sqrt{U^2+4} - 2U)^2 - U^2 - 4 \\
&= \frac{1}{4U^2} + U^2 + 4 + \frac{1}{U} \cdot \sqrt{U^2+4} - U^2 - 4 \\
&= \frac{1}{4U^2} + \frac{\sqrt{U^2+4}}{U} \longrightarrow \infty \quad U \longrightarrow 0.
\end{aligned}
$$

Contradiction with $2 \leq m_1 \leq 4$ from previous discussion. Thus, we have the conclusion that $m_2 \leq 4$.

If $m_2 < 3$, i.e. $m_2 \leq 2$, since m_2 must be a integer.

$$
\begin{aligned}
4U^2 + 4 - 2U \cdot \sqrt{2N_0 + U^2} + 2U \cdot \sqrt{U^2 + 2m_1 + 4} - 1 &\leq 2 \\
4U^2 - 2U \cdot \sqrt{2N_0 + U^2} + 2U \cdot \sqrt{U^2 + 2m_1 + 4} &\leq -1
\end{aligned}
$$

for all $U \in [0, 1]$.

However, when $U = 0$, we have left side$=0 \leq -1$. Contradiction!

Therefore, based on all the discussion we have above, we have $3 \leq m_2 \leq 4$. Similar proof for m_4, m_6, m_8,... In another words, we have $3 \leq m_j \leq 4$, for $j = 2, 4, 6, 8, \ldots$.

Proof is done.$\square$

According to the two propositions above, we can consider the Player II's total loss. Since Player II only wins twice in each cycle (each cycle contains m_{2k-1} plus m_{2k} stages which is of length at least 5) and since Player I plays 0 at stage 1, Player II's win$\leq 2 + \frac{N-N_0}{5} \cdot 2 - 2$. Thus, Player II's total loss is $\geq N - \frac{N-N_0}{5} \cdot 2$, i.e Player II's total loss is $\geq \frac{3}{5}N + \frac{2N_0}{5} - 2$. Therefore the Cesaro loss is

$$
CL_N(\underline{a}) = \frac{3}{5} + O(\frac{1}{N}).
$$

At the other side, the Bayes envelope $g_N \wedge (1-g_N) \le \frac{1}{2}$. Therefore,

$$D_N = CL_N(\underline{a}) - g_N \wedge (1-g_N) \ge \frac{1}{10} + O(\frac{1}{N}).$$

Example 2.3 *(Simulation of PARP without refreshing randomness)*

Suppose the randomness used in Player II's strategy $U = 0.7$. Player I and Player II play as we describe at the beginning of this section. Then, the Player II's average loss sequence and Bayes envelope at each stage are showed by the graph below. and the simulation of the first 27 stages of Player I and Player II are listed by the following table.

Figure 2.3: Non refresh randomness U=0.7.

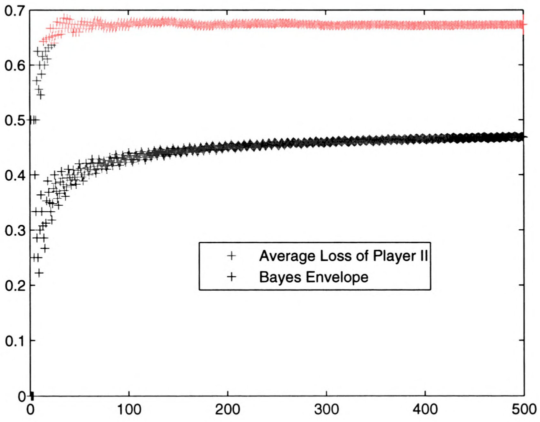


Table 2.3: Non refresh randomness example with U=0.7.

stage	Player I	g_t	bar_t*	Player II	$Loss_t$
1	0	0	1.2000	0	0
2	0	0	0.9950	0	0
3	1	0.3333	0.9041	0	0
4	1	0.5000	0.8500	0	1
5	1	0.6000	0.8130	0	1
6	1	0.6667	0.7858	0	1
7	1	0.7143	0.7646	0	1
8	1	0.7500	0.7475	0	1
9	1	0.7778	0.7333	1	0
10	0	0.7000	0.7214	1	1
11	0	0.6364	0.7111	0	0
12	1	0.6667	0.7021	0	1
13	1	0.6923	0.6941	0	1
14	1	0.7143	0.6871	0	1
15	1	0.7333	0.6807	1	0
16	0	0.6875	0.6750	1	1
17	0	0.6471	0.6698	1	1
18	0	0.6111	0.6650	0	0
19	1	0.6316	0.6606	0	1
20	1	0.6667	0.6528	0	1
21	1	0.6818	0.6492	1	0
22	0	0.6522	0.6460	1	1
23	0	0.6250	0.6429	1	1
24	0	0.6000	0.6400	0	0
25	1	0.6154	0.6373	0	1
26	1	0.6296	0.6347	0	1
27	1	0.6429	0.6323	0	1
28	1	0.6552	0.6300	1	0
29	0	0.6333	0.6278	1	1
...	...	...	...	...	...

* $bar_t = 0.5 + \frac{U}{\sqrt{(t)}}$ is the threshold for decision based on g_t.

All the discussion and simulation example show that it is necessary to refresh the randomness at each stage when we use Hannan type strategy to make decision.

Chapter 3

PARP Strategy for k-extended Envelope Problem

3.1 Introduction

Practical forecasting problems are of great variety. Sometimes we suspect that Nature or the market (as our Player I) makes its decision by some patterns. The decision on one stage may be affected by the previous k stage decisions. For example, the market gives rise to a certain stock price. This may raise investors confidence and this confidence or followup may make the market give another increase the next day.

Therefore, we are motivated to study such kinds of pattern behavior. Suppose the Player I's moves on a_t are affected by a_{t-k}, a_{t-k+1}, $\ldots$, a_{t-1}, then with this situation, our Bayes envelope is called *k-extended Bayes envelope*, and the corresponding forecasting problem is called *k-extended Bayes envelope problem.*

In this chapter, we will give definitions and extensions of PAP strategy and PARP strategy for the two-extended envelope problem. Although we focus on the two-extended envelope problem, it is easy to generalize the two-extended envelope case to k-extended cases.

3.2 Envelopes including Extended Envelopes

We have already introduced the simple envelope R for the evaluation of average loss in the repeated play of matching pennies. Hannan (1957, Sec 3) defines the simple envelope at stage N as the total loss $N \cdot CL_N$ to Player II had II known the empirical distribution of I's moves $a_1, a_2, \ldots, a_N$ and played Bayes against this distribution at each stage $t = 1, 2, \ldots, N$.

We consider what are called extended envelopes for repeated play, first introduced by Swain (1965) and Johns (1967) for the repeated play of a statistical decision problem and first analyzed and ordered in general terms in Gilliland and Hannan (1969). Extended envelopes can be defined as minimum average loss across specified sets of strategies including those chosen to take advantage of possible Markov-type structure in the empirical distribution of I's moves.

Example 3.2.1 Repeated Play of Matching Pennies.
Consider the repeated play of matching pennies and the collection of strategies $S = \{\underline{p}^{(0)}, \underline{p}^{(1)}, \underline{p}^{(2)}, \underline{p}^{(3)}\}$ where $\underline{p}^{(0)}$ and $\underline{p}^{(1)}$ were defined in example 1.1 and

$$p_1^{(2)} = a_N \quad and \quad p_t^{(2)} = a_{t-1}, \quad t = 2, 3, \ldots, N$$

and

$$p_1^{(3)} = 1 - a_N \quad and \quad p_t^{(3)} = 1 - a_{t-1}, \quad t = 2, 3, \ldots, N$$

These may be thought of as a *stay strategy* and *switch strategy*, respectively, although the moves by strategy $\underline{p}^{(i)}$ at stage 1 are not possible given the rules for the repeated play. The extended envelope of order 2 is given by

$$R^{(2)}(\underline{a}_N) = min\{CL_N(\underline{a}_N, \underline{p}_N) | \underline{p} \in S\}$$

As the minimum over a larger set of strategies,

$$R^{(2)}(\underline{a}_N) \leq R^{(1)}(\underline{a}_N) = R(\underline{g}_N) \quad for \quad all \quad \underline{a}_N$$

Thus, 2-extended envelope $R^{(2)}$ is a more stringent envelope against which to compare the average loss of a Player II strategy.

For the explicit evaluation of the extended envelope $R^{(2)}$ it is useful to consider all consecutive pairs contained in the sequence $\underline{a}_N$ understood to be wrapped in a circle so that a_N precedes a_1. There are N consecutive, overlapping pairs. Let n_{ij} denote the count of the pairs with first component i and second component j, $i, j = 0, 1$. Then it follows that $n_1 := n_{10} + n_{11} = Ng_N$ =number of 1's in the sequence $\underline{a}_N$ and $n_0 := n_{00} + n_{01} = N(1 - g_N)$ =number of 0's in the sequence $\underline{a}_N$, and $n_{01} = n_{10}$.

Proposition 3.2.1. *If n_{ij} are defined as above, then*

$$NR^{(2)}(\underline{a}_N) = (n_{01} \wedge n_{00}) + (n_{11} \wedge n_{10})$$

Proof:

From the definition of n_{ij} and CL_N, we have

$$\begin{aligned}
N \cdot CL_N(\underline{a}_N, \underline{p}_N^{(0)}) &= n_1 = n_{10} + n_{11} \\
N \cdot CL_N(\underline{a}_N, \underline{p}_N^{(1)}) &= n_0 = n_{01} + n_{00} \\
N \cdot CL_N(\underline{a}_N, \underline{p}_N^{(2)}) &= n_{01} + n_{10} = 2n_{01} \\
N \cdot CL_N(\underline{a}_N, \underline{p}_N^{(3)}) &= n_{00} + n_{11}
\end{aligned}$$

Proof can be completed by considering the four situations: ($n_{01} \leq n_{00}$ and $n_{01} \leq n_{11}$), ($n_{01} \leq n_{00}$ and $n_{01} \geq n_{11}$), ($n_{01} \geq n_{00}$ and $n_{01} \leq n_{11}$) and ($n_{01} \geq n_{00}$ and $n_{01} \geq n_{11}$).

Example 3.2.2 Let $N = 17$, and $\underline{a}_{17}$: (0,1,1,1,0,1,0,1,1,0,0,0,1,1,0,1,0). In this case $n_{00} = 3$, $n_{01} = n_{10} = 5, n_{11} = 4$.

Then 2-extended envelope $17 \cdot R^{(2)}(\underline{a}_{17}) = (3 \wedge 5) + (4 \wedge 5) = 7$. While, $17 \cdot R(9/17) = 17 \cdot R^{(1)}(\underline{a}_{17}) = (3+5) \wedge (4+5) = 8$.

Consider another example where the sequence of Player I moves shows greater second order dependency, Markov structure. Take $N = 17$, and

$\underline{a}_{17}$: (0.1.0.1.0.1.0.1.1.0.0.1.1.1.0.1.0). In this case $n_{00} = 2$, $n_{01} = n_{10} = 6, n_{11} = 3$.

Then 2-extended envelope $17 \cdot R^{(2)}(\underline{a}_{17}) = (2 \wedge 6) + (3 \wedge 6) = 5$. While, $17 \cdot R(9/17) = 17 \cdot R^{(1)}(\underline{a}_{17}) = (2+6) \wedge (3+6) = 8$.

The extended envelope idea can be based on three-tuples, four-tuples, etc. and leads to an ordering for a family of k envelopes

$$R^{(k)}(\underline{a}_N) \leq R^{(k-1)}(\underline{a}_N) \leq \cdots \leq R^{(1)}(\underline{a}_N) = g_N \wedge (1 - g_N).$$

Regret relative to the 2-extended envelope is

$$D_N^{(2)}(\underline{a}, \underline{p}) = CL_N(\underline{a}_N, \underline{p}_N) - R^{(2)}(\underline{a}_N).$$

This Chapter includes the adaptation of the PARP strategy (defined in Chapter 2) to matching pennies that achieves uniform Hannan consistency with the 2-extended envelope, i. e.,

$$D_N^{(2)}(\underline{a}, \underline{parp}) = O(N^{-\frac{1}{2}}) \quad uniformly \quad in \quad \underline{a}.$$

Remark:

The wrapping of the sequence $\underline{a}_N$ gives an ordering to the resulting envelopes, an idea exploited in Gilliland and Hannan (1969). However, in developing Hannan consistent strategies in repeated play with the k-extended envelope, there are only $t - k$ past k-tuples of Player I consecutive moves available to Player II for basing a move at stage t, $t = k+1, k+2, \ldots$. Thus, the regrets studied in the following sections are relative to envelopes based on the empirical distribution of the $N - k + 1$ (not N) k-

tuples $(a_1, a_2, \ldots, a_k)$, $(a_2, a_3, \ldots, a_{k+1})$, $\ldots$ $(a_{N-k+1}, a_{N-k+2}, \ldots, a_N)$. The difference in regrets compared to those relative to envelopes based on the N k-tuples from the wrapped sequences is for fixed k of order $O(1/N)$ since only k of the k-tuples are omitted at each stage N.

In the rest of this chapter we consider the repeated play of matching pennies. In section 3.3 we continue discussion of the 2-extended envelope. In section 3.4 and 3.5 we derive exact expressions for the Cesaro loss from PAP and PAP+ applied to empirical distribution of past. In section 3.6, we show how PARP applies to give Hannan consistency for the 2-extended problem.

3.3 Two-extended envelope problem

Let's consider Player I's moves in an actual game in $N = 13$ trials.

Table 3.1: An example of Player I's actual moves in N=13 trials

stage	1	2	3	4	5	6	7	8	9	10	11	12	13
Payer I	0	1	0	1	1	0	1	0	1	0	0	1	0

Now as Player II, we want to predict Player I's move on the $(N+1)^{th}$ trial.

In order to do the prediction, we have considered PAP and also discussed the strategy based on random past (PARP) which is randomized from the empirical distribution in previous chapters. However, another question arises: Does Player I more likely play 1 following a 0 or more likely play 0 following 0. In another words, is it possible that Player I is playing on a certain 'pair' pattern.

Here, we explore a strategy to deal with such kind of 'pair' pattern moves of Player I, i.e. our Bayes responses will be based on the past 'pairs'.

In general, Player I 's move is sequence $\underline{a}$ with N trials:

$$\underline{a}_N : a_1, \quad a_2, \quad a_3, \ldots, \quad a_{N-1}, \quad a_N.$$

By pairing Player I's move on each trial and its next trial, we have a set $\widetilde{a}_N$:

$$\widetilde{\underline{a}}_N : \{(a_1, a_2), \quad (a_2, a_3), \quad (a_3, a_4), \ldots, \quad (a_{N-1}, a_N)\}.$$

For each pair in $\widetilde{\underline{a}}_N$, we take the first coordinates as the condition. Then, we can get two partition sets of $\widetilde{\underline{a}}_N$ with respect to the condition of each pair 0 or 1, and during this partition procedure, we keep the order of these pairs.

Therefore, by partition with the first coordinate as the condition, we have:

$$A_N : \{(a_i, a_j)| \quad a_i = 0, \quad , i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N\}$$

$$B_N : \{(a_i, a_j)| \quad a_i = 1, \quad , i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N\}$$

Suppose Player I's move on the N^{th} is $a_N = 0$. Our Bayes response b_{N+1} is the move under the condition, preceding move is 0. Naturally, our prediction should be based on Player I's pair move under the same condition, i.e, on set A_N.

In set A_N, all the second coordinates of each pair are Player I's move with condition: the preceding move is 0. Therefore, we take out the second coordinates of each pair, keep their orders and put them together to form a new sequence $\widetilde{A}_N$. The new sequence $\widetilde{A}_N$ contains all the behavior of Player I with the condition: the preceding move is 0.

Take the example at the beginning of this section:

$$\underline{a}_N : 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0$$

We have:

$$A_N : \{(0, 1), \quad (0, 1), \quad (0, 1), \quad (0, 1), \quad (0, 0)\}$$

$$B_N : \{(1, 0), \quad (1, 1), \quad (1, 0), \quad (1, 0), \quad (1, 0), \quad (1, 0)\}$$

Since $a_N = 0$, like what we discussed above, we consider A_N, and take out all the second coordinates of each pairs in A_N to form a new sequence:

$$\widetilde{A}_N : 1, \quad 1, \quad 1, \quad 1, \quad 0.$$

If we also include the pair (a_N, a_{N+1}), $\widetilde{A}_N$ becomes:

$$\widetilde{A}_N : 1, \quad 1, \quad 1, \quad 1, \quad 0, \quad a_{N+1}.$$

where a_{N+1} is the move we want to predict.

In sequence $\widetilde{A}_N$, each term is an individual play with the same condition: proceeding move is 0. In another word, we can consider each term of $\widetilde{A}_N$ as an individual move. Therefore, we can use the strategies we have discussed already in previous chapter and the PARP from Gilliland and Jung (2006).

3.4 PAP Strategy in two-extended envelope problem

We apply PAP strategy on the sets $\widetilde{A}_N$ and $\widetilde{B}_N$ which we have constructed in last section.

Theorem: In this kind of two-extended envelope problem, when the PAP strategy is used to do the $(N+1)$ stage prediction, the Cesaro expected loss sequence for his PAP strategy $\underline{p}$ is given by:

$$\begin{aligned} CL_N^{(2)} = \frac{0.5}{N} &+ \frac{1}{N} \cdot (n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N}) + \frac{0.5}{N} \cdot (\nu_{n1} + \mu_{n2}) \\ &+ \frac{0.5}{N} \cdot [n'_{00,N} \neq n'_{01,N}] + \frac{0.5}{N} \cdot [n'_{10,N} \neq n'_{11,N}] \end{aligned}$$

where

$$
\begin{aligned}
n'_{ij,N} &= number\, of\, (a_k, a_{k+1}) = (i,j)\ , for\, k = 1,2,3,\dots,N-1; \\
n1 &= (n'_{00,N} + n'_{01,N}), \\
\nu_{n1} &= \{number\, of\, times\, of\, g_k = 0.5\ |k = 1,2,3,\dots,n1\}
\end{aligned}
$$

g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence A_N.

$$
\begin{aligned}
n2 &= (n'_{10,N} + n'_{11,N}) \\
\mu_{n2} &= \{number\, of\, times\, of\, g_k = 0.5\ |k = 1,2,3,\dots,n2\}
\end{aligned}
$$

g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence B_N. Furthermore,

$$
\begin{aligned}
\min_{\underline{a}_N} &\ \{\ CL_N^{(2)}(\underline{a}_N, \underline{p}_N)|fixed\quad n'_{ij,N},\ i = 0,1,\ ,j = 0,1\} \\
&= \frac{0.5}{N} + \frac{1}{N}\cdot(n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N})
\end{aligned}
$$

and

$$
\begin{aligned}
\max_{\underline{a}_N} &\ \{\ CL_N^{(2)}(\underline{a}_N, \underline{p}_N)|\, fixed\, n'_{ij,N},\ i = 0,1,\ ,j = 0,1\} \\
&= \frac{0.5}{N} + 1.5\cdot\frac{1}{N}\cdot(n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N}) \\
&+ \frac{0.5}{N}\cdot([n'_{00,N} \neq n'_{01,N}] + [n'_{10,N} \neq n'_{11,N}])
\end{aligned}
$$

Proof:

For any sequence $\underline{a}_N$, according to four patterns in two-extended envelope problem, (0,0),(0,1),(1,0),(1,1), we can always create set $\widetilde{\underline{a}_N}$ from sequence $\underline{a}_N$ where

$$\widetilde{\underline{a}}_N : \{(a_1, a_2), \quad (a_2, a_3), \quad (a_3, a_4), \ldots, \quad (a_{N-1}, a_N)\}.$$

Since the first coordinate in a pair is 0 or 1, $\widetilde{\underline{a}_N}$ can be divided into two subsequences:

$$A_N : \{(a_i, a_j) |\, a_i = 0, \quad , i = 1, 2, \ldots, N-1,\ j = 2, \ldots, N\}$$

$$B_N : \{(a_i, a_j) |\, a_i = 1, \quad , i = 1, 2, \ldots, N-1,\ j = 2, \ldots, N\}$$

Therefore, we notice that $A_N \cap B_N = \emptyset$ and $A_N \cup B_N = \widetilde{\underline{a}_N}$. Thus,

$$\begin{aligned} N \cdot CL_N^{(2)} &= \textit{Loss on } a_1 + \textit{Loss on second coordinates } A_N \\ &+ \textit{Loss on second coordinates } B_N \end{aligned}$$

Since we always flip a coin on b_1 as the start, Loss on $a_1 = 0.5$.

For sequence A_N, since we only consider loss on second coordinate in each pair, we can form all the second coordinates into a sequence. By using past strategy $\underline{p}$, according to Gilliland and Jung (2006),

$$\begin{aligned} \textit{Loss on second coordinates } A_N &= n1 \cdot \left(\frac{n'_{00,N}}{n'_{00,N} + n'_{01,N}} \wedge \frac{n'_{01,N}}{n'_{00,N} + n'_{01,N}}\right) \\ &= 0.5 \cdot \nu_{n1} + 0.5 \cdot \left[\frac{n'_{00,N}}{n'_{00,N} + n'_{01,N}} \neq 0.5\right] \end{aligned}$$

Since $n1$ is the the number of the pairs in the sequence A_N, i.e. $n1 = (n'_{00,N} + n'_{01,N})$, therefore we can simplify the formula above,

$$\begin{aligned} \textit{Loss on second coordinates } A_N &= (n'_{00,N} \wedge n'_{01,N}) + 0.5 \cdot \nu_{n1} \\ &+ 0.5 \cdot [n'_{00,N} \neq n'_{01,N}] \end{aligned}$$

where,

$$\nu_{n1} = \{number\ of\ times\ of\ g_k = 0.5 \mid k = 1, 2, 3, \ldots, n1\}.$$

g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence A_N.

Similar for sequence B_N,

$$\begin{aligned} Loss\ on\ second\ coordinates\ B_N &= (n'_{10.N} \wedge n'_{11.N}) + 0.5 \cdot \mu_{n2} \\ &+ 0.5 \cdot [n'_{10.N} \neq n'_{11.N}] \end{aligned}$$

where

$$\begin{aligned} n2 &= (n'_{10.N} + n'_{11.N}) \\ \mu_{n2} &= \{number\ of\ times\ of\ g_k = 0.5 \mid k = 1, 2, 3, \ldots, n2\} \end{aligned}$$

g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence B_N.

Thus, the total loss can be written as the following form:

$$\begin{aligned} N \cdot CL_N^{(2)} = 0.5 &+ (n'_{00.N} \wedge n'_{01.N} + n'_{10.N} \wedge n'_{11.N}) + 0.5 \cdot (\nu_{n1} + \mu_{n2}) \\ &+ 0.5 \cdot [n'_{00.N} \neq n'_{01.N}] + 0.5 \cdot [n'_{10.N} \neq n'_{11.N}] \end{aligned}$$

i.e.

$$\begin{aligned} CL_N^{(2)} = \frac{0.5}{N} &+ \frac{1}{N} \cdot (n'_{00.N} \wedge n'_{01.N} + n'_{10.N} \wedge n'_{11.N}) + \frac{0.5}{N} \cdot (\nu_{n1} + \mu_{n2}) \\ &+ \frac{0.5}{N} \cdot [n'_{00.N} \neq n'_{01.N}] + \frac{0.5}{N} \cdot [n'_{10.N} \neq n'_{11.N}] \end{aligned}$$

By Gilliland and Jung (2006), for fixed n1 and $n1 \cdot \frac{n'_{00,N}}{n'_{00,N}+n'_{01,N}}$, minimum loss on the second coordinates on A_N is achieved by the minimum number of ν_{n1}. And

$$\begin{aligned} &min\{loss \quad on \quad the\ second\ coordinates\ on\ A_N\} \\ &= n1 \cdot (\frac{n'_{00,N}}{n'_{00,N}+n'_{01,N}} \wedge \frac{n'_{01,N}}{n'_{00,N}+n'_{01,N}}) + 0.5 \end{aligned}$$

Similar for the sequence B_N.

$$\begin{aligned} &min\{loss \quad on \quad the\ second\ coordinates\ on\ B_N\} \\ &= n2 \cdot (\frac{n'_{10,N}}{n'_{10,N}+n'_{11,N}} \wedge \frac{n'_{11,N}}{n'_{10,N}+n'_{11,N}}) + 0.5 \end{aligned}$$

i.e.

$$\begin{aligned} \min_{\underline{a}_N} \quad &\{ \quad CL_N^{(2)}(\underline{a}_N, \underline{p}_N) | fixed\ N g_{ij}^{(N)},\ i=0,1,\ ,j=0,1\} \\ &= \frac{1.5}{N} + \frac{1}{N} \cdot (n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N}) \end{aligned}$$

Similarly, we can easily derive for the explicit form for the maximum of $CL_N^{(2)}$,

$$\begin{aligned} &max\{loss \quad on \quad the\ second\ coordinates\ on\ A_N\} \\ &= n1 \cdot (\frac{n'_{00,N}}{n'_{00,N}+n'_{01,N}} \wedge \frac{n'_{01,N}}{n'_{00,N}+n'_{01,N}}) + 0.5 \cdot [n'_{00,N} \neq n'_{01,N}] \end{aligned}$$

$$\begin{aligned} &max\{loss \quad on \quad the\ second\ coordinates\ on\ B_N\} \\ &= n2 \cdot (\frac{n'_{10,N}}{n'_{10,N}+n'_{11,N}} \wedge \frac{n'_{11,N}}{n'_{10,N}+n'_{11,N}}) + 0.5 \cdot [n'_{00,N} \neq n'_{01,N}] \end{aligned}$$

Therefore,

$$\begin{aligned}\max_{\underline{a}_N}\{CL_N^{(2)}(\underline{a}_N,\underline{p}_N) \quad | \quad & fixed\ Ng_{ij}^{(N)},\ i=0,1,\ ,j=0,1\} \\ = & \frac{0.5}{N}+1.5\cdot\frac{1}{N}\cdot(n'_{00,N}\wedge n'_{01,N}+n'_{10,N}\wedge n'_{11,N}) \\ + & \frac{0.5}{N}\cdot([n'_{00,N}\neq n'_{01,N}]+[n'_{10,N}\neq n'_{11,N}])\end{aligned}$$

Proof is done.□

Comments 3.4.1: Since the maximum of the $CL_N^{(2)}$ is achieved by the sequence with maximum of ν_N, in another word, max loss on the second coordinate on A_N /n1 is achieved by the maximum of ν_N, i.e.

$$n'_{00,N}=n'_{01,N}$$

as many times as they can in the sequence A_N.

This indicate that the sequence $A_N:\{(0,0),\quad(0,1),\quad(0,0),\quad(0,1),\quad(0,0),\dots\}$ or conversely $A_N:\{(0,1),\quad(0,0),\quad(0,1),\quad(0,0),\quad(0,1),\dots\}$.

Similar for the sequence B_N,
the maximum case is the sequence: $B_N:\{(1,0),\quad(1,1),\quad(1,0),\quad(1,1),\quad(1,0),\dots\}$ or conversely $B_N:\{(1,1),\quad(1,0),\quad(1,1),\quad(1,0),\quad(1,1),\dots\}$.

If we transfer the two sequence back to original sequence form, the maximum case is:

$$\underline{a}_N:0\quad 0\quad 1\quad 1\quad 0\quad 0\quad 1\quad 1\quad 0\quad 0\quad 1\quad 1\quad 0\quad 0,\dots$$

or

$$\underline{a}_N:1\quad 1\quad 0\quad 0\quad 1\quad 1\quad 0\quad 0\quad 1\quad 1\quad 0\quad 0\quad 1\quad 1,\dots$$

For the minimum case, since the minimum of $CL_N^{(2)}$ is obtained by the sequence with the minimum of ν_N i.e.

$$n'_{01,N} = n'_{10,N}, \quad n'_{10,N} = n'_{11,N}$$

As few times they can, i.e. $A_N : \{(0,1), \quad (0,1), \quad (0,1), \quad (0,1), \quad (0,1), \ldots\}$ or $B_N : \{(1,0), \quad (1,0), \quad (1,0), \quad (1,0), \quad (1,0), \ldots\}$. In another word, the original sequence is:

$$\underline{a}_N : 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0, \ldots.$$

or reverse the position of 0 and 1.

Corollary 3.4.1: When we take the dimension of the envelope as a higher dimension k, k is a fixed positive integer, we can use the same idea as the theorem we proved above. Here we can take k=3 as an example to show the possibility of this generalization.

For k=3, we have $2^3 = 8$ triple patterns which is the combination of the three dimension with each coordinates has 0 and 1 two choices:

$$(0,0,0) \quad (0,0,1) \quad (0,1,0) \quad (0,1,1) \quad (1,0,0) \quad (1,0,1) \quad (1,1,0) \quad (1,1,1)$$

to transform original sequence $\underline{a}$ into

$$\widetilde{\underline{a}}_N : \{(a_1, a_2, a_3), \quad (a_2, a_3, a_4), \quad (a_3, a_4, a_5), \ldots, \quad (a_{n-2}, a_{N-1}, a_N)\}.$$

Then, we can construct the sets A_1, A_2, A_3 and A_4:

$$A_{1,N} \quad : \{(a_i, a_j, a_k)| \quad a_i = 0,\ a_j = 0\ , i = 1, \ldots, N-2,\ j = 2, \ldots, N-1,\ k = 3, \ldots N\}$$

$$A_{2,N} \quad : \{(a_i, a_j, a_k)| \quad a_i = 0,\ a_j = 1\ , i = 1, \ldots, N-2,\ j = 2, \ldots, N-1,\ k = 3, \ldots N\}$$

$$A_{3,N} \quad : \{(a_i, a_j, a_k)| \quad a_i = 1,\ a_j = 0\ , i = 1, \ldots, N-2,\ j = 2, \ldots, N-1,\ k = 3, \ldots N\}$$

$$A_{4,N} \quad : \{(a_i, a_j, a_k)| \quad a_i = 1,\ a_j = 1\ , i = 1, \ldots, N-2,\ j = 2, \ldots, N-1,\ k = 3, \ldots N\}$$

As we do in the theorem we proved above:

$$\begin{aligned} N \cdot CL_N^{(3)} &= Loss\ on\ a_1 + Loss\ on\ a_2 \\ &+ \sum_{i=1}^{4} Loss\ on\ 3^{rd}\ coordinates\ on\ A_{i,N} \end{aligned}$$

And

$$\begin{aligned} &Loss \quad on \quad 3^{rd}\ coordinates\ on\ A_{1,N} \\ &= n1 \cdot \left(\frac{n'_{000,N}}{n'_{000,N} + n'_{001,N}} \wedge \frac{n'_{001,N}}{n'_{000,N} + n'_{001,N}}\right) + 0.5 \cdot \nu_{n1} \\ &+ 0.5 \cdot \left[\frac{n'_{001,N}}{n'_{000,N} + n'_{001,N}} \neq 0.5\right] \\ &= (n'_{000,N} \wedge n'_{001,N}) + 0.5 \cdot \nu_{n1} + [n'_{000,N} \neq n'_{001,N}] \end{aligned}$$

where $n'_{i,j,l}$ is the number of $(a_k, a_{k+1}, a_{k+2}) = (i, j, l)$ for $k = 1, 2, 3, \ldots, N-2$; $n1$ is the number of the pairs in the sequence A_N.i.e. $n1 = (n'_{000,N} + n'_{001,N})$.

Apply the similar ideas on A_2, A_3 and A_4, and plug them into the formula of

$CL_N^{(3)}$, we have:

$$\begin{aligned}
CL_N^{(3)} &= \frac{1}{N} + \frac{1}{N} \cdot (n'_{000,N} \wedge n'_{001,N} + n'_{010,N} \wedge n'_{011,N} \\
&+ n'_{100,N} \wedge n'_{101,N} + n'_{110,N} \wedge n'_{111,N}) \\
&+ \frac{0.5}{N} \cdot (\nu_{n1} + \mu_{n2} + \omega_{n3} + \gamma_{n4}) \\
&+ \frac{0.5}{N} \cdot ([n'_{000,N} \neq n'_{001,N}] + [n'_{010,N} \neq n'_{011,N}] \\
&+ [n'_{100,N} \neq n'_{101,N}] + [n'_{110,N} \neq n'_{111,N}])
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\min CL_N^{(3)} &= \frac{3}{N} + \frac{1}{N} \cdot (n'_{000,N} \wedge n'_{001,N} + n'_{010,N} \wedge n'_{011,N} \\
&+ n'_{100,N} \wedge n'_{101,N} + n'_{110,N} \wedge n'_{111,N})
\end{aligned}$$

and

$$\begin{aligned}
\max CL_N^{(3)} &= \frac{1}{N} + 1.5 \cdot \frac{1}{N} \cdot (n'_{000,N} \wedge n'_{001,N} + n'_{010,N} \wedge n'_{011,N} \\
&+ n'_{100,N} \wedge n'_{101,N} + n'_{110,N} \wedge n'_{111,N}) \\
&+ \frac{0.5}{N} \cdot ([n'_{000,N} \neq n'_{001,N}] + [n'_{010,N} \neq n'_{011,N}] \\
&+ [n'_{100,N} \neq n'_{101,N}] + [n'_{110,N} \neq n'_{111,N}])
\end{aligned}$$

With the discussion in theorem, comments and corollary above, we can see that using PAP strategy which does not involve any random process, just makes the forecasting decision based on the exact empirical data, will be trapped in some special cases like the maximum example we just showed in comments after the theorem. Therefore, reasonably we would like to more innovative idea to avoid the trapping situations.

3.5 PAP+ Strategy in two-extended envelope problem

Similar with what we discussed in chapter 2, after we construct and studied the PAP strategy in 2-extended envelope problem, we come to consider to play against past plus present (PAP+) strategy.

Theorem: In this kind of two-extended envelope problem, when the PAP+ strategy is used to do the (N) stage prediction, the Cesaro expected loss sequence for his PAP+ strategy $\underline{p}$ is given by:

$$CL_N^{(2)}(\underline{p}) = \frac{0.5}{N} \quad + \quad \frac{1}{N}(n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N}) - \frac{0.5}{N} \cdot (\nu_{n1} + \mu_{n2})$$

where

$$\begin{aligned} n'_{ij,N} &= number\ of\ (a_k, a_{k+1}) = (i,j)\ ,for\ k = 1,2,3,\ldots,N-1; \\ n1 &= (n'_{00,N} + n'_{01,N}), \\ \nu_{n1} &= \{number\ of\ times\ of\ g_k = 0.5\ |k = 1,2,3,\ldots,n1\} \end{aligned}$$

g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence A_N.

$$\begin{aligned} n2 &= (n'_{10,N} + n'_{11,N}) \\ \mu_{n2} &= \{number\ of\ times\ of\ g_k = 0.5\ |k = 1,2,3,\ldots,n2\} \end{aligned}$$

Here, g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence B_N.

Proof:

Similarly in the proof of PAP strategy,for any sequence $\underline{a}_N$, according to four patterns in two-extended envelope problem, (0,0), (0,1), (1,0), (1,1), we defined the four proportions with respect to these four patterns, and also we can always create set $\widetilde{\underline{a}_N}$ from sequence $\underline{a}_N$, where

$$\widetilde{\underline{a}}_N : \{(a_1, a_2), \quad (a_2, a_3), \quad (a_3, a_4), \ldots, \quad (a_{N-1}, a_N)\}.$$

By first coordinate in a pair is 0 or 1, $\widetilde{\underline{a}_N}$ can be divided into two subsequences:

$$A_N : \{(a_i, a_j)| \quad a_i = 0, \quad , i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N\}$$

$$B_N : \{(a_i, a_j)| \quad a_i = 1, \quad , i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N\}$$

Therefore, we notice that $A_N \cap B_N = \emptyset$ and $A_N \cup B_N = \widetilde{\underline{a}}_N$.

Since any a_i, $i = 2, 3, \ldots, N$ and the second coordinates in a pair appear once and only once either in A_N or B_N, and decision on a_i only depend on $\{(a_j, a_k) \mid a_j = a_{i-1}, j = 1, \ldots, i-2, \; k = 2, 3, \ldots, i-1\}$ i.e. loss on decision of a_i in $\underline{a}_N$is equal to the loss on decision of a_i in A_N.

Thus,

$$\begin{aligned} N \cdot CL_N^{(2)} &= \textit{Loss on } a_1 \textit{ and } a_2 + \textit{Loss on second coordinates } A_N \\ &+ \textit{Loss on second coordinates } B_N \end{aligned}$$

Since we always flip a coin on b_1 as the start, Loss on a_1=0.5.

$$\textit{Loss on second coordinates } A_N = (n'_{00,N} \wedge n'_{01,N}) - 0.5 \cdot \nu_{n1}$$

where,

$$\nu_{n1} = \{number\ of\ times\ of\ g_k = 0.5 \mid k = 1, 2, 3, \ldots, n1\}.$$

and here g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence A_N. Similarly, for set B_N

$$Loss\ on\ second\ coordinates\ B_N = (n'_{00.N} \wedge n'_{01.N}) - 0.5 \cdot \mu_{n2}$$

where,

$$\mu_{n2} = \{number\ of\ times\ of\ g_k = 0.5 \mid k = 1, 2, 3, \ldots, n2\}.$$

and here g_k is the empirical proportion of 1 as the second coordinate through stage k in the sequence B_N.

In this way, for PAP+ strategy, the total loss is

$$N \cdot CL_N^{(2)}(p) = 0.5 \quad + \quad (n'_{00.N} \wedge n'_{01.N} + n'_{10.N} \wedge n'_{11.N}) - 0.5 \cdot (\nu_{n1} + \mu_{n2})$$

i.e.

$$CL_N^{(2)}(\underline{p}) = \frac{0.5}{N} \quad + \quad \frac{1}{N}(n'_{00.N} \wedge n'_{01.N} + n'_{10.N} \wedge n'_{11.N}) - \frac{0.5}{N} \cdot (\nu_{n1} + \mu_{n2}).$$

Proof is done. $\square$

The explicit form of the expected loss of PAP+ strategy above shows that when ν_{n1}, and μ_{n2} reach their maximum $n1/2$ and $n2/2$, the expected loss approach to its minimum, which is

$$\min CL_N^{(2)}(\underline{p}) = \frac{0.5}{N} + \frac{1}{N}(n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N}) - 0.25$$

and when when ν_{n1}, and μ_{n2} reach their minimum 0, the expected loss approach to its maximum, which is

$$\max CL_N^{(2)}(\underline{p}) = \frac{0.5}{N} + \frac{1}{N}(n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N}).$$

3.6 PARP Strategy in two-extended envelope problem

As we did in last section for PAPast strategy, we partition the original sequence $\underline{a}_N$ into two sets A_N and B_N, according to the previous stage value as the condition. Then, we extract all the second coordinates of each pair pattern within these two condition subsets, to form two subsequence correspondingly. On these two subsequences, instead of PAPast strategy, we use PARP strategy ideas.

For each fixed n, when we forecast the next stage i.e. the $(N+1)$ stage, in sequence a_N, we are going to do the forecast base on the N^{th} stage a_N's value. If $a_N = 0$, we will use the subsequence $\widetilde{A}_N$ from the empirical data set A_N. By PARP strategy idea, we will do bootstrap sampling in this subsequence, and use the majority from the bootstrap sampling result as our forecast for a_{N+1}. Different from the PARP strategy we introduced in previous chapters, this is an conditional PARP strategy. The object of applying PARP strategy in not the original sequence a_N any more, but the conditional subsequence. With this idea, we have to investigate the asymptotic properties of this new methodology.

Theorem: If we apply PARP strategy on two-extended envelope problem, let p^*

be the PARP decision, then

$$|N \cdot CL_N^{(2)}(\underline{a}, p^*) - R_N^{(2)}(p)| \leq A + B \cdot (\sqrt{n'_{00,N} \wedge n'_{01,N}} + \sqrt{n'_{10,N} \wedge n'_{11,N}})$$

where A and B are constants, $R_N^{(2)}(p) = n'_{00,N} \wedge n'_{01,N} + n'_{10,N} \wedge n'_{11,N}$ is the two-extended Bayes envelope.

Proof:

We partition the original sequence $\underline{a}$ into two subsets:

$$A_N : \{(a_i, a_j)| \quad a_i = 0, \quad , i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N\}$$

$$B_N : \{(a_i, a_j)| \quad a_i = 1, \quad , i = 1, 2, \ldots, N-1, \quad j = 1, 2, \ldots, N\}$$

Furthermore, we take out the second coordinates of each subsets to form two new subsequences $\widetilde{A}_N$ and $\widetilde{B}_N$. Thus, applying PARP strategy on original sequence conditional on the previous stage, is equivalent to using PARP strategy on $\widetilde{A}_N$ and $\widetilde{B}_N$.

By Gilliland and Jung (2006), when we use PARP strategy on $\widetilde{A}_N$ to predict a_{N+1}, we will have two constant A1 and B1 such that

$$CL_{n1}(\widetilde{A}, p^*) \leq \frac{1}{n1}(A1 + B1 \cdot \sqrt{n1 \cdot g_{n1}^N \wedge (1 - g_{n1}^N)})$$

where $CL_{n1}(\widetilde{A}, p^*)$ is the Cesàro loss on the random past strategy for sequence $\widetilde{A}_N$, $g_{n1}^N = \frac{n'_{01,N}}{n'_{01,N} + n'_{00,N}}$,

$$n'_{01,N} = number\ of\ (0, 1)\ in\ \widetilde{A}_N.$$

$$n'_{00,N} = number\ of\ (0, 0)\ in\ \widetilde{A}_N,$$

and $n1 = (n'_{01,N} + n'_{00,N})$, i.e. the number of pairs in $\widetilde{A}_N$.

Similarly, if $a_N = 1$, i.e. our Bayes response would be based on the set B_N instead of A_N, and when we use the random past strategy on the sequence $\widetilde{B}_N$. We have A2, B2 such that

$$CL_{n2}(\widetilde{B},p^*) \leq \frac{1}{n2}(A2 + B2 \cdot \sqrt{n2 \cdot g_{n2}^N \wedge (1 - g_{n2}^N)})$$

where $g_{n2}^N = \frac{n'_{11,N}}{n'_{10,N}+n'_{11,N}}$,

$$n'_{10,N} = number\ of\ (1,0)\ in\ \widetilde{B}_N,$$

$$n'_{11,N} = number\ of\ (1,1)\ in\ \widetilde{B}_N,$$

and $n2 = (n'_{11,N} + n'_{10,N})$.

Since $CL_{n1}(\widetilde{A},p^*)$ is the Cesaro loss with condition $a_N = 0$, and $CL_{n2}(\widetilde{B},p^*)$ is the Cesaro loss with condition $a_N = 1$, the total Cesaro loss should be:

$$CL_N(\underline{a},p^*) = P(a_N = 0) \cdot CL_{n1}(\widetilde{A},p^*) + P(a_N = 1) \cdot CL_{n2}(\widetilde{B},p^*)$$

By empirical distribution, we use $\frac{(n'_{00,N}+n'_{01,N})}{N-1}$ to estimate $P(a_N = 0)$, and $\frac{(n'_{01,N}+n'_{11,N})}{N-1}$ to estimate $P(a_N = 1)$.

Thus,

$$N \cdot CL_N(\underline{a},p^*) = (n'_{01,N} + n'_{00,N}) \cdot CL_{n1}(\widetilde{A},p^*) + (n'_{10,N} + n'_{11,N}) \cdot CL_{n2}(\widetilde{B},p^*)$$

Since the extended envelope $R_N^{(2)}(p) = n'_{00.N} \wedge n'_{01.N} + n'_{10.N} \wedge n'_{11.N}$. We have

$$\begin{aligned}
|N \cdot CL_N(\underline{a}, p^*) - R_N^{(2)}(p)| &= |(n'_{01.N} + n'_{00.N}) \cdot CL_{n1}(\widetilde{A}, p^*) - n'_{00.N} \wedge n'_{01.N} \\
&+ (n'_{10.N} + n'_{11.N}) \cdot CL_{n2}(\widetilde{B}, p^* - n'_{10.N} \wedge n'_{11.N}| \\
&= |(n'_{01.N} + n'_{00.N}) \cdot (CL_{n1}(\widetilde{A}, p^*) - g_{n1}^N \wedge (1 - g_{n1}^N)) \\
&+ (n'_{10.N} + n'_{11.N}) \cdot (CL_{n2}(\widetilde{B}, p^* - g_{n2}^N \wedge (1 - g_{n2}^N))| \\
&\leq (n'_{01.N} + n'_{00.N}) \cdot \frac{1}{n1}(A1 + B1 \cdot \sqrt{n1 \cdot g_{n1}^N \wedge (1 - g_{n1}^N)}) \\
&+ (n'_{10.N} + n'_{11.N}) \cdot \frac{1}{n2}(A2 + B2 \cdot \sqrt{n2 \cdot g_{n2}^N \wedge (1 - g_{n2}^N)}) \\
&\leq (A1 + B1 \cdot \sqrt{n'_{00.N} \wedge n'_{01.N}}) \\
&+ (A2 + B2 \cdot \sqrt{n'_{10.N} \wedge n'_{11.N}})
\end{aligned}$$

Let $A = A1 + A2$, and $B = \max\{B1, B2\}$, then

$$|N \cdot CL_N(\underline{a}, p^*) - R_N^{(2)}(p)| \leq A + B \cdot (\sqrt{n'_{00.N} \wedge n'_{01.N}} + \sqrt{n'_{10.N} \wedge n'_{11.N}})$$

for all Player I move sequence $\underline{a}$ and all N ($N > 1$).

Proof is done.□

Comments 3.6.1: The theorem above shows that $|CL_N(\underline{a}, p^*) - R_N^{(2)}(p)/N|$ has a uniform bound which is $O(N^{-1/2})$ in Player I sequence of move,i.e., the Cesaro loss of PARP strategy in two-extended envelope problem converges to the Bayes envelope with convergence rate $O(N^{-1/2})$.

Example 3.6.1 Suppose Player I is playing 0 or 1 for each time, we suspect that Player I's move at each stage is effected by his move on last stage. Therefore, it is reasonable for us to form this forecasting problem as a two-extended envelope problem.

Table 3.2: Two-extended envelope problem

stage k	a_k	p_k	$CL_k^{(2)}(\underline{a}, p)$	p_k^*	$CL_k^{(2)}(\underline{a}, p^*)$
1	1	0.5	0.5	0.5	0.5
2	1	0.5	0.5	0.5	0.5
3	0	1	0.667	1	0.667
4	1	0.5	0.625	0.5	0.625
5	0	0.5	0.6	0.5	0.6
6	0	1	0.667	1	0.667
7	1	0.5	0.643	0.5	0.643
8	0	0.3	0.604	0.26	0.595
9	1	0.67	0.574	0.74	0.557
10	1	0.25	0.592	0.16	0.586
11	1	0.4	0.592	0.32	0.595
12	0	0.5	0.585	0.5	0.587
13	1	0.75	0.559	0.84	0.553

We observed Player I 's moves through 13 stages:

$$\underline{a}_N : 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1$$

.

We apply both PAP and PARP strategy after each stage to make the decision for the next step. After the 13^{th} stage, we collect all the results in the table 3.1 showing where Player I move a_k, the PAP decision from Player I's past move is p, the PARP decision is p^*, and the expected Cesaro losses for both PAP and PARP strategy are also listed in this table.

Chapter 4

Discovering Hannan

This chapter gives the results of a search of the literature that had the goal of finding published results that were in Hannan (1957) and not recognized as being there. We find several examples. To some extent, the cryptic style and notations of Hannan (1957) explain the failure of other researchers to fully exploit the Hannan work and to recognize the specific theorems that he proved. Motivation for this search is provided in part by Gina Kolta's New York Times article *Pity the Scientist Who Discovers the Discovered*. February 5, 2006.

In Section 4.1, we discuss some of the Kolta (2006) article and mention that Chen (1997) established the direct connection of the Foster and Vohra (1993) result to Hannan (1957). In Section 4.2 we show that the main result in Feder, Merhav and Gutman (1992) is contained in Hannan (1957). In Section 4.3 we consider for the first and only time the finite horizon version of repeated play (see Section 1.2) and connect results on minmax regret found in Cover (1967), Chung (1994) and Cesa-Bianchi and Lugosi (1999) to Hannan (1957).

4.1 Foster and Vohra: Selecting Forecasters

Kolta's lead paragraph mentions the Foster and Vohra (1993) paper "A Randomized Rule for Selecting Forecasters." The strategy proposed in the Foster and Vohra paper has the structure of a Hannan strategy, much in appearance like those covered by his Theorems 3, 4 and 6. (The structure of Hannan-type strategies was explained in Chapter 2, Section 3.)

Hannan's theorems claimed and proved the conclusions for strategies built on being Bayes versus random perturbations of the multinomial empirical counts $(t-1)G_{t-1}$ of Player I's pure moves in repeated play of a game where Player I has m possible pure moves. Chen (1997) reexamined the Hannan theorems and proofs and showed that the proofs actually cover the case where the empirical distributions G_{t-1} are replaced by the empirical distributions of randomization distributions taking values in the probability simplex in R^m. A randomization distribution x is a probability distribution over the m pure moves and $(t-1)G_{t-1}$ is replaced by $X_{t-1} = x_1+x_2+\cdots+x_{t-1}$. Then Chen (1997, Section 4.3) shows how the Foster and Vohra strategy is a Hannan strategy so that bounds on its regret and asymptotics are a direct consequence of her reinterpretation of the Hannan theorems. Following Chen's work, Gilliland and Hannan (1999, 2008) improved on Chen, mainly through the demonstration of good bounds and Hannan consistency for strategies in the repeated play of the dual of the S-game. This component easily subsumes the expert selection problem considered by Foster and Vohra without a weakening of bounds (larger constants) that is inherent in Chen's approach.

Vohra is quoted in the following paragraph from Kolta (2006) in which Hannan's name is misspelled:

> In 1957, for example, a statistician named James Hanna called his theorem Bayesian Regret. He had been preceded by David Blackwell, also a statistician, who called his theorem Controlled Random Walks. Other,

> later papers had titles like "On Pseudo Games", "How to Play an Unknown Game", "Universal Coding" and "Universal Portfolios" Dr. Vohra said, adding, "It's not obvious how you do a literature search for this result."

As mentioned previously, Hannan and Blackwell used different approaches in the construction of their strategies, and it is likely that neither one named his theorems. Moreover, the term "Bayesian Regret" was probably not ever used by Hannan. "Controlled Random Walks" is the title of a talk and a subsequent proceedings paper by Blackwell (1956) that give a general result that can be applied to produce Hannan-consistent strategies. "On Pseudo Games", "How to Play an Unknown Game", "Universal Coding" and "Universal Portfolios" denote different general but related topics. The quoted paragraph might leave the false impression that exactly one result or theorem has been given the different names.

4.2 Feder, Merhav and Gutman:Universal Prediction

Feder, Merhav and Gutman (1992) considered the problem of predicting the next stage of an individual binary sequence using finite memory. And in the section III of this paper, they gave out the definition of this predictor, which is called "*S-State Universal Sequential Predictor*," in the following way:

$$\widehat{X}_{t+1} = \begin{cases} \text{“0”}, & \text{with probability } \phi(\widehat{p}_t(0)), \\ \text{“1”}, & \text{with probability } \phi(\widehat{p}_t(1)) = 1 - \phi(\widehat{p}_t(0)). \end{cases}$$

where $\phi(\cdot)$ is given by

$$\phi(\alpha) = \begin{cases} 0, & 0 \le \alpha < \frac{1}{2} - \epsilon, \\ \frac{1}{2\epsilon}[\alpha - \frac{1}{2}] + \frac{1}{2}, & \frac{1}{2} - \epsilon \le \alpha \le \frac{1}{2} + \epsilon, \\ 1, & \frac{1}{2} + \epsilon < \alpha \le 1. \end{cases}$$

They allow ϵ to depend on t, i.e., $\epsilon = \epsilon_t$, and use $\widehat{\pi}_1(x_1^n)$ to represent the expected fraction errors made by this scheme over the sequence x^n where the expectation is with respect to the randomization in the definition of $\widehat{x}_t$. $\widehat{\pi}_1(x_1^n)$ is our Cesaro loss.

In the major theorem, they proved that, for $\epsilon = \epsilon_t = 1/(2\sqrt{t+2})$,

$$\widehat{\pi}_1(x_1^n) \le \pi_1(x_1^n) + \delta_1(n).$$

where $\delta_1(n) = O(1/\sqrt{n})$, $\pi_1(x_1^n) = 1/n \cdot \min\{N_n(0), N_n(1)\}$, and $N_n(0)$ and $N_n(1)$ are count of zeros and ones, respectively, along the sequence $\underline{X}$. Thus in their major theorem, they showed that the expected error converges to the simple Bayes envelope with convergence rate $O(1/\sqrt{n})$.

However, we can show that their strategy is equivalent to a special case of Hannan's strategy. Recall a Hannan-type strategy in matching pennies problem, which is discussed in chapter 2 with Player II's predictor:

$$\sigma(1 - g_{t-1}, g_t) = \begin{cases} 0, & Z_2 - Z_1 \le \frac{1-2g_{t-1}}{h_t}, \\ 1, & Z_2 - Z_1 > \frac{1-2g_{t-1}}{h_t}. \end{cases}$$

where g_{t-1} is the proportion of 1's in Player I's play from stage 1 to stage t-1. In another words, Player II's predictor is

$$\widehat{X}_{t+1} = \begin{cases} \text{"0"}, & \text{with probability } P(Z_2 - Z_1 \le \frac{1-2g_{t-1}}{h_t})), \\ \text{"1"}, & \text{with probability } P(Z_2 - Z_1 > \frac{1-2g_{t-1}}{h_t}). \end{cases}$$

Now we can show that Feder, Merhav and Gutman's strategy's probability of playing 1, $\phi(\widehat{p}_t(1)) = 1-\phi(\widehat{p}_t(0))$, is the same as probability of playing 1 in Hannan's strategy $P(Z_2 - Z_1 > \frac{1-2g_{t-1}}{h_t})$ with a specified distribution for $Z = (Z_1, Z_2)$.

Take Z_1 to be a random variable from uniform distribution on $[0,1]$, let Z_2 be degenerate at $1/2$, and let $\alpha = \widehat{p}_t(1)$ the proportion of 1 in empirical distribution of Player I moves from stage 1 to t-1. Then, we have

$$\phi(\alpha) = P(Z_1 < \frac{1}{2} + \frac{1}{2\varepsilon}(\alpha - \frac{1}{2})) = \begin{cases} 0, & 0 \le \alpha < \frac{1}{2} - \epsilon, \\ \frac{1}{2\epsilon}[\alpha - \frac{1}{2}] + \frac{1}{2}, & \frac{1}{2} - \epsilon \le \alpha \le \frac{1}{2} + \epsilon, \\ 1, & \frac{1}{2} + \epsilon < \alpha \le 1. \end{cases}$$

and letting $\alpha = g_{t-1}$, $\frac{1}{2\varepsilon} = \frac{2}{h_t}$, i.e. $h_t = 4\varepsilon$, we have

$$\begin{aligned} P(Z_1 < \frac{1}{2} + \frac{1}{2\varepsilon}(\alpha - \frac{1}{2})) &= P(Z_1 - \frac{1}{2} < \frac{2}{h_t}(\alpha - \frac{1}{2})) \\ &= P(Z_1 - \frac{1}{2} < \frac{1}{h_t}(2g_{t-1} - 1)) \\ &= P(Z_1 - Z_2 < \frac{1}{h_t}(2g_{t-1} - 1) \quad where \quad Z_2 = \frac{1}{2} \\ &= P(Z_2 - Z_1 > \frac{1}{h_t}(1 - 2g_{t-1}) \quad where \quad Z_2 = \frac{1}{2} \end{aligned}$$

This shows that probability of playing 1 in the Feder, Merhav and Gutman's strategy is the same as the probability of playing 1 in Hannan's strategy where $Z_1 \sim Uniform[0,1]$ and Z_2 is degenerated at $1/2$. Furthermore, since probability of playing 0, Prob(playing 0)=1-Prob(playing 1), we also have $\phi(\widehat{p}_t(0)) = P(Z_2 - Z_1 \le \frac{1}{h_t}(1-2g_{t-1})$, which is probability of playing 0 in Hannan's strategy.

Therefore, the Feder, Merhav and Gutman(1992) strategy is the same as a Hannan's strategy (1957).

4.3 Minimax Regret

In this section and only in this section we consider repeated play with a finite horizon N known to Player II in advance of the repeated play. Player II's concern is with regret at stage N and II's sequence of recursive functions p_N can depend on N. There is a considerable literature on predicting individual sequences in the context of finite horizon repeated play and minimax regret strategies. A minimax regret strategy can be constructed from backward reduction starting at N. Since

$$ND_N(\underline{a}, \underline{p}) = \sum_{t=1}^{N-1} L(a_t, p_t(\underline{a}_{t-1}) + \{L(a_N, p_N(\underline{a}_{N-1})) - NR(G_N)\},$$

$p_N(\underline{a}_{N-1})$ can be chosen to minimize

$$\max\{L(a_N, p_N(\underline{a}_{N-1})) - NR(G_N) | a_N \in A\}.$$

Then the resulting max is added to $L(a_{N-1}, p_{N-1}(\underline{a}_{N-2}))$ and $p_{N-1}(\underline{a}_{N-2})$ chosen to minimize the maximum possible total over all a_{N-1}. Continuing in this way, a minimax regret strategy $\underline{p}^{mM} = (p_1^{mM}, p_2^{mM}, \ldots, p_N^{mM})$ is determined. The minimax regret strategy results in constant regret across all sequences $\underline{a}$ and that common value is denoted by D_N^{mM} and is called minimax regret.

Hannan (1957, Section 4) constructed minimax strategies for the finite horizon repeated play of a game where Player I has m moves. He illustrates his results with an example of "matching m-sided pennies." The translation of his minimax regret to predicting a binary sequence requires taking $m = 2$, his $p_1 = p_2 = \frac{1}{2}$ and noting that his loss of 2 comes from a match not a mismatch. Then the translation of Hannan (1957, (8), p. 115) gives,

$$ND_N^{mM} = \frac{N}{2} - E\min\{Y, N - Y\} \tag{4.1}$$

where $Y \sim Binomial(N, \frac{1}{2})$.

Proposition 4.3.1. *For the repeated play of matching pennies with a finite horizon,*

$$D^{mM}_{2k+1} = D^{mM}_{2k} = \frac{1}{2}\frac{(2k)!}{4^k k!k!} = \frac{1}{2}P(Y^* = k), \tag{4.2}$$

where $Y^* \sim Binomial(2k, 1/2)$. Furthermore, D^{mM}_{2k} is a decreasing sequence with limit 0, $\sqrt{2k}D^{mM}_{2k} \sim 1/\sqrt{2\pi}$, and

$$\sqrt{2k}\cdot D^{mM}_{2k} = \frac{1}{\sqrt{2\pi}} + O(\frac{1}{k}). \tag{4.3}$$

Proof: Hannan (1957, Theorem 2, p. 111) develops an asymptotic lower bound for D^{mM}_N in the general case that he considers, that is, in the repeated play of an $m \times n$ game. The constant h in his Theorem 2 bound can be shown to be 1 for matching two-sided pennies with the loss matrix we use, so that Theorem 2 implies that $\liminf_N N^{1/2} D^{mM}_N \geq (2\pi)^{1/2}$.

We develop an expression for D^{mM}_N. Since $\min\{Y, N-Y\} = N - \max\{Y, N-Y\}$, (4.1) can be expressed by:

$$D^{mM}_N = \frac{1}{N}\cdot E\max\{Y, N-Y\} - \frac{1}{2}. \tag{4.4}$$

A calculation using the symmetry of the distribution $B(N, \frac{1}{2})$ about $N/2$ shows that with $N = 2k+1$,

$$\max\{Y, 2k+1-Y\} = \begin{cases} 2k+1, & \text{if } Y = 0, \quad 2k+1 \\ 2k, & \text{if } Y = 1, \quad 2k, \\ 2k-1, & \text{if } Y = 2, \quad 2k-1, \\ \dots, & \dots \\ \dots & \dots \\ k+1, & \text{if } Y = k, \quad k+1. \end{cases}$$

Therefore,

$$
\begin{aligned}
E\max\{Y, 2k+1-Y\} &= 2\cdot\frac{1}{2^{2k+1}}\cdot\sum_{j=0}^{k}\binom{2k+1}{j}\cdot(2k+1-j)\\
&= 2\cdot\frac{1}{2^{2k+1}}\cdot\sum_{j=0}^{k}\frac{(2k+1)!}{j!(2k-j)!}\cdot(2k+1-j)\\
&= \frac{2k+1}{2^{2k}}\cdot\sum_{j=0}^{k}\frac{(2k)!}{j!(2k-j)!}\\
&= (2k+1)\cdot(\frac{1}{2}+\frac{1}{2}P(Y^*=k)).
\end{aligned}
$$

where $Y^* \sim Binomial(2k, \frac{1}{2})$, i.e.

$$E\max\{Y, 2k+1-Y\} = (2k+1)\cdot\{\frac{1}{2}+\frac{1}{2}\frac{(2k)!}{4^k k!k!}\},$$

And with $N = 2k$, similarly we have

$$E\max\{Y, 2k-Y\} = 2k\cdot\{\frac{1}{2}+\frac{1}{2}\frac{(2k)!}{4^k k!k!}\}.$$

It follows that minimax regret is given by

$$D^{mM}_{2k+1} = D^{mM}_{2k} = \frac{1}{2}\frac{(2k)!}{4^k k!k!} = \frac{1}{2}P(Y^*=k), \tag{4.5}$$

where $Y^* \sim Binomial(2k, \frac{1}{2})$, $k = 1, 2, 3, \ldots$.

Furthermore, since for all $k = 1, 2, \ldots$, $D^{mM}_{2k} \geq 0$, and

$$\frac{D^{mM}_{2(k+1)}}{D^{mM}_{2k}} = \frac{\frac{1}{2}\cdot\frac{2(k+1)!}{4^k(k+1)!(k+1)!}}{\frac{1}{2}\cdot\frac{2(k)!}{4^k(k)!(k)!}} = \frac{2k+1}{2k+2} < 1$$

This shows $D^{mM}_{2(k+1)}$ is smaller than D^{mM}_{2k}, i.e., D^{mM}_{2k} is a decreasing sequence.

From the Stirling Formula it follows that

$$D_N^{mM} \sim \frac{1}{\sqrt{2\pi N}}. \tag{4.6}$$

We deduce (4.3) from the well-known result on the rate of convergence of the Walli's product sequence to $\pi/2$, i.e.,

$$W_k = \frac{\pi}{2} - \frac{\pi}{8k} + o(\frac{1}{k}), \ \ as\ k \to \infty$$

where $W_k = \frac{\pi_k}{2}$, $\pi_k = \frac{4k+2}{\alpha_k^2}$, $\alpha_k = \frac{(2k+1)!}{4^k k! k!}$. (For example, see Hirschhorn (2005).) The sequence $\{\alpha_k\}$ was encountered at the end of section 2.1 and was studied extensively in Frame and Gilliland (1985) where a continued fraction representation is found. Note that

$$\begin{aligned} \sqrt{2k} \cdot D_{2k}^{mM} &= \frac{1}{2} \cdot \sqrt{\frac{2k}{2k+1}} \cdot \sqrt{\frac{2}{\pi_k}} \\ &= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{2k}{2k+1}} \cdot \frac{\sqrt{\pi}}{\sqrt{\pi_k}}. \end{aligned}$$

Then $\sqrt{2\pi} \cdot \sqrt{2k} \cdot D_{2k}^{mM} - 1 = \sqrt{\frac{2k}{2k+1}} \cdot \frac{\sqrt{\pi}}{\sqrt{\pi_k}} - 1$.

And we claim that

$$\sqrt{2k} \cdot D_{2k}^{mM} - \frac{1}{\sqrt{2\pi}} \leq \frac{A}{\sqrt{2\pi}}, \quad where \quad A = O(\frac{1}{k}).$$

Since

$$\sqrt{\frac{2k}{2k+1}} \cdot \frac{\sqrt{\pi}}{\sqrt{\pi_k}} - 1 \leq \frac{\sqrt{\pi}}{\sqrt{\pi_k}} - 1$$

Proof is done if we show that $\frac{\sqrt{\pi}}{\sqrt{\pi_k}} = 1 + O(\frac{1}{k})$, i.e. $\frac{\pi}{\pi_k} = 1 + O(\frac{1}{k})$.

Since Wallis' sequence,

$$W_k = \frac{\pi}{2} - \frac{\pi}{8k} + o(\frac{1}{k}) \quad as \quad k \sim \infty$$

and $W_k = \frac{\pi_k}{2}$, then we have

$$\pi_k = \pi - \frac{\pi}{4k} + o(\frac{1}{k})$$

i.e.,

$$\frac{\pi_k}{\pi} = 1 - \frac{1}{4k} + o(\frac{1}{k}).$$

In another form,

$$\begin{aligned}
\frac{\pi}{\pi_k} &= \frac{1}{1 - 1/(4k) + o(1/k)} \\
&= 1 + \frac{1 - o(1)}{4k - 1 + o(1)} \\
&= 1 + O(\frac{1}{k})
\end{aligned}$$

Proof is done. □

Here is a table of initial values of minimax regret. Recall that $D^{mM}_{2h+1} = D^{mM}_{2h}$, and note that $1/\sqrt{2\pi} = 0.39894$.

We have given a simple expression for the minimax regret. There is interest in the minimax regret strategy for Player II, that is, the strategy $\underline{p}^{mM}$, that minimizes the maximum regret for the finite horizon N.

For the simple case that we are considering, the strategy can be deduced by specializing Hannan (1957, (4), p. 114) to the binary case $m = 2$ with $p_1 = p_2 = \frac{1}{2}$. Also, the Hannan loss is 2 for a match so that his Bayes envelope and procedures must be reinterpreted. His y_1^j is our p_j^{mM}. Hannan (1957, (4), p. 114) written for the

Table 4.1: Convergence of minimax regret

N	D_N^{mM}	$\sqrt{N}D_N^{mM}$
1	$\frac{1}{2}$	0.5
2	$\frac{1}{4}$	0.35355
4	$\frac{3}{16}$	0.375
6	$\frac{5}{32}$	0.38273
8	$\frac{35}{256}$	0.38670
10	$\frac{63}{512}$	0.38911
12	$\frac{462}{4096}$	0.39073
14	$\frac{1716}{16384}$	0.39189
16	$\frac{6435}{65536}$	0.39276
∞	0	0.39894

binary sequence case that we are considering is

$$p_j^{mM}(\underline{a}_{j-1}) = \frac{1}{2}[1 + ER((\sum_{t=1}^{j-1} a_t + Y)/N - ER((\sum_{t=1}^{j-1} a_t + Y^* + Y)/N))] \qquad (4.7)$$

where $Y \sim Binomial(N-j, \frac{1}{2})$, and $Y^* \sim Binomial(1, \frac{1}{2})$ are independent. Expectation over Y^* results in

$$p_j^{mM}(\underline{a}_{j-1}) = \frac{1}{2}[1 + ER((\sum_{t=1}^{j-1} a_t + Y)/N) - R((\sum_{t=1}^{j-1} a_t + 1 + Y)/N))] \qquad (4.8)$$

Since $R(\pi) = \min\{\pi, 1-\pi\}$, we see that

$$R((\sum_{t=1}^{j-1} a_t + Y)/N) - R((\sum_{t=1}^{j-1} a_t + 1 + Y)/N) = \begin{cases} 1, & \text{if } Y \geq N/2 - \sum_{t=1}^{j-1} a_t, \\ 0, & \text{if } Y \geq N/2 - \sum_{t=1}^{j-1} a_t - \frac{1}{2}, \\ -1, & \text{if } Y \leq N/2 - \sum_{t=1}^{j-1} a_t - 1, \end{cases}$$

It follows that

$$p_j^{mM}(\underline{a}_{j-1}) = \frac{1}{2} + \frac{1}{2}[P(Y \geq N/2 - \sum_{t=1}^{j-1} a_t) - P(Y \leq N/2 - 1 - \sum_{t=1}^{j-1} a_t) \tag{4.9}$$

where $Y \sim Binomial(N-j, \frac{1}{2})$ if $j = 1, 2, \ldots, N-1$ and Y is taken as 0 if $j = N$. Note that $p_N^{mM}(\underline{a}_{N-1})$ is simply the PAP strategy, that is, at the last stage N, the Player II minimax regret strategy plays the Player I majority choice in the first $N-1$ stages.

In the minimax strategy, is the probability used for playing 1 is larger than the probability in Hannan's strategy and in the PARP strategy. Figure 4.1 and 4.2 show these three probabilities for N=5 and N=10.

Table 4.2: Player I's play sequence with N=5.

stage	1	2	3	4	5
$\underline{a}$	1	1	1	1	0

Figure 4.1: Hannan, PARP and Minimax probability for N=5

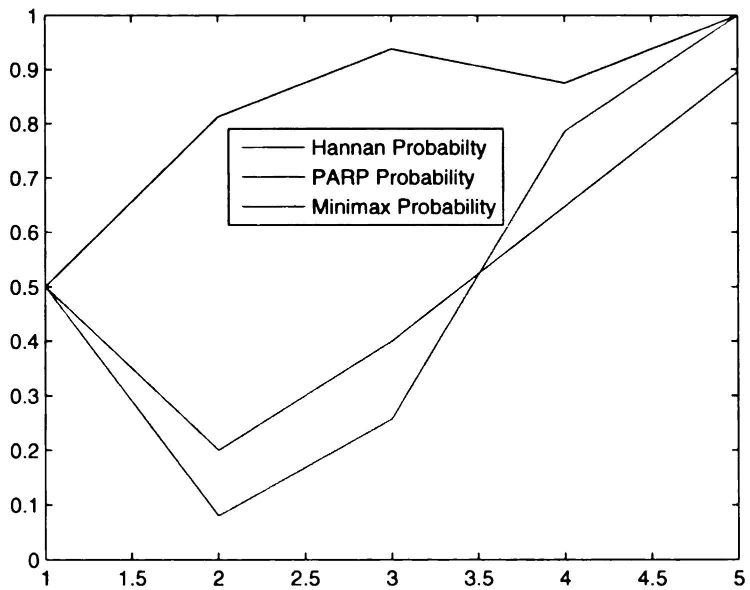
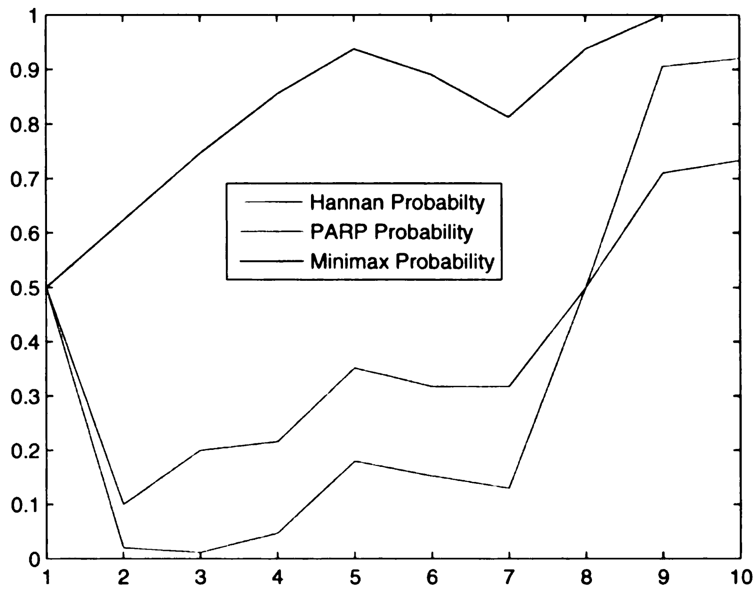


Figure 4.2: Hannan, PARP and Minimax probability for N=10



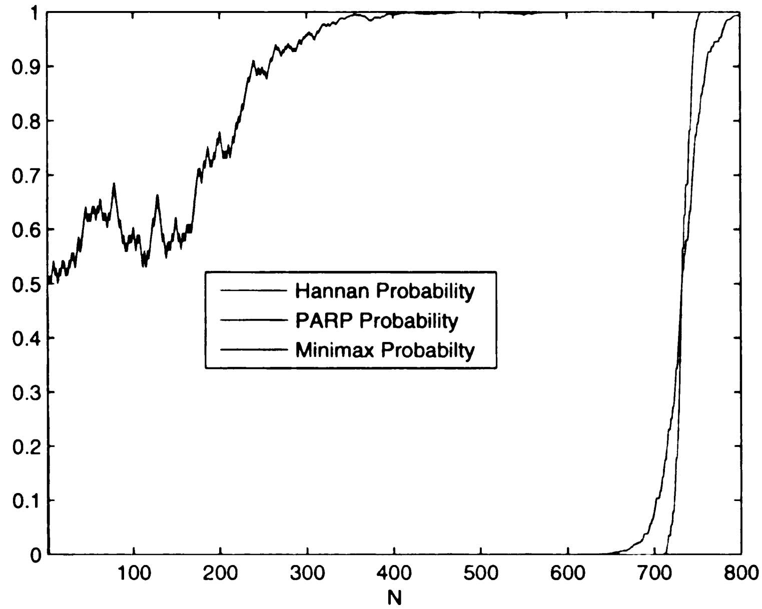
Asymptotically Hannan's probability is smoother than probability in PARP strategy, and Minimax probability is bigger than both of them as shown by Figure 4.3. We generate a binary sequence with N=100 from Bernoulli(1,1/2) as Player I's play sequence and the asymptotic behavior of Hannan, PARP and Minimax probability are illustrated by figure 4.3.

Cover (1967) develops many interesting results concerning strategies for predicting binary sequences. Cover measures the performance of strategies by gain through the number of matches (not by loss through number of misses) so his regret is the negative of the one we consider. The display following his (4.13) is of the minimax regret value (4.4) with the asymptotic result $D_N^{mM} \sim 1/\sqrt{2\pi N}$ noted. Cover (1967, (4.11)) gives

Table 4.3: Player I's play sequence with N=10.

stage	1	2	3	4	5	6	7	8	9	10
$\underline{a}$	1	1	1	1	0	0	1	1	0	1

Figure 4.3: Hannan, PARP and Minimax probability for N=800



the minimax regret strategy, which we will show is the same as derived by Hannan (1957).

His predictor achieving $\widehat{s}$ is:

$$\widehat{p}_k(i) = \frac{1}{2} + N \cdot (\frac{1}{2})^{N-k+1} \sum_{j=1}^{N-k} (\widehat{s}(\frac{i+j+1}{N} - \widehat{s}(\frac{i+j}{N})) \binom{N-k}{j}.$$

And it follows that Cover (1967) (4.13), $\widehat{s}$ can be specified as the simple Bayes envelope, i.e.

$$\widehat{s}(\eta) = \max\{\eta, 1-\eta\}, \quad 0 < \eta < 1.$$

Let $i = k-1$, then

$$\widehat{p}_k(k-1) = \frac{1}{2} + N \cdot (\frac{1}{2})^{N-k+1} \sum_{j=1}^{N-k} (\widehat{s}(\frac{k+j}{N}) - \widehat{s}(\frac{k+j-1}{N})) \binom{N-k}{j}.$$

where

$$\widehat{s}(\frac{k+j}{N}) = \begin{cases} \frac{k+j}{N}, & \text{if } \frac{N}{2} - k \le j \le N-k \\ 1 - \frac{k+j}{N}, & \text{if } 0 \le j \le \frac{N}{2} - k, \end{cases}$$

$$\widehat{s}(\frac{k+j-1}{N}) = \begin{cases} \frac{k+j-1}{N}, & \text{if } \frac{N+1}{2} - k \le j \le N-k \\ 1 - \frac{k+j-1}{N}, & \text{if } 0 \le j \le \frac{N+1}{2} - k, \end{cases}$$

Since j is an integer, based on the fact that

$$\widehat{s}(\frac{k+j}{N}) - \widehat{s}(\frac{k+j-1}{N}) = \begin{cases} 1, & \text{if } \frac{N+1}{2} - k \le j \le N-k \\ -1, & \text{if } 0 \le j \le \frac{N}{2} - k, \end{cases}$$

we have

$$\begin{aligned} \widehat{p}_k(k-1) &= \frac{1}{2} + \frac{1}{2} \cdot [\sum_{j=\frac{N+1}{2}-k}^{N-k} \binom{N-k}{j} \cdot (\frac{1}{2})^{N-k} - \sum_{j=0}^{\frac{N+1}{2}-k} \binom{N-k}{j} \cdot (\frac{1}{2})^{N-k}] \\ &= \frac{1}{2} + \frac{1}{2} \cdot [P(Y \ge \frac{N+1}{2} - k) - P(Y \le \frac{N}{2} - k)] \end{aligned}$$

where $Y \sim Binomial(N-k, 1/2)$.

This shows that Cover's strategy $\widehat{p}_k(i)$ is the same as Hannan's strategy (4.9), if the envelope is the simple Bayes envelope.

Chung (1994) and Cesa-Bianchi and Lugosi (1999) studied sequential randomized prediction for an arbitrary binary sequence. In latter with n replaced by N, the prediction at each time $t = 1, 2, 3, \ldots, N$, it is given by

$$p_t(\underline{y}^{t-1}) = \frac{1}{2} + \frac{1}{2} E[\inf_{\mathfrak{F}} L_F(\underline{y}^{t-1} 0 \underline{Y}^{N-t}) - \inf_{\mathfrak{F}} L_F(\underline{y}^{t-1} 1 \underline{Y}^{N-t})] \tag{4.10}$$

$\mathfrak{F}$ is the set of experts, and $\underline{y}^{t-1} 0 \underline{Y}^{N-t}$ and $\underline{y}^{t-1} 1 \underline{Y}^{N-t}$ represent the following sequences respectively:

$$y_1, y_2, \ldots, y_{t-1}, 0, Y_t, \ldots, Y_N$$

and

$$y_1, y_2, \ldots, y_{t-1}, 1, Y_t, \ldots, Y_N$$

where the Y_j are independent random variables which follow $Bernoulli(1/2)$.

In Chung (1994) with $T = N$, it was expressed as

$$p_t(\underline{y}^{t-1}) = \frac{1}{2} + \frac{1}{2}E[\max_i \Phi_{i,N}(\underline{y}^{t-1}, 0, Y_{t+1}^N) - \max_i \Phi_{i,N}(\underline{y}^{t-1}, 1, Y_{t+1}^N)] \tag{4.11}$$

where $\max_i \Phi_{i,N}(\underline{a})$ gives the maximum pay-off for sequence $\underline{s}$ among i experts given the total stages N.

Since these two predictions are essentially the same, we take Cesa-Bianchi and Lugosi (1999) strategy as an example and specialize it to the simple Bayes envelope. Then

$$\inf_{\mathfrak{F}} L_F(y^{t-1}0Y^{N-t}) - \inf_{\mathfrak{F}} L_F(y^{t-1}1Y^{N-t}) = \begin{cases} 1, & \text{if } S \geq N/2 - (t-1)\cdot g_{t-1}, \\ 0, & \text{if } S \geq N/2 - (t-1)\cdot g_{t-1}\text{-}1/2, \\ -1, & \text{if } S \geq N/2 - (t-1)\cdot g_{t-1} - 1, \end{cases}$$

where $S = \sum Y_i \sim Binomial(N-t, 1/2)$, since $Y_i \sim Bernoulli(1/2)$, $i = t+1, \ldots, N$, and g_{t-1} is the proportion of 1 from stage 1 to stage t-1.

In this way, (4.10) is converted to

$$p_t^{mM}(\underline{y}^{t-1}) = \frac{1}{2} + \frac{1}{2}[P(S \geq N/2 - (t-1)\cdot g_{t-1}) - P(S \leq N/2 - 1 - (t-1)\cdot g_{t-1}) \tag{4.12}$$

where $S \sim Binomial(N-t, \frac{1}{2})$.

This shows that if we specify the envelope to be the simple Bayes envelope, minimax regret results in both Chung (1994) and Cesa-Bianchi and Lugosi (1999) are the same as found in (4.9), Hannan (1957).

Chapter 5

Expert Selection Problem

5.1 Introduction and Review

Nowadays, all kinds of consulting services are booming, especially in financial services. There are many financial companies and agencies giving advice everyday to all kinds of investors. They are using different and complicated system or algorithms to analyze the financial market of different financial products and to forecast the market of these products. As experts with experience and knowledge in finance, each of them is trying to persuade the individual investors to take his/her advice. However, surrounded by so many experts' advice, as an investor, how can one make a decision? This is called *expert problem*.

Littlestone (1988) generalized the earlier researcher's idea to an arbitrary set of experts. However, in his strategy, randomness is not included in the forecasting process. His strategy concerns picking the the expert whose forecasting record is the best, as the best expert in the set of experts, and using this best expert's prediction as the final forecast. He showed that as long as there exist one expert whose forecasting is correct in all stages, the final decision maker will not make more than $log_2 N$ mistakes, where N is the total number stages.

To remove the restriction in Littlestone(1988), i.e., to consider the case that among all experts there is no expert always correct. Littlestone and Warmuth (1989) introduced the idea of weighted majority algorithm. In this strategy, they assign a weight to each expert. Once an expert makes a mistake in forecasting, he will receive a penalty: his new weight is old weight multiplied by k, $0 < k < 1$. i.e., to reduce the weight on his advice in our final decision. Furthermore, Warmuth and Haussler, et al (1993) considered the weighted majority strategy in the situation that each expert's forecaster is a probability distribution on set $\{0, 1\}$. In all these strategies, randomness is not involved in selecting the expert's actual forecast.

Hannan in 1957 first proposed the idea of bringing randomness into sequential forecasting problems. In Hannan's strategy, a random factor is added to the empirical distribution, and a predictor based on this adjusted empirical distribution is used as the forecaster for next stage of the play in a repeated game problem, as we have introduced in previous chapters. This idea can be introduced into expert selection problem.

Foster and Vorha (1993), proposed a randomized rule for selecting experts. They first proposed this expert actual selection problem instead of predicting a probability distribution of experts' set, or combining the experts' advices. However, the randomized strategy they proposed is equivalent to Hannan's strategy, which was proved by Chen (1997) and improved in Gilliland and Hannan (1999, 2008).

By using bootstrap sampling, we introduce the PARP strategy to the expert selection problem, especially the two experts selection problem, which will be discussed in section 5.3. Section 5.2 will introduce an example of usage of expert selection in the real world, a methodology called focus forecasting. And at the end of this chapter, a simulation example of using PARP strategy in financial forecasting is given.

5.2 Focus Forecasting

5.2.1 Introduction

In inventory management, forecasting is essential. As a new concept of forecasting, the term *focus forecast* was raised by Bernard Smith (1978). With around 30 years of usage so far, this method which is described as a heuristic methodology and it is used widely in industrial area. Over 800 companies in 47 countries worldwide are using Demand Solution which is designed around focus forecasting in their inventory management request. And this method is described to be a simple simulation approaches to optimization, to be more practical, more easy to understand and a simple system to work.

Focus forecasting constructs a pool of alternative decision rules for forecasting one stage ahead. At every stage, all the decision rules or models in the pool, are tested by the empirical data generated before this stage, and the rule with the smallest error in selected for the decision.

Therefore, focus forecasting simulates every time it forecasts. It is a dynamic simulation. It uses a computer to simulate every time, and compares the errors of all the rules, to pick one to use in the current forecast. Regardless the seasonal or trend type of time series data, focus forecasting itself just picks the one best strategy based on the empirical test against recent history data.

In inventory management, the traditional method is exponential smoothing, which is taught to almost every student in inventory management and is still the most widely used forecasting method in the world today. However, focus forecasting doesn't use the exponential smoothing to approximate moving average. The reason Smith states in Bernard Smith (1978):

> In those early computers, storing a twelve moth inventory history to calculate a moving average was expensive, inaccurate, and dangerous. So Bob

> Brown used exponential smoothing... Focus Forecasting doesn't use exponential smoothing to approximate moving average. Why? Well, computers today don't make mistakes. They are nearly 100 percent accurate...

However, Gardner and Anderson (2001), compared the focus forecasting and exponential smoothing showing that exponential smoothing is substantially more accurate than the Demand Solutions approach. Although there are some criticism on focus forecasting in academic field, we still can notice some interesting ideas in focus forecasting, which is Play Against Past strategy's idea.

5.2.2 Methodology

Focus forecasting constructs a pool of decision rules or strategies. Some of these rules are designed for recognizing trend, some of them are designed for recognizing seasonality.

For example, 'whatever the demand was in the past three months will probably be the demand in the next three months', this would be a rule to recognize trend instead of seasonality. While if a simple rule as 'whatever percentage increase or decrease we had over last year in the last three months will probably be the percentage increase or decrease over last year in the next three months', would be a rule of recognizing seasonality.

Gardner, Anderson-Flether and Wicks (2001) listed the seventeen decision rules included in Demand Solutions. And there rules are functions of the previous quarterly data:

1. Next quarter will equal last quarter.
2. Next quarter will equal last quarter plus a growth factor.
3. Next quarter will equal the same quarter a year ago.
4. Next quarter will equal the same quarter a year ago plus a growth factor.
5. Next quarter will equal the average of the last two quarters.
6. Next quarter will equal the average of the last two quarters plus a

growth factor.
7. Next quarter will equal the average of the last two quarters with the last quarter double weighted.
8. Next quarter will equal the last quarter plus the difference of the corresponding quarters last year.
9. Next quarter will equal the average of the last three quarters, with the last quarter double-weighted, and which seasonal adjustment.
10. Next quarter will equal the average of the same quarter in the last two years plus a growth factor.
11. Next quarter will equal the average of the last quarter of the current year plus the difference of the corresponding quarters from the last year plus the difference of the corresponding quarters from two years ago.
12. Next quarter will equal the average quarter of the last year.
13. Next quarter will equal the average quarter of the last year plus a growth factor.
14. Next quarter will equal the average quarter of the last two years.
15. Next quarter will equal the average quarter of the last two years with seasonal adjustment.
16. Next quarter will equal the average quarter of the last year plus the change from the average quarter two years ago.
17. Next quarter will equal the average quarter last year, plus the change from the average quarter two years ago, with seasonal adjustment.

Then, during the simulation procedure, an error of measurement for each strategy for each time will be computed, and the final forecast strategy is selected among these decision rules.

From these decision rules' definition, we can easily notice that although the final decision is selected among these rules, final decision is a function of the history data, i.e. past data, since all the decision rules are function of past data. In another words, focus forecasting is using Play Against Past strategy to make the decision.

From our discussion about PAP strategy and its failure, we could see that there are some situations in which focus forecasting fails.

5.3 Two experts selection problem

Suppose there are two experts who give out predictions against the market each day. Their errors probabilities are recorded at the end of the day:

$$Expert \quad 1: X_1, \quad X_2, \quad X_3, \dots \quad X_{n-1}, \quad X_n, \dots$$

$$Expert \quad 2: Y_1, \quad Y_2, \quad Y_3, \dots \quad Y_{n-1}, \quad Y_n, \dots$$

where we assume the errors are bounded. Without loss of generality we take $\{X_i\}$ and $\{Y_i\} \in [0, 1]$.

Selecting an expert for each stage is repeated play of the component game where
Player I selects a pure $a = (x, y) \in [0, 1]^2$
Player II selects a coordinate $b \in \{1, 2\}$
and the loss function for Player II is

$$L(a, b) = X[b = 1] + Y[b = 2].$$

If π is a probability distribution on $[0, 1]^2$, then the Bayes risk of b is

$$L(\pi, b) = \int L(a, b) d\pi(a) = E_\pi(X)[b = 1] + E_\pi(Y)[b = 2].$$

Bayes risk is any choice b to minimize thus.

$$b = \begin{cases} 1, & \text{if } E_\pi(X) < E_\pi(Y) \\ arbitrary, & \text{if } E_\pi(X) = E_\pi(Y) \\ 2, & \text{if } E_\pi(X) > E_\pi(Y) \end{cases}$$

and minimum Bayes risk is

$$R(\pi) = E_\pi(X) \wedge E_\pi(Y).$$

In repeated play, Player II uses $b_t(\underline{a}_{t-1})$ for stage t, $t = 2, 3, \ldots$ and the simple envelope is $\overline{X}_N \wedge \overline{Y}_N$ where $(\overline{X}_N, \overline{Y}_N)$ is the average of the (X_1, Y_1), $(X_2, Y_2), \ldots,$ (X_N, Y_N).

5.3.1 Reduction to Z-problem

Let $Z_i = X_i - Y_i$, then the Bayes envelope of two-experts selection problem $R_N = \overline{X_N} \wedge \overline{Y_N}$ is equivalent to $R_N = \overline{Z_N} \wedge 0 + \overline{Y_N}$. Thus, two experts' forecasting error sequences are equivalent to:

$$\underline{Z} : Z_1, \quad Z_2, \quad Z_3, \ldots \quad Z_{N-1}, \quad Z_N, \ldots$$

and the zero sequence since Y_i are fixed and given by the past history data. In this way, original two experts problem is transformed into a one-dimensional game.

Each term of this Z-sequence can be arbitrary number from -1 to 1, then by previous result,

$$N \cdot D_N(\underline{Z}, \underline{parp}) = \sum_{k=1}^{N} (P_k \cdot Z_k) - (\sum_{k=1}^{N} Z_k) \wedge 0.$$

where $\underline{P}$ is the probability sequence P_1, $P_2, \ldots,$ P_N. $P_k = Prob(\overline{Z}^*_{k-1} \leq 0)$, for $k = 2, 3, \ldots, N$, P_1 =arbitrary number from $[0, 1]$, and $\overline{Z}^*_{k-1}$ is the average of the random sample from $\{Z_1, Z_2, \ldots, Z_{k-1}\}$.

Let $0 \leq A_N \uparrow$ and $0 \leq B_N \uparrow$ be such that

$$-A_N \leq N \cdot D_N \leq B_N.$$

If we remove all the 0 terms from this Z-sequence to form a new sequence $\underline{\tilde{Z}}$ which is the subsequence of original Z-sequence, let $m_N \leq N$ be the number of the non-zero terms, also be the number of the terms in sequence $\underline{\tilde{Z}}$.

Then, consider P is the probability through the past summation, we have

$$-A_{m_N} \leq N \cdot D_N(\underline{Z}, \underline{P}) = m_N \cdot D_{m_N}(\underline{\tilde{Z}}, \underline{P}) \leq B_{m_N}.$$

Therefore, we have a stronger bound for original regret:

$$-A_{m_N} \leq N \cdot D_N(\underline{Z}, \underline{P}) \leq B_{m_N}.$$

When $\{X_i\}$, $\{Y_i\} \in \{0, 1\}$, $Z_i \in \{-1, 1\}$ which is equivalent to $\{0, 1\}$ matching binary bits problem. In fact, any two state game,i.e. two Players' action set is $\{a, b\}$, is equivalent to $\{0, 1\}$ matching binary bits problem.

Lemma 5.3.1. *If in a repeated game $Z_k \in \{-1, 1\}$, $P_k \in [0, 1]$, $L(Z, P) = Z \cdot P$, then the regret of this game is equal to the regret of matching binary bits, i.e.,*

$$D_N(\underline{Z}, \underline{P}) = D^*_N(\underline{X}, \underline{P})$$

where $X \in \{0, 1\}$, $P \in [0, 1]$ is the matching binary bits problem.

Proof: By the definition of the loss function,

$$L_k(Z_k, P_k) = \begin{cases} -P_k & \text{if } Z_k = -1 \\ P_k & \text{if } Z_k = 1. \end{cases}$$

Where $P_k = Prob(\overline{Z}_{k-1} \leq 0)$. And from matching binary bits game, $X \in \{0, 1\}$, $P \in [0, 1]$, and

$$L^*_k(X_k, P_k) = \begin{cases} P_k & \text{if } X_k = 0 \\ 1 - P_k & \text{if } X_k = 1. \end{cases}$$

where $P_k = Prob(\overline{X}_{k-1} \geq \frac{1}{2})$. Then, let $q_k = 1 - P_k = Prob(\overline{X}_{k-1} \leq \frac{1}{2})$.

$$L_k^*(X_k, P_k) = X_k \cdot q_k + (1 - X_k) \cdot (1 - q_k) = 2X_k \cdot q_k + 1 - q_k + X_k.$$

Since there exist the one to one transformation mapping:

$$X = \frac{1}{2}(Z + 1) \quad or \quad Z = 2X - 1.$$

Then $\overline{Z}_k = 2\overline{X}_k - 1$, i.e. $Prob(\overline{Z}_{k-1} \leq 0) = Prob(2\overline{X}_k - 1 \leq 0) = q_k$.

Thus,

$$\begin{aligned} L_N(\underline{Z}_N, \underline{P}_N) &= \frac{1}{N}\sum_{i=1}^{N}(2X_i - 1) \cdot P_i \\ &= \frac{1}{N}\sum_{i=1}^{N}(2X_i - 1) \cdot P_i + 1 - X_i - (1 - X_i) \\ &= L_N^*(\underline{X}_N \cdot \underline{q}_N) - (1 - \overline{X}_N) \end{aligned}$$

Therefore, since $\overline{Z}_k = 2\overline{X}_k - 1$ the Bayes envelope of matching binary bits game:

$$\overline{X}_N \wedge (1 - \overline{X}_N = (2\overline{X}_k - 1) \wedge 0 + (1 - \overline{X}_N).$$

the regret of Z-sequence is:

$$\begin{aligned} D_N(Z, P) &= L_N(\underline{Z}_N, \underline{P}_N) - \overline{Z}_N \wedge 0 \\ &= L_N^*(\underline{X}_N \cdot \underline{q}_N) - (1 - \overline{X}_N) - (2\overline{X}_k - 1) \wedge 0 \\ &= L_N^*(\underline{X}_N, \underline{q}_N) - ((2\overline{X}_k - 1) \wedge 0 + (1 - \overline{X}_N)) \\ &= D_N^*(X, P). \end{aligned}$$

Proof is done.□

Comments 5.3.1. *With proof in lemma, study of two-experts selection system, is* *equivalent to the study on sequence $\underline{Z}$, where $Z_i = X_i - Y_i$. And this is used in following sections.*

5.3.2 Worst case discussion

In Gilliland and Jung (2006), asymptotic convergence property was proved by considering the worst case $\underline{a}$ for the strategy. For Matching Binary Bits problem, the worst case for both PAP strategy and PARP is:

$$\underline{a}: 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1, \ldots$$

By establishing a bound for the modified regret of this worst case, a uniform bound of the regret for all situations of Player I's play sequences $\underline{a}$ was developed.

To study the asymptotic convergence property of PARP strategy in the two-experts selection problem, it is reasonable to seek and analyze a worst case.

Lemma 5.3.2. *The worst case of modified regret of PARP strategy is not achieved at boundary, i.e.* $\max_{\underline{a}} D_N(\underline{a}, \underline{b})$ *is not achieved on the boundary.*

Proof: Suppose n=3, so two experts system is:

$$(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$$

the modified regret of Play Against Random Past strategy is defined as:

$$\begin{aligned} 3 \cdot D_3(\underline{a}, \underline{b}) &= \frac{1}{2}(X_1 + Y_1) + X_2[X_1 \le Y_1] + X_3 \cdot P_2 + Y_3 \cdot (1 - P_2) \\ &- (X_1 + X_2 + X_3) \wedge (Y_1 + Y_2 + Y_3), \end{aligned}$$

where the probability $P_2 = Prob(\overline{X}^*_2 \le \overline{Y}^*_2)$ and $[\cdot]$ is an indicator function.

By discussion of last section, two-experts selection problem is isomorphic to Z-problem. Let $Z_i = X_i - Y_i$, then

$$\begin{aligned} 3 \cdot D_3(\underline{a}, \underline{b}) &= \frac{1}{2}(Z_1 + 2Y_1) + Z_2[X_1 \leq Y_1] + Y_2 + Z_3 \cdot P_2 + Y_3 \\ &- (Z_1 + Z_2 + Z_3) \wedge 0 - (Y_1 + Y_2 + Y_3), \end{aligned}$$

where the density function $P_2 = Prob(\overline{Z}_2^* \leq 0)$. i.e.

$$3 \cdot D_3(\underline{a}, \underline{b}) = \frac{1}{2}Z_1 + Z_2 \cdot [Z_1 < 0] + Z_3 \cdot P_2 - (Z_1 + Z_2 + Z_3) \wedge 0$$

According to Play Against Random Past strategy, the probability mass function of $\overline{Z}_2^*$ from bootstrap sample $\{Z_1, Z_2\}$ is:

$$p_2 = \begin{cases} \frac{1}{4} & \text{when choose } Z_1 \text{ twice} \\ \frac{1}{2} & \text{when choose once } Z_1 \text{ and once } Z_2. \\ \frac{1}{4} & \text{when } Z_2 \text{ twice} \end{cases}$$

Suppose $(Z_1 + Z_2 + Z_3) < 0$ i.e. Bayes envelope $R_3 = Z_1 + Z_2 + Z_3$.

Further more, assume $Z_1 > 0$, then

$$\begin{aligned} 3 \cdot D_3(\underline{a}, \underline{b}) &= \frac{1}{2}Z_1 + Z_2 \cdot 0 + Z_3 \cdot P_2 - (Z_1 + Z_2 + Z_3) \\ &= -\frac{1}{2}Z_1 - Z_2 - Z_3 \cdot (1 - P_2) \end{aligned}$$

Since $Z_1 > 0$, to achieve maximum of the regret

$$\max_{Z_1 > 0, Z_1 + Z_2 + Z_3 < 0} 3 \cdot D_3(\underline{a}, \underline{b}),$$

Z_2 and Z_3 should be negative. When $Z_1 > 0$, $Z_2 < 0$,

$$P_2 = \begin{cases} \frac{1}{4} & \text{when } Z_1 + Z_2 > 0 \\ \frac{3}{4} & \text{when } Z_1 + Z_2 \leq 0. \end{cases}$$

If $P_2 = \frac{1}{4}$,

$$\begin{aligned} \max_{Z_1>0, Z_1+Z_2+Z_3<0} 3 \cdot D_3(\underline{a}, \underline{b}) &= -\frac{1}{2}Z_1 - Z_2 - (-1) \cdot (1 - \frac{1}{4}) \\ &\approx 1.25 \end{aligned}$$

where $Z_1 = 1$, $Z_2 = -0.9999$, and $Z_3 = -1$, since $Z_1 > 0$, $Z_2 < 0$ and $Z_1 + Z_2 > 0$. The nearer to -1 Z_2 is, the better, but Z_2 can not be -1.

If $P_2 = \frac{3}{4}$,

$$\begin{aligned} \max_{Z_1>0, Z_1+Z_2+Z_3<0} 3 \cdot D_3(\underline{a}, \underline{b}) &= -\frac{1}{2}Z_1 - Z_2 - (-1) \cdot (1 - \frac{3}{4}) \\ &\approx 1.25 \end{aligned}$$

where $Z_1 = 0.0001$, $Z_2 = -1$, and $Z_3 = -1$, since $Z_1 > 0$, $Z_2 < 0$ and $Z_1 + Z_2 \leq 0$. The nearer to 0 Z_1 is, the better, but Z_1 can not be 0.

For the case of $Z_1 \leq 0$,

$$\begin{aligned} 3 \cdot D_3(\underline{a}, \underline{b}) &= \frac{1}{2}Z_1 + Z_2 \cdot 1 + Z_3 \cdot P_2 - (Z_1 + Z_2 + Z_3) \\ &= -\frac{1}{2}Z_1 - Z_3 \cdot (1 - P_2) \end{aligned}$$

Base on the definition of probability mass function P_2, we have

$$P_2 = \begin{cases} \frac{1}{4} & \text{when } Z_1 + Z_2 > 0, \text{ since } Z_1 \leq 0, Z_2 > 0 \\ \frac{3}{4} & \text{when } Z_1 + Z_2 \leq 0. \end{cases}$$

When $P_2 = \frac{1}{4}$,

$$\begin{aligned} \max_{Z_1 \leq 0, Z_1+Z_2+Z_3<0} 3 \cdot D_3(\underline{a}, \underline{b}) &= -\frac{1}{2} Z_1 - (-1) \cdot (1 - \frac{1}{4}) \\ &\approx 1.25 \end{aligned}$$

where $Z_1 = -0.9999$, $Z_2 = 1$, and $Z_3 = -1$, since $Z_1 \leq 0$, $Z_2 > 0$ and $Z_1 + Z_2 > 0$. The nearer to -1 Z_1 is, the better, but Z_1 can not be -1.

When $P_2 = \frac{3}{4}$,

$$\begin{aligned} \max_{Z_1 \leq 0, Z_1+Z_2+Z_3<0} 3 \cdot D_3(\underline{a}, \underline{b}) &= -\frac{1}{2} \cdot (-1) - (-1) \cdot (1 - \frac{1}{4}) \\ &= 1.25 \end{aligned}$$

where $Z_1 = -1$, $Z_2 = \{any \quad value \in [-1, 1] | Z_1 + Z_2 \leq 0\}$, and $Z_3 = -1$.

All the calculations above, shows even for n=3, the maximum of the modified regret of PARP Strategy is not achieved at the boundary of the problem domain $[0, 1]^n \times [0, 1]^n$ which is equivalent to the domain $[-1, 1]^n$ for the Z-problem.

Proof is done.$\square$

Therefore, the proof of Hannan consistency of PARP strategy for two-experts selection problem can not be studied through the worst case idea.

5.3.3 Hannan consistency of PARP for Certain classes

We concern the asymptotic convergence property of PARP strategy's regret under different classes of sequences. With discussion in Z-problem, we notice that the original problem is equivalent to Z-problem, i.e. we only need to discuss the convergence property of Z-sequence.

$$\underline{Z} : Z_1, \quad Z_2, \quad Z_3, \ldots \quad Z_{N-1}, \quad Z_N, \ldots$$

The most easy case is the one expert is superior to the other one, i.e. in sequence $\underline{Z}$, all Z_t has the same sign. Let $Z_t > 0$ without loss of generality. Then, for all $N > 0$,

$$N \cdot D_N(\underline{Z}, \underline{parp}) = \sum_{k=1}^{N} Z_k \cdot P(\overline{Z}^*_{k-1} \leq 0) - N\overline{Z}_N \wedge 0 = \sum_{k=1}^{N} Z_k \cdot P(\overline{Z}^*_{k-1} \leq 0) - 0 = 0.$$

where $\overline{Z}^*_{k-1}$ is the sample mean of bootstrap sample in PARP strategy.

This means if one expert's prediction is always better than the other, the regret of PARP strategy is always 0.

More difficult situation is the two experts are competing with each other.

Theorem: For any sequence $\underline{Z}$, $Z_k \in [-1, 1]$, if $\sigma_k \geq C1 > 0$, the regret of PARP strategy converges to 0 with order $O(\frac{1}{\sqrt{N}})$, i.e.,

$$D_N(\underline{Z}, \underline{parp}) \to 0 \quad with \quad rate \quad O(\frac{1}{\sqrt{N}}).$$

where $C1$ is a constant and σ_k^2 is the variance of sequence $\{Z_1, \ldots, Z_k\}$.

Proof:

By Berry-Esseen theorem in Loeve (1963, pp 288),

$$N \cdot D_N(\underline{Z}, \underline{p}^*) = \sum_{k=1}^{N} Z_k \cdot (\Phi(-\frac{\overline{Z}_{k-1} \cdot \sqrt{k-1}}{\sigma_{k-1}})) + \sum_{k=1}^{N} Z_k \cdot \frac{C}{\sqrt{k-1}} \cdot \frac{\rho_k}{\sigma_k^3}$$

where C is a constant, $\rho_k = E|Z^*_{i,k} - \overline{Z}_k|^3$.

Now we consider the term:

$$\sum_{k=1}^{N} Z_k \cdot (\Phi(-\frac{\overline{Z}_{k-1} \cdot \sqrt{k-1}}{\sigma_{k-1}}))$$

Let $A_{k-1} = -\frac{1}{\sigma_{k-1}} \cdot (\overline{Z}_{k-1} \cdot \sqrt{k-1})$, and since $Z_k = k \cdot \overline{Z}_k - (k-1) \cdot \overline{Z}_{k-1}$, we can write the term as:

$$\sum_{k=1}^{N}(k\cdot\overline{Z}_k-(k-1)\cdot\overline{Z}_{k-1})\cdot\Phi(A_{k-1}).$$

It follows that

$$\begin{aligned}\sum_{k=1}^{N}(k\cdot\overline{Z}_k &- (k-1)\cdot\overline{Z}_{k-1})\cdot\Phi(A_{k-1})\\ &= \sum_{k=1}^{N}k\cdot\overline{Z}_k\cdot(\Phi(A_{k-1})-\Phi(A_k))-0+N\overline{Z}_N\cdot\Phi(A_N)\\ &= \sum_{k=1}^{N}k\cdot\overline{Z}_k\cdot\int_{A_k}^{A_{k-1}}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}dx+N\overline{Z}_N\cdot\Phi(A_N)\end{aligned}$$

Therefore,

$$\begin{aligned}N\cdot D_N(\underline{Z}.\underline{parp}) &= \sum_{k=1}^{N}k\cdot\overline{Z}_k\cdot\int_{A_k}^{A_{k-1}}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}dx+N\overline{Z}_N\cdot\Phi(A_N)\\ &+ \sum_{k=1}^{N}Z_k\cdot\frac{C}{\sqrt{k-1}}\cdot\frac{\rho_k}{\sigma_k^3}.\end{aligned}$$

Let Part I$=\sum_{k=1}^{N}k\cdot\overline{Z}_k\cdot\int_{A_k}^{A_{k-1}}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}dx$, Part II$=N\overline{Z}_N\cdot\Phi(A_N)$, and Part III$=\sum_{k=1}^{N}Z_k\cdot\frac{C}{\sqrt{k-1}}\cdot\frac{\rho_k}{\sigma_k^3}$.

For Part I, suppose $\sigma_{k-1}<\sigma_k$ without loss of generality. Then

$$\begin{aligned}|A_{k-1}-A_k| &\le |-\frac{1}{\sigma_{k-1}}\cdot(\overline{Z}_{k-1}\cdot\sqrt{k-1}-\overline{Z}_k\cdot\sqrt{k})|\\ &= |-\frac{1}{\sigma_{k-1}\sqrt{k}}\cdot(\overline{Z}_{k-1}\cdot\sqrt{k-1}\sqrt{k}-\overline{Z}_k(k-1)+\overline{Z}_k(k-1)-\overline{Z}_k k)|\\ &= \frac{1}{\sigma_{k-1}\sqrt{k}}|Z_k+(k-1)\overline{Z}_{k-1}-\overline{Z}_{k-1}\cdot\sqrt{k-1}\cdot\sqrt{k}|\\ &\le \frac{1}{\sigma_{k-1}\sqrt{k}}(|Z_k|+|\sqrt{k-1}\cdot\overline{Z}_{k-1}\cdot(\sqrt{k-1}-\sqrt{k})|)\end{aligned}$$

Since $\sqrt{k-1}-\sqrt{k}=\frac{C}{\sqrt{k-1}}$, $\sigma_k \geq C1 > 0$

$$|A_{k-1}-A_k| \leq \frac{1}{\sigma_{k-1}\sqrt{k}}(|Z_k| + C\cdot|\overline{Z}_{k-1}|) \leq O(\frac{1}{\sqrt{k}}) + O(\frac{1}{\sqrt{k}})\frac{1}{\sigma_{k-1}}\cdot|\overline{Z}_{k-1}|.$$

Thus,

$$\int_{A_k}^{A_{k-1}} \frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}dx \leq |A_{k-1}-A_k|\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{A_k^2}{2}}$$

if we assume $\phi(A_k)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{A_k^2}{2}} > \phi(A_{k-1}) = \frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{A_{k-1}^2}{2}}$ without loss of generality, then

$$Part \quad I \leq \sum_{k=1}^{N} O(\frac{1}{\sqrt{k}}) + O(\frac{1}{\sqrt{k}})\cdot\frac{1}{\sqrt{2\pi}}\cdot k\cdot\frac{\overline{Z}_k^2}{\sigma_{k-1}}\cdot e^{-\frac{k\overline{Z}_k^2}{2\sigma_k^2}}$$

since $k\cdot\frac{\overline{Z}_k^2}{\sigma_{k-1}}\cdot e^{-\frac{k\overline{Z}_k^2}{2\sigma_k^2}} \leq C$ and $0 < C1 \leq \sigma_k < C$, we have
Part I$\leq \sum_{k=1}^{N} O(1/\sqrt{k}) = O(\sqrt{N})$.

For Part II$=N\overline{Z}_N\cdot\Phi(A_N)$,
if $\overline{Z}_N \leq 0$, then Part II≤ 0;
if $\overline{Z}_N > 0$, then

$$N\overline{Z}_N\cdot\Phi(A_N) = N\overline{Z}_N\cdot\Phi(-\frac{\overline{Z}_N\sqrt{N}}{\sigma_N}) = \sqrt{N}\cdot(\sqrt{N}\overline{Z}_N)\cdot\Phi(-\frac{\overline{Z}_N\sqrt{N}}{\sigma_N}).$$

Let $w_N=\sqrt{N}\cdot\overline{Z}_N$, then by Feller (1964, pp 166),

$$w_N\cdot\int_{-\infty}^{-\frac{w_N}{\sqrt{2}}}\sqrt{\frac{1}{2\pi}}\cdot e^{-\frac{x^2}{2}}dx = w_N\cdot(1-\int_{-\infty}^{\frac{w_N}{\sqrt{2}}}\sqrt{\frac{1}{2\pi}}\cdot e^{-\frac{x^2}{2}}dx) \sim w_N\cdot\frac{\sqrt{2}}{\sqrt{2\pi}w_N}e^{-\frac{w_N^2}{4}} \leq C$$

Thus,

$$\begin{aligned} Part \quad II &= \sqrt{N} \cdot w_N \cdot \Phi(-\frac{w_N}{\sigma_N}) \leq \sqrt{N} \cdot w_N \cdot \Phi(-\frac{w_N}{\sqrt{2}}) \\ &= \sqrt{N} \cdot w_N \cdot \int_{-\infty}^{-\frac{w_N}{\sqrt{2}}} \sqrt{\frac{1}{2\pi}} \cdot e^{-\frac{x^2}{2}} dx \\ &\leq \sqrt{N} \cdot C. \end{aligned}$$

i.e. $Part \quad II \leq O(\sqrt{N})$.

For Part III$= \sum_{k=1}^{N} Z_k \cdot \frac{C}{\sqrt{k-1}} \cdot \frac{\rho_k}{\sigma_k^3}$,

by the definition of $Z_{i,k}^*$ and $|Z_k| \in [-1, 1]$,

$$\begin{aligned} \rho_k &\approx \frac{1}{k} \sum_{i=1}^{k} |Z_{i,k}^* - \overline{Z}_k|^3 \\ &\leq C \end{aligned}$$

Then,

$$|Z_k \cdot \frac{\rho_k}{\sigma_k^3}| \leq C.$$

i.e. $Part \quad III \leq O(\sqrt{N})$.

With discussion on Part I, Part II, and Part III, we have $D_N(\underline{Z}, \underline{parp}) \leq O(1/\sqrt{N})$.

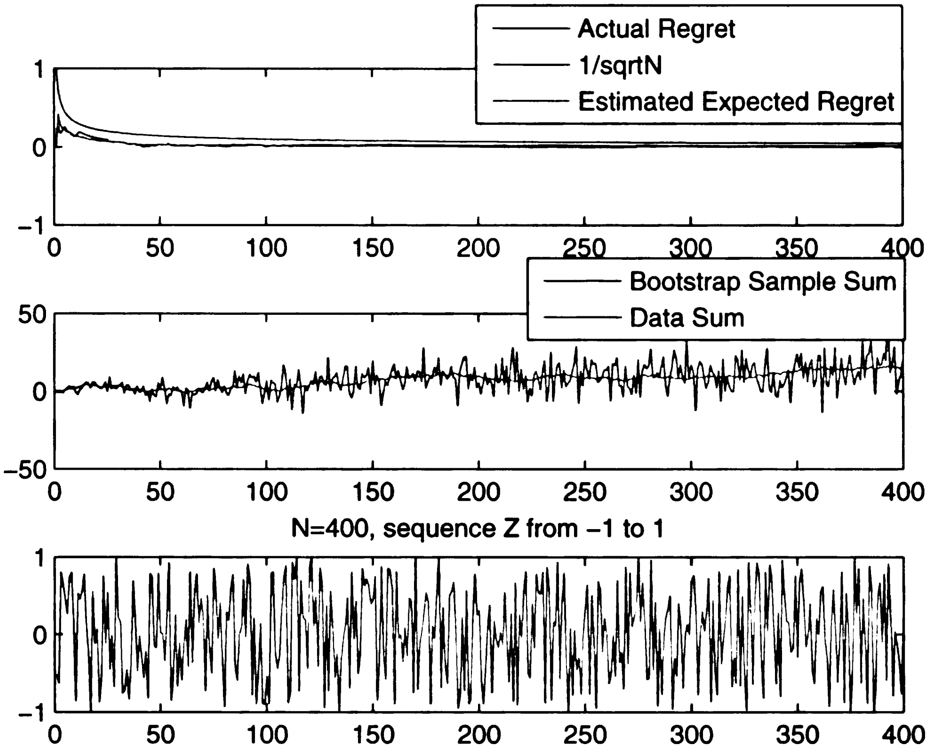
i.e.

$$D_N(\underline{Z}, \underline{parp}) \rightarrow 0 \quad with \quad rate \quad O(\frac{1}{\sqrt{N}}).$$

Proof is done.□

The group of three figures below shows the simulation of PARP strategy for sequence $\underline{z}$, where $z \in [-1, 1]$ and $\sigma_k > 0.1$, which is shown in the figure in row3. The figure of row 1 is actual Cesaro loss-Bayes Envelope (in blue) vs expected Cesaro loss - Bayes Envelope (in red) and the function $1/\sqrt{N}$. The figure in row 2 is S^* (in blue) vs S (in green).

Figure 5.1: Simulation



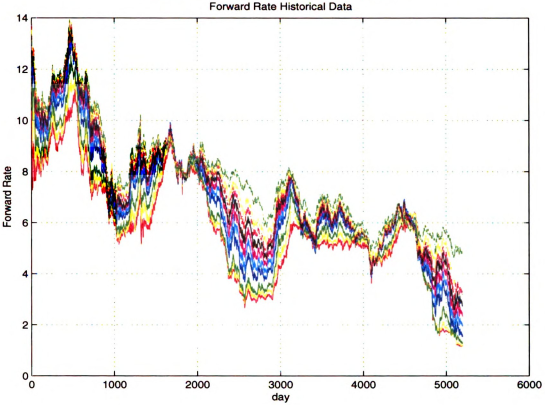
5.4 Examples of the application of PARP strategy

As a statistical decision strategy in expert selection problem, PARP strategy can be applied to many kinds of practical problems. Here we only give an example of the application of this strategy in two-expert system.

By Hull (2002) in finance, there is a constant effort to predict future or forward prices of stocks, bonds, options and commodities; the ability to predict future behavior provides important information about the underlying structure of these securities.

In interest rate market, many different types of interest rates are regularly quoted. These include mortgage rates, deposit rates, prime borrowing rates, and so on. As a member of interest rate market, the n-year *zero rate* or *spot rate* is defined as the rate of interest earned on an investment that starts today and last for n years. All the interest and principal is realized at the end of n years. There are no intermediate payments. Forward rates or forward interest rates are the rates of interest implied by current zero rates for periods of time in the future. The graph shows the movements

Figure 5.2: Forward Rate History Data

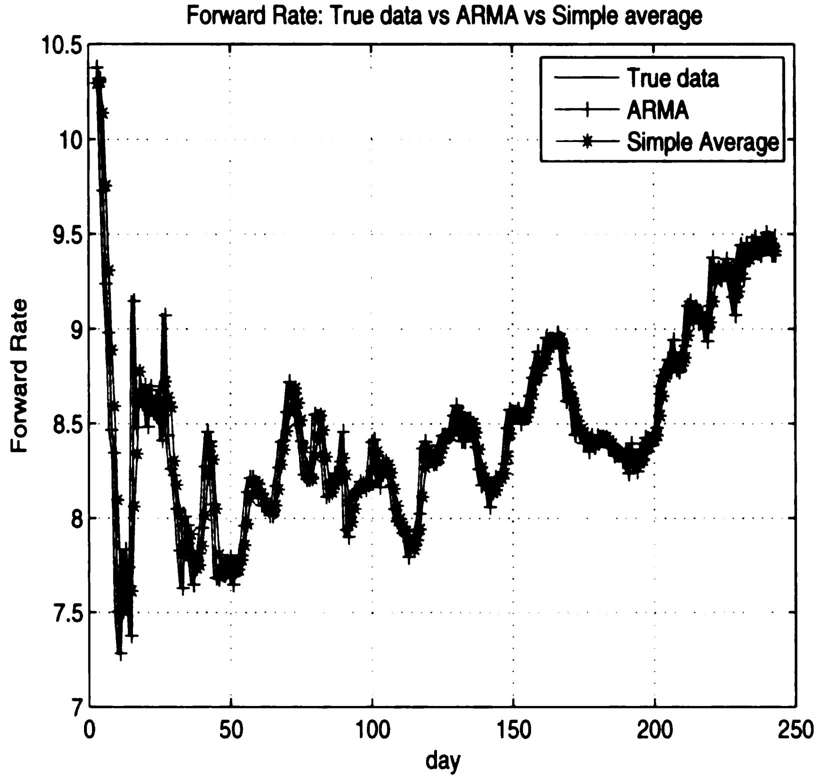


of various forward rates of US market data from 1983 to 2003. It includes forward rates for 3 months, 6 months, 1 year, 2 years, 3 years, 4 years, 5 years, 7 years, 10 years and 30 years.

We are going to use forward rate for 3 month as an example to show how PARP strategy is applied on it. In our two expert system, Expert I is ARMA model and Expert II is just the average rate for the recent last 3 days. Blue line represent the true data; red line is Expert I, i.e. ARMA model; Green line is Expert II which is average rate of past 3 days.

We take 1 cycle=120 work days, then list two experts' errors and the PARP strategy's error in the table 5.1. From the table above, it is easily to observe that

Figure 5.3: Forward Rate prediction: True data vs ARMA vs Simple Average



PARP strategy automatically choose the better model,i.e. Expert I which has more precise predictions.

In this example, since Expert I the ARMA model is superior to Expert II at most of time, it is reasonable that PARP strategy converges to Expert I's decision, and the graph 5.4 also shows that average loss of PARP strategy converges to the Bayes Envelopes, which agrees with the theoretical proof in previous section with the sample standard deviation $\sigma_k > 0.01$ for all $k = 1, 2, \ldots, 120$ in this example.

Situation is more complicated if the two experts' forecast are quite close. For example, Expert I is still ARMA model, but Expert II is GARCH(1,1) model with their prediction showed in the graph.

We still keep 1 cycle=120 work days,then the comparison between experts and PARP strategy's errors are showed in Table 5.2.

The simulation shows the PARP strategy works well. The graph 5.6 also shows

Table 5.1: Forward Rate prediction: ARMA vs Simple Average

Number of Cycle	Ave Loss Expert I	Ave loss Expert II	Ave Loss of PARP
cycle 1	0.0759	0.1083	0.0776
cycle 2	0.0711	0.0820	0.0699
cycle 3	0.0698	0.0830	0.0749
cycle 4	0.0499	0.0719	0.0522
cycle 5	0.0434	0.0591	0.0440

Figure 5.4: Forward Rate: Bayes Envelope vs PARP Average Loss for cycle 5

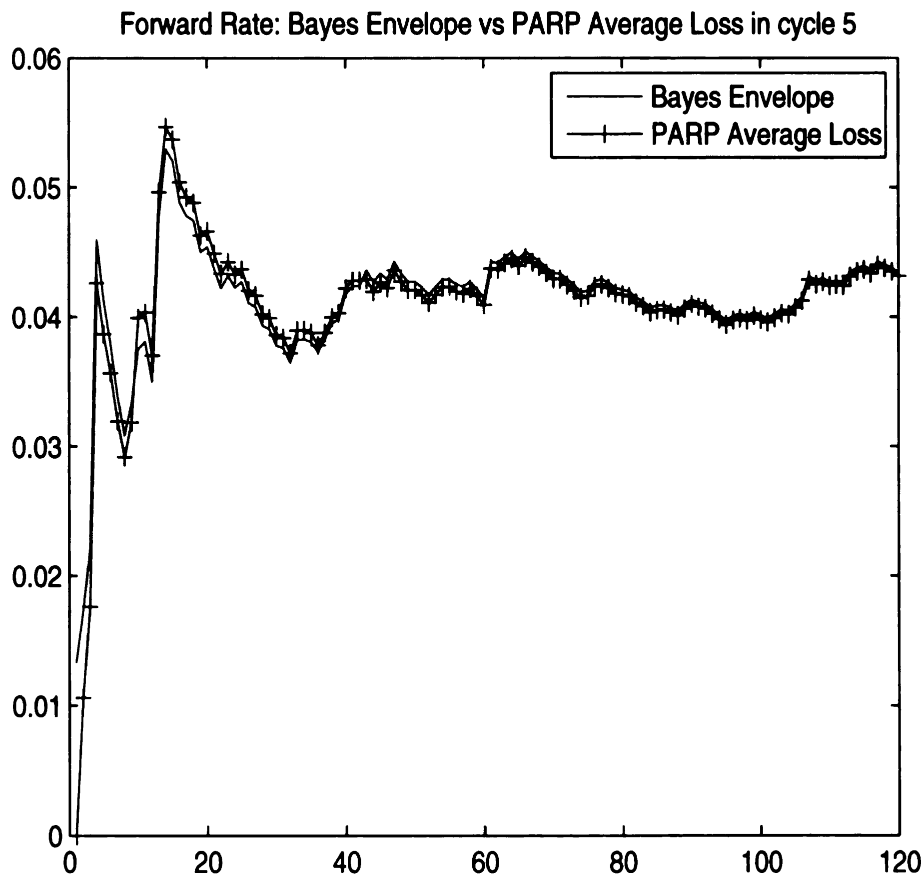


Table 5.2: Forward Rate prediction: ARMA vs GARCH(1,1)

Number of cycles	Ave loss of Expert I	Ave loss Expert II	Ave Loss of PARP
cycle 1	0.0707	0.0719	0.0706
cycle 2	0.0500	0.0503	0.0500
cycle 3	0.0697	0.0714	0.0689
cycle 4	0.0635	0.0640	0.0635
cycle 5	0.0673	0.0676	0.0666

Figure 5.5: Forward Rate: True data vs ARMA vs GARCH(1,1)

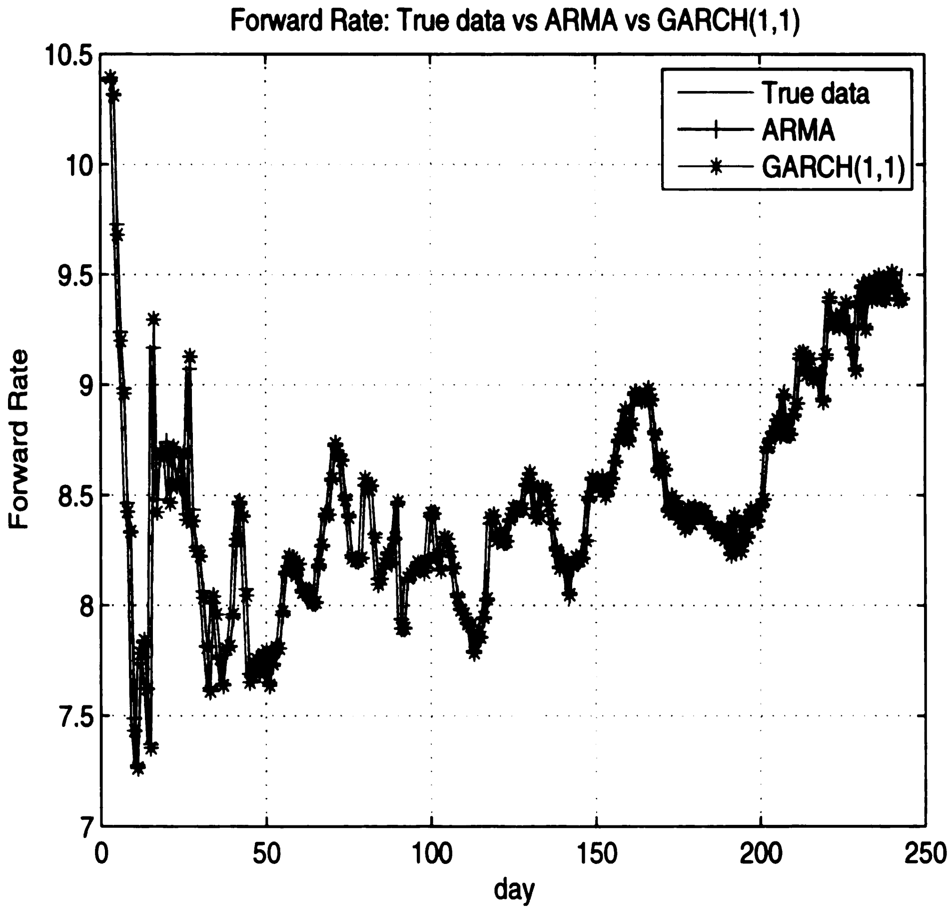
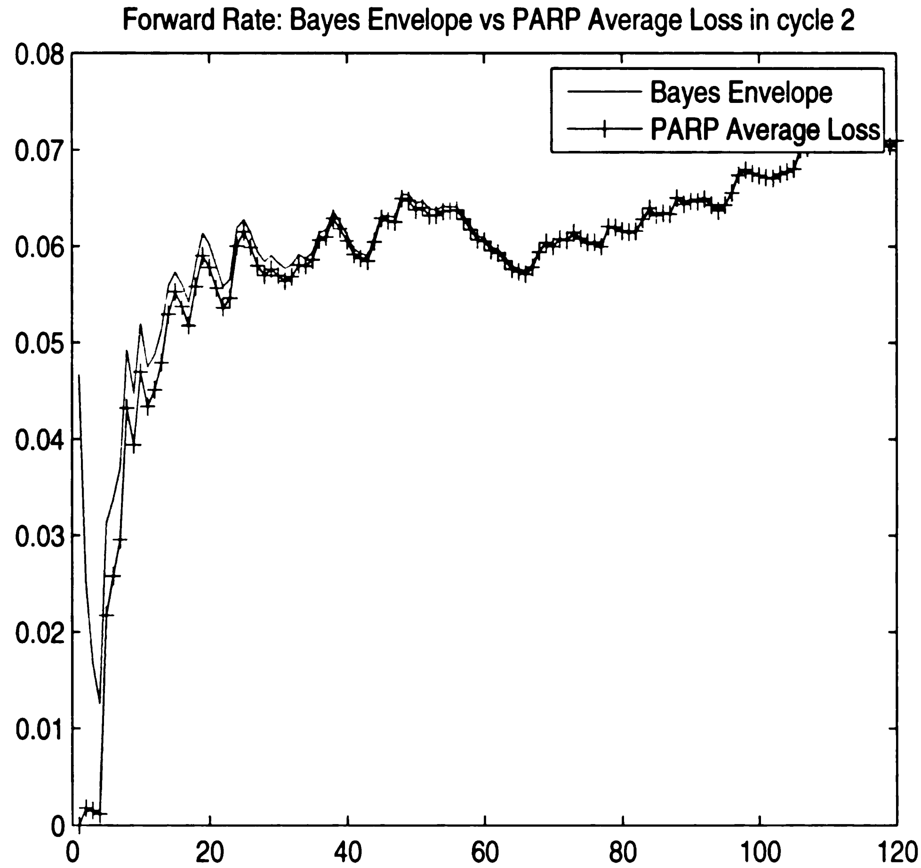


Figure 5.6: Forward Rate: Bayes Envelope vs PARP Average Loss for cycle 2



that average loss of PARP strategy converges to the Bayes Envelopes, which agrees with the theoretical proof in previous section with $\sigma_k > 0.001$ for all $k = 1, 2, \ldots, 120$ in this example.

5.5 Future work

Future work will include work on the two-expert selection problem and the k-expert selection problem. For the two-expert selection problem, the goal is to extend proofs of Hannan consistency for the PARP strategy to the general case covering all sequences $\underset{\sim}{z} \in [-1,1]^{\infty}$. To accomplish this, we need to get approximations of $P(\overline{Z}_k^* \leq 0)$ for the general case.

We are looking forward to understanding and discovering more properties about the distribution of $\overline{Z}_k^*$ in the future. One possible approach may be creating some bins on the domain of Z_k, i.e., make $[-1,1]$ in to several categories, in order to make the domain of Z_k a discrete set. Another one may be considering the change of $P(\overline{Z}_i^* \leq 0)$ from stage $i = k$ to stage $i = k+1$. These ideas will be worked on and discussed in the future.

There are still a lot of open problems in this field as well. For example, since sometimes more recent past moves are more important to the decision, time-weighted PARP strategy can be constructed and its Hannan consistency can be studied in the future. Also, in non-symmetric repeated game, construction of PAP and PARP strategies and their asymptotic properties are very interesting and can investigate in the future as well.

Bibliography

[1] N. Cesa-Bianchi, G. Lugosi, On prediction of individual sequences, *The Annals of Statistics*, Vol. 27, No.6, pp. 1865–1895, 1999.

[2] L. Chen, Application of play against past strategies in repititions of a game. *PhD Dissertation, Department of Statistics and Probability, Michigan State Univeristy*, 1997.

[3] T. H. Chung, Minimax learning in iterated games via distributional majorization. Ph. D dissertation, Stanford University, 1994.

[4] T. M. Cover, Behavior of sequential predictors of binary sequences. *Transactions of the Fourth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Publishing House of the Czechoslovak Academy of Sciences*, pp. 263–272, 1967.

[5] M. Feder, N. Merhav and M. Gutman, Universal prediction of individual sequences, *IEEE Trans. Inform. Theory* 38 (1992) (4), pp. 1258-1270.

[6] W. Feller, An introduction to probability theory and its applications. Vol. I, Second Edition: Wiley and Sons, NY, 1964

[7] D. P. Foster, R. V. Vohra, A randomization Rule for selection forecasters *Operations Reserach*, 41, No. 4, pp. 704–709.

[8] J. S. Frame, Dennis C. Gilliland, Votes and a half-binomial, *Discrete Applied Mathematics*, 12, pp. 31–40, 1985.

[9] E. Gardner, E. Anderson-Flether, A. Wicks, Further results on focus forecasting vs. exponential smoothing, *International Journal of Forecating*, 17, pp. 287–293, 2001.

[10] D. Gilliland, I. Jung, Play against the random past for matching binary bits, *J. Statistical Theory and Applications*, 5, No.3, pp. 282–291, 2006.

[11] D. Gilliland and J. Hannan, Play against past strategies in repeated play of $(T, S, \cdot)$ game, 1999, unpublished.

[12] D. Gilliland and J. Hannan, Play against past strategies in repeated play of a S-game and its dual, Michigan State University, Department of Statistics and Probability, RM-603, 2008.

[13] D. Gilliland, Bootstrap fictitious play in repeated games, *Proc. Joint Stattistical Meetings*, Toronto, Aug, 8–12, 2004.

[14] D. Gilliland, Asymptotic risk stability resulting from play against the past in a sequence of decision problems, *IEEE Transcations on information theory*, Vol.IT-18, No.5, pp. 614–617, September 1972.

[15] D. Gilliland, J. Hannan, On an extended compound decision problem, *The Annals of Mathematical Statistics*, Vol 40, No.50, pp. 1536–1541, 1969.

[16] C. M. Grinstead, J. L. Snell, Introduction to Probability, 2008. *www.dartmouth.edu/ chance/teaching$_a$ids/books_articles*

[17] J. F. Hannan, Approximation to Bayes risk in repeated play *Contributions to the Theory of Games, Princeton University Press*, 1957.

[18] M. D. Hirschhorn, Comments on the paper "Wallis sequence..." by Lampret. Aust. Math. Soc. Gazette 32, 194, 2005.

[19] J. Hull, Options, Futures & other derivatives, fifth edition, 2002.

[20] M. V. Johns Jr., Two-action compound decision problems. *Proc. Fifth Berkley Symp. Math. Statist. Prob.* 1 463-478, University of California Press, 1967.

[21] N. Littlestone, Learing quickly when irrelavent attributes abound: a new linear-threshold algorithm, *Machine Learning*, 2:285-318, 1988.

[22] N. Littlestone and M. K. Warmuth, The weighted majority algarithm. *Symposium on Foudations of Computer Science*, 30: 256–261, November 1989.

[23] M. Loeve, Probability Theory, Third Edition, D. Van Nostrand Company, Inc, NJ, 1963.

[24] J. Robinson, An iterative method of solving a game, *Annals of Mathematics*, 54, pp. 296–301, 1951.

[25] B. Smith, Focus corecasting computer techniques for inventory control: revised for the twenty-first centrury, 1997.

[26] D. Swain, Bounds and rates of convergence for the extended compound estimate problem in the sequence case. *Tech. Report* No. 81, Department of Statistics, Standford Unviersity, 1965.

[27] N. N. Vorob'ev, Game theory lectures for ecnomitsts and systems scientists, Springer-Verlag.