MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

On a simple strategy weakly forcing the strong law of large numbers in the bounded forecasting game

Masayuki Kumon and Akimichi Takemura

METR 2005–20 August 2005

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/mi/mi-e.htm

On a simple strategy weakly forcing the strong law of large numbers in the bounded forecasting game

Masayuki Kumon
Risk Analysis Research Center
Institute of Statistical Mathematics
and
Akimichi Takemura
Graduate School of Information Science and Technology
University of Tokyo

August, 2005

Abstract

In the framework of the game-theoretic probability of Shafer and Vovk (2001) it is of basic importance to construct an explicit strategy weakly forcing the strong law of large numbers (SLLN) in the bounded forecasting game. We present a simple finite-memory strategy based on the past average of Reality's moves, which weakly forces the strong law of large numbers with the convergence rate of $O(\sqrt{\log n/n})$. We also give a detailed analysis of the paths of Skeptic's capital process for the case of the fair-coin game when our strategy is used. We show that if Reality violates SLLN, then the exponential growth rate of Skeptic's capital process is explicitly described in terms of the Kullback divergence between the average of Reality's moves when she violates SLLN and the average when she observes SLLN.

Keywords and phrases: Azuma-Hoeffding-Bennett inequality, capital process, game-theoretic probability, Kullback divergence, large deviation.

1 Introduction

The book by Shafer and Vovk (2001) established the whole new field of game-theoretic probability and finance. Their framework provides an attractive alternative foundation of probability theory. Compared to the conventional measure theoretic probability, the game theoretic probability treats the sets of measure zero in a very explicit way when proving various probabilistic laws, such as the strong law of large numbers. In a game-theoretic proof, we can explicitly describe the behavior of the paths on a set of measure zero, whereas in measure-theoretic proofs the sets of measure zero are often simply ignored. This feature of game-theoretic probability is well illustrated in the explicit construction of Skeptic's strategy forcing SLLN in Chapter 3 of Shafer and Vovk (2001).

However the strategy given in Chapter 3 of Shafer and Vovk (2001), which we call a mixture ϵ-strategy in this paper, is not yet satisfactory, in the sense that it requires combination of infinite number of "accounts" and it needs to keep all the past moves of Reality in memory. We summarize their construction in Appendix in a somewhat more general form than in Chapter 3 of Shafer and Vovk (2001). In fact in Section 3.5 of their book, Shafer and Vovk pose the question of required memory for strategies forcing SLLN.

In this paper we prove that a very simple single strategy, based only on the past average of Reality's moves is weakly forcing SLLN. Furthermore it weakly forces SLLN with the convergence rate of $O(\sqrt{\log n/n})$. In this sense, our result is a substantial improvement over the mixture ϵ-strategy of Shafer and Vovk. Since ϵ-strategies are used as essential building blocks for the "defensive forecasting" ([9]), the performance of defensive forecasting might be improved by incorporating our simple strategy.

Our thinking was very much influenced by the detailed analysis by Takeuchi ([7] and Chapter 5 of [6]) of the optimum strategy of Skeptic in the games, which are favorable for Skeptic. We should also mention that the intuition behind our strategy is already discussed several times throughout the book by Shafer and Vovk (see e.g. Section 5.2). Our contribution is in proving that the strategy based on the past average of Reality's moves is actually weakly forcing SLLN.

We also give a detailed analysis of the behavior of the paths of the log capital process when our strategy is used in the case of the fair-coin game. We shall prove the remarkable fact that if Reality violates SLLN, i.e., if the sample average of Reality's moves does not converge to 0, then the exponential growth rate of the Skeptic's capital is explicitly described by the Kullback divergence between the average of Reality's moves when she violates SLLN and the average when she observes SLLN.

In this paper we only consider weakly forcing by a strategy. A strategy weakly forcing an event E can be transformed to a strategy forcing E as in Lemma 3.1 of Shafer and Vovk (2001). We do not present anything new for this step of the argument.

The organization of this paper is as follows. In Section 2 we formulate the bounded forecasting game and motivate the strategy based on the past average of Reality's moves as an approximately optimum ϵ-strategy. In Section 3 we prove that our strategy is weakly forcing SLLN. In Section 4 we give a detailed analysis of the paths of Skeptic's capital process, when our strategy is used for the fair-coin game. In Appendix we give a summary of the mixture ϵ-strategy in Chapter 3 of Shafer and Vovk (2001).

2 Approximately optimum single ϵ-strategy for the bounded forecasting game

Consider the bounded forecasting game in Section 3.2 of Shafer and Vovk (2001).

Bounded Forecasting Game

Protocol:

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \dots$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $x_n \in [-1, 1]$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n x_n$.
END FOR

For a fixed ϵ, $|\epsilon| < 1$, the ϵ-strategy sets $M_n = \epsilon \mathcal{K}_{n-1}$. Under this strategy Skeptic's capital process $\mathcal{K}_n$ is written as $\mathcal{K}_n = \prod_{i=1}^n (1 + \epsilon x_i)$ or

$$\log \mathcal{K}_n = \sum_{i=1}^n \log(1 + \epsilon x_i).$$

For sufficiently small $|\epsilon|$, $\log \mathcal{K}_n$ is approximated as

$$\log \mathcal{K}_n \simeq \epsilon \sum_{i=1}^n x_i - \frac{1}{2}\epsilon^2 \sum_{i=1}^n x_i^2.$$

The right-hand side is maximized by taking

$$\epsilon = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2}.$$

In particular in the fair-coin game, where x_n is restricted as $x_n = \pm 1$, approximately optimum ϵ is given as

$$\epsilon = \bar{x}_n = \frac{1}{n}\sum_{i=1}^n x_i.$$

Actually, as shown by Takeuchi ([7]), it is easy to check that $\epsilon = \bar{x}_n$ exactly maximizes $\prod_{i=1}^n (1 + \epsilon x_i)$ for the case of the fair-coin game.

Of course, the above approximately optimum ϵ is chosen in hindsight, i.e., we can choose optimum ϵ after seeing the moves $x_1, \dots, x_n$. However it suggests to choose M_n based on the past average $\bar{x}_{n-1}$ of Reality's moves. Therefore consider a strategy

$$M_n = c\bar{x}_{n-1}\mathcal{K}_{n-1}. \tag{1}$$

In the next section we prove that for $0 < c \le 1/2$ this strategy is weakly forcing SLLN. The restriction $0 < c \le 1/2$ is just for convenience for the proof and the above argument suggests that $c = 1$ is the best choice, at least for the fair-coin game. In Section 4 we will investigate the behavior of the paths of the capital process for the fair-coin game in detail when (1) with $c = 1$ is used.

Compared to a single fixed ϵ-strategy $M_n = \epsilon \mathcal{K}_{n-1}$ or the mixture ϵ-strategy in Chapter 3 of Shafer and Vovk (2001), letting $\epsilon = c\bar{x}_{n-1}$ depend on $\bar{x}_{n-1}$ seems to be reasonable from the viewpoint of effectiveness of Skeptic's strategy. The basic reason is that as $\bar{x}_{n-1}$

deviates more from the origin, Skeptic should try to exploit this bias in Reality's moves by betting a larger amount. Clearly this reasoning is shaky because for each round Skeptic has to move first and Reality can decide her move after seeing Skeptic's move. However in the next section we show that the strategy in (1) is indeed weakly forcing SLLN with the convergence rate of $O(\sqrt{\log n/n})$.

3 Weakly forcing SLLN by past averages

In this section we prove the following result.

Theorem 3.1 *In the bounded forecasting game, if Skeptic uses the strategy (1) with* $0 < c \le 1/2$, *then* $\limsup_n \mathcal{K}_n = \infty$ *for each path* $\xi = x_1 x_2 \dots$ *of Reality's moves such that*

$$\limsup_n \frac{\sqrt{n}|\bar{x}_n|}{\sqrt{\log n}} > 1. \tag{2}$$

This theorem states that the strategy (1) weakly forces that $\bar{x}_n$ converges to 0 with the convergence rate of $O(\sqrt{\log n/n})$. Therefore it is much stronger than the mixture ϵ-strategy in Chapter 3 of Shafer and Vovk (2001), which only forces convergence to 0. A corresponding measure theoretic result was stated in Theorem 1 of Azuma (1967) as discussed in Remark 3.1 at the end of this section. The rest of this section is devoted to a proof of Theorem 3.1.

By comparing $1, 1/2, 1/3, \dots,$ and the integral of $1/x$ we have

$$\log(n+1) = \int_1^{n+1} \frac{1}{x} dx \le 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \le 1 + \int_1^n \frac{1}{x} dx = 1 + \log n.$$

Next, by summing up the term

$$\frac{1}{(k-1)k} = \frac{1}{k-1} - \frac{1}{k}$$

from $k = i$ to n, we have

$$\sum_{k=i}^{n} \frac{1}{(k-1)k} = \frac{1}{i-1} - \frac{1}{n} \tag{3}$$

or

$$\frac{1}{i-1} = \sum_{k=i}^{n} \frac{1}{(k-1)k} + \frac{1}{n}.$$

Now consider the sum

$$\sum_{i=2}^{n} \frac{i}{i-1} \bar{x}_i^2 = \sum_{i=2}^{n} \frac{1}{(i-1)i} (x_1 + \dots + x_i)^2. \tag{4}$$

When we expand the right-hand side, the coefficient of the term $x_j x_k$, $j < k$, is given by

$$2 \times \left(\frac{1}{(k-1)k} + \frac{1}{k(k+1)} + \cdots + \frac{1}{(n-1)n} \right) = 2 \times \left(\frac{1}{k-1} - \frac{1}{n} \right).$$

We now consider the coefficient of x_j^2. In (4) we have the sum from $i = 2$ and we need to treat x_1 separately. The coefficient of x_1^2 is $1 - 1/n$ from (3). For $j \geq 2$, the coefficient of x_j^2 is given by $1/(j-1) - 1/n$, as in the case of the cross terms. Therefore

$$\begin{aligned}
\sum_{i=2}^{n} \frac{i}{i-1} \bar{x}_i^2 &= 2 \sum_{1 \leq j < i \leq n} \Big(\frac{1}{i-1} - \frac{1}{n} \Big) x_j x_i + \Big(1 - \frac{1}{n}\Big) x_1^2 + \sum_{i=2}^{n} \Big(\frac{1}{i-1} - \frac{1}{n} \Big) x_i^2 \\
&= 2 \sum_{1 \leq j < i \leq n} \frac{1}{i-1} x_j x_i - \frac{2}{n} \sum_{1 \leq j < i \leq n} x_j x_i + x_1^2 + \sum_{i=2}^{n} \frac{1}{i-1} x_i^2 - \frac{1}{n} \sum_{i=1}^{n} x_i^2 \\
&= 2 \sum_{i=2}^{n} \bar{x}_{i-1} x_i - n \bar{x}_n^2 + x_1^2 + \sum_{i=2}^{n} \frac{1}{i-1} x_i^2. \qquad (5)
\end{aligned}$$

Write $\bar{x}_0 = 0$. Then the first sum on the right-hand side can be written from $i = 1$. Under this notational convention (5) is rewritten as

$$\sum_{i=1}^{n} \bar{x}_{i-1} x_i = \frac{1}{2} \sum_{i=2}^{n} \frac{i}{i-1} \bar{x}_i^2 + \frac{n}{2} \bar{x}_n^2 - \frac{1}{2} \Big(x_1^2 + \sum_{i=2}^{n} \frac{1}{i-1} x_i^2 \Big). \qquad (6)$$

Now the capital process $\mathcal{K}_n$ of (1) is written as

$$\mathcal{K}_n = \prod_{i=1}^{n} (1 + c \bar{x}_{i-1} x_i).$$

As in Chapter 3 of Shafer and Vovk (2001) we use

$$\log(1 + t) \geq t - t^2, \qquad |t| \leq 1/2.$$

Then for $0 < c \leq 1/2$ we have

$$\begin{aligned}
\log \mathcal{K}_n &= \sum_{i=1}^{n} \log(1 + c \bar{x}_{i-1} x_i) \\
&\geq c \sum_{i=1}^{n} \bar{x}_{i-1} x_i - c^2 \sum_{i=1}^{n} \bar{x}_{i-1}^2 x_i^2,
\end{aligned}$$

and under the restriction $|x_n| \leq 1, \forall n$, we can further bound $\log \mathcal{K}_n$ from below as

$$\log \mathcal{K}_n \geq c \sum_{i=1}^{n} \bar{x}_{i-1} x_i - c^2 \sum_{i=1}^{n} \bar{x}_{i-1}^2. \qquad (7)$$

By considering the restriction $|x_n| \le 1$, (6) is bounded from below as

$$
\begin{aligned}
\sum_{i=1}^n \bar{x}_{i-1} x_i &\ge \frac{1}{2}\sum_{i=2}^n \frac{i}{i-1}\bar{x}_i^2 + \frac{n}{2}\bar{x}_n^2 - \frac{1}{2}\Big(1+\sum_{i=2}^n \frac{1}{i-1}\Big) \\
&\ge \frac{1}{2}\sum_{i=2}^n \frac{i}{i-1}\bar{x}_i^2 + \frac{n}{2}\bar{x}_n^2 - \frac{1}{2}(2+\log(n-1)) \\
&\ge \frac{1}{2}\sum_{i=2}^n \bar{x}_i^2 + \frac{n}{2}\bar{x}_n^2 - \frac{1}{2}(2+\log(n-1)) \\
&\ge \frac{1}{2}\sum_{i=1}^n \bar{x}_i^2 + \frac{n}{2}\bar{x}_n^2 - \frac{1}{2}(3+\log(n-1)).
\end{aligned}
$$

Since $c \le 1/2$, substituting this into (7) yields

$$
\begin{aligned}
\log \mathcal{K}_n &\ge c\Big(\frac{1}{2}-c\Big)\sum_{i=1}^n \bar{x}_{i-1}^2 + c\frac{n}{2}\bar{x}_n^2 - \frac{c}{2}(3+\log(n-1)) \\
&\ge c\frac{n}{2}\bar{x}_n^2 - \frac{c}{2}(3+\log(n-1)) \\
&\ge \frac{c}{2}(n\bar{x}_n^2 - \log n) - \frac{3}{2}c \\
&= \frac{c}{2}\log n\left(\frac{n\bar{x}_n^2}{\log n} - 1\right) - \frac{3}{2}c.
\end{aligned}
$$

Now if $\limsup_n \sqrt{n}|\bar{x}_n|/\sqrt{\log n} > 1$, then $\limsup_n \log \mathcal{K}_n = +\infty$ because $\log n \uparrow \infty$. This proves the theorem.

Remark 3.1 *In the framework of the conventional measure theoretic probability, a strong law of large numbers analogous to Theorem 3.1 can be proved using Azuma-Hoeffding-Bennett inequality (Appendix A.7 of Vovk, Gammerman and Shafer (2005), Section 2.4 of Dembo and Zeitouni (1998), Azuma (1967), Hoeffding (1963), Bennett (1962)). Let* $X_1, X_2, \dots$ *be a sequence of martingale differences such that* $|X_n| \le 1$, $\forall n$. *Then for any* $\epsilon > 0$

$$P(|\bar{X}_n| \ge \epsilon) \le 2\exp(-n\epsilon^2/2).$$

Fix an arbitrary $\alpha > 1/2$. *Then for any* $\epsilon > 0$

$$\sum_n P(|\bar{X}_n| \ge \epsilon(\log n)^\alpha/\sqrt{n}) \le \sum_n \exp(-\frac{\epsilon^2}{2}(\log n)^{2\alpha}) \ < \infty.$$

Therefore by Borel-Cantelli $\sqrt{n}|\bar{X}_n|/(\log n)^\alpha \to 0$ *almost surely. Actually Theorem 1 of Azuma (1967) states the following stronger result*

$$\limsup_{n\to\infty} \frac{\sqrt{n}\bar{x}_n}{\sqrt{\log n}} \le \sqrt{2} \qquad \text{a.s.} \tag{8}$$

Although our Theorem 3.1 is better in the constant factor of $\sqrt{2}$, Azuma's result (8) and our result (2) are virtually the same. However we want to emphasize that our game theoretic proof requires much less mathematical background than the measure theoretic proof.

4 Growth rate of Skeptic's capital process when SLLN is violated

In this section we study the behavior of the paths of the log capital process $\log \mathcal{K}_n$ in the case of the fair-coin game ($x_n = \pm 1$), when the strategy (1) with $c = 1$ is used. We are interested in the main term (i.e. $O(n)$ term) of $\log \mathcal{K}_n$ when Reality violates SLLN, namely, either $\lim_n \bar{x}_n \neq 0$ exists or

$$\liminf_n \bar{x}_n < \limsup_n \bar{x}_n.$$

In this section our interest is in the latter case. For $0 < p,\, q < 1$, let

$$D(p\|q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$$

denote the Kullback divergence between p and q. Note that $D(p\|q) \geq 0$ and the equality holds if and only if $p = q$. We will prove the following remarkable fact.

Theorem 4.1 *Consider the fair-coin game, where Skeptic uses the strategy (1) with $c = 1$. For each path $\xi = x_1 x_2 \ldots$ of Reality's moves such that*

$$0 < \liminf_n \bar{x}_n < \limsup_n \bar{x}_n < 1,$$

we have

$$\frac{1}{n} \log \mathcal{K}_n - D\Big(\frac{1+\bar{x}_n}{2} \,\Big\|\, \frac{1}{2}\Big) \to 0 \qquad (n \to \infty).$$

We give a proof of this theorem in Section 4.2. In Section 4.1 we give a preliminary consideration of $\log \mathcal{K}_n$ when its increment can be approximated by an integral, which naturally leads to the Kullback divergence.

At this point we need to make the following technical remark. If Skeptic uses (1) with $c = 1$ from the very beginning, he becomes bankrupt at time n ($\mathcal{K}_n = 0$) if

$$x_1 = \cdots = x_{n-1} = 1, x_n = -1 \quad \text{or} \quad x_1 = \cdots = x_{n-1} = -1, x_n = 1.$$

Therefore, when we say that Skeptic uses $c = 1$, we actually mean that Skeptic sets M_n as

$$M_n = \begin{cases} \bar{x}_{n-1} \mathcal{K}_{n-1} & \text{if} \quad |\bar{x}_{n-1}| < 1 \\ c_0 \bar{x}_{n-1} \mathcal{K}_{n-1} & \text{if} \quad |\bar{x}_{n-1}| = 1, \end{cases}$$

where $0 < c_0 < 1$. For a similar reason, in Theorem 4.1 we are avoiding the case that $0 = \liminf_n \bar{x}_n$ or $1 = \limsup_n \bar{x}_n$. In fact these cases and the case that $\lim_n \bar{x}_n \neq 0$ exists should be simpler than our case considered in Theorem 4.1

4.1 Integral approximation to the increment of monotone log capital process

In this section, we give a preliminary result on the increment of $\log \mathcal{K}_n$, when $\bar{x}_n$ increases or decreases monotonically, i.e., when there are successive 1's or -1's in x_n's. The reason to consider this case is as follows. Write

$$l_\infty = \liminf_n \bar{x}_n \ < \ u_\infty = \limsup_n \bar{x}_n.$$

Then $\bar{x}_n$ wanders between l_∞ and u_∞ infinitely often. $\bar{x}_n$ moves the fastest when $\bar{x}_n$ oscillates between l_∞ and u_∞ monotonically.

Fix a sufficiently large n and let $v = \bar{x}_n$, $-1 < v < 1$. We consider the increment $\log \mathcal{K}_{n+N} - \log \mathcal{K}_n$, when $x_{n+1} = \cdots = x_{n+N} = 1$ or $x_{n+1} = \cdots = x_{n+N} = -1$.

First consider the case that $x_{n+1} = \cdots = x_{n+N} = 1$. Then $\bar{x}_{n+k}$ is written as

$$\bar{x}_{n+k} = \frac{nv+k}{n+k}, \quad k = 0, 1, \dots, N-1,$$

and

$$\log \mathcal{K}_{n+N} - \log \mathcal{K}_n = \sum_{k=0}^{N-1} \log\left(1 + \frac{nv+k}{n+k}\right) = n \sum_{k=0}^{N-1} \log\left(1 + \frac{v+k/n}{1+k/n}\right)\frac{1}{n} \ . \tag{9}$$

Note that $\log(1 + (v+x)/(1+x))$ is monotone increasing in $x \geq 0$. Then the summation on the right-hand side of (9) is approximated as

$$\int_0^{N/n} \log\left(1 + \frac{v+x}{1+x}\right)dx < \sum_{k=0}^{N-1} \log\left(1 + \frac{v+k/n}{1+k/n}\right)\frac{1}{n} < \int_{1/n}^{(N+1)/n} \log\left(1 + \frac{v+x}{1+x}\right)dx.$$

Also note that the difference between the right-hand side and the left-hand side is bounded from above by C/n, where C can be chosen independently of v and N. Therefore

$$\Big| \sum_{k=0}^{N-1} \log\left(1 + \frac{nv+k}{n+k}\right) - n \int_0^{N/n} \log\left(1 + \frac{v+x}{1+x}\right)dx \Big| < C.$$

The definite integral can be explicitly written as follows. Using

$$\int_0^c \log(a+bx)dx = \frac{1}{b}[(a+bc)\log(a+bc) - a\log a] - c,$$

we have

$$\begin{aligned}\int_0^{N/n} \log\left(1 + \frac{v+x}{1+x}\right)dx = &\frac{1}{2}[(1+v+2c)\log(1+v+2c) - (1+v)\log(1+v)] \\ &- (1+c)\log(1+c),\end{aligned} \tag{10}$$

where $c = N/n$.

Next consider the case $x_{n+1} = \cdots = x_{n+N} = -1$. Then $\bar{x}_{n+k}$ is written as

$$\bar{x}_{n+k} = \frac{nv-k}{n+k}, \quad k = 0, 1, \ldots, N-1,$$

and

$$\log \mathcal{K}_{n+N} - \log \mathcal{K}_n = \sum_{k=0}^{N-1} \log\Big(1 - \frac{nv-k}{n+k}\Big).$$

The similar consideration as above yields

$$\Big| \sum_{k=0}^{N-1} \log\Big(1 - \frac{nv-k}{n+k}\Big) - n\int_0^{N/n} \log\Big(1 - \frac{v-x}{1+x}\Big)dx\Big| < C.$$

The definite integral can be explicitly written as

$$\begin{aligned}\int_0^{N/n} \log\Big(1 - \frac{v-x}{1+x}\Big)dx &= \frac{1}{2}[(1-v+2c)\log(1-v+2c) - (1-v)\log(1-v)] \\ &\quad - (1+c)\log(1+c), \qquad (11)\end{aligned}$$

where $c = N/n$.

Consider the case that $\bar{x}_n$ increases from $l \simeq l_\infty$ at time n to $u \simeq u_\infty$ at time $n+N$. Then

$$u = \bar{x}_{n+N} = \frac{nl+N}{n+N} = \frac{l+c}{1+c},$$

and

$$1 + c = \frac{1-l}{1-u}, \quad 1 + l + 2c = \frac{(1-l)(1+u)}{1-u}\ . \qquad (12)$$

It should be noted $n + N = (1 + c)n$ is of the order $O(n)$, i.e., it takes $O(n)$ steps even if $\bar{x}_{n+k}$ moves from l to u monotonically. By using (10) and (12), we have except for $O(1)$ terms

$$\begin{aligned}\log \mathcal{K}_{n+N} - \log \mathcal{K}_n &= n\Big[\frac{(1-l)(1+u)}{2(1-u)}\log(1+u) + \frac{1-l}{2}\log(1-u) \\ &\qquad - \frac{1+l}{2}\log(1+l) - \frac{1-l}{2}\log(1-l)\Big] \\ &= \frac{n(1-l)}{1-u}\Big[\frac{1+u}{2}\log(1+u) + \frac{1-u}{2}\log(1-u)\Big] \\ &\qquad - n\Big[\frac{1+l}{2}\log(1+l) + \frac{1-l}{2}\log(1-l)\Big] \\ &= (n+N)D\Big(\frac{1+u}{2} \Big\| \frac{1}{2}\Big) - nD\Big(\frac{1+l}{2} \Big\| \frac{1}{2}\Big). \qquad (13)\end{aligned}$$

Next consider the case that $\bar{x}_n$ decreases from u to l. Then

$$l = \bar{x}_{n+N} = \frac{nu-N}{n+N} = \frac{u-c}{1+c},$$

and

$$1+c=\frac{1+u}{1+l},\quad 1-u+2c=\frac{(1+u)(1-l)}{1+l}\ .$$

By using (11) and the above, we have except for $O(1)$ terms

$$\begin{aligned}\log\mathcal{K}_{n+N}-\log\mathcal{K}_n &= n\Big[\frac{1+u}{2}\log(1+l)+\frac{(1+u)(1-l)}{2(1+l)}\log(1-l)\\ &\qquad -\frac{1+u}{2}\log(1+u)-\frac{1-u}{2}\log(1-u)\Big]\\ &= \frac{n(1+u)}{1+l}\Big[\frac{1+l}{2}\log(1+l)+\frac{1-l}{2}\log(1-l)\Big]\\ &\qquad -n\Big[\frac{1+u}{2}\log(1+u)+\frac{1-u}{2}\log(1-u)\Big]\\ &= (n+N)D\Big(\frac{1+l}{2}\ \Big\|\ \frac{1}{2}\Big)-nD\Big(\frac{1+u}{2}\ \Big\|\ \frac{1}{2}\Big).\end{aligned}\tag{14}$$

It should be noted that both (13) and (14) can be written as

$$\log\mathcal{K}_{n+N}-\log\mathcal{K}_n=(n+N)D\Big(\frac{1+\bar{x}_{n+N}}{2}\ \Big\|\ \frac{1}{2}\Big)-nD\Big(\frac{1+\bar{x}_n}{2}\ \Big\|\ \frac{1}{2}\Big)+O(1).\tag{15}$$

4.2 Increment of log capital process for an arbitrary path

In this section, we first show that the main term of (15) remains the same for any path not necessarily monotonic between $\bar{x}_n$ and $\bar{x}_{n+N}$. Theorem 4.1 is a simple consequence of the following result.

Theorem 4.2 *For sufficiently large n, when $\bar{x}_n$ moves from $\bar{x}_n=l$ to $\bar{x}_{n+N}=u$, $l\neq u$, through any path,*

$$\log\mathcal{K}_{n+N}-\log\mathcal{K}_n=(n+N)D\Big(\frac{1+u}{2}\ \Big\|\ \frac{1}{2}\Big)-nD\Big(\frac{1+l}{2}\ \Big\|\ \frac{1}{2}\Big)+o(N).\tag{16}$$

Proof: For notational simplicity we assume $l<u$ and prove this theorem. The case $l>u$ can be proved similarly. As discussed in the previous subsection, it takes at least $O(n)$ steps for moving from $\bar{x}_n=l$ to $\bar{x}_{n+N}=u$. We divide the vertical $\bar{x}_n$-axis through l and u by the horizontal lines with equal distance $1/\sqrt{n}$. Since $\bar{x}_n$ can wander below l or above u, we extend the grid of equidistant horizontal lines to $[0,l]$ and $[u,1]$. Next, we denote the times when $\bar{x}_m, n\le m\le N$, hits at different horizontal lines (admitting $\pm 1/n$ rounding error) as $n=n_0<n_1<n_2<\cdots<n_M=n+N$, successively. Then $\bar{x}_m, n_j\le m<n_{j+1}$, moves in the interval $(\bar{x}_{n_j}-1/\sqrt{n}, \bar{x}_{n_j}+1/\sqrt{n})$, i.e.,

$$|\bar{x}_m-\bar{x}_{n_j}|<\frac{1}{\sqrt{n}},\quad n_j\le m<n_{j+1},$$

and $\bar{x}_{n_{j+1}}=\bar{x}_{n_j}\pm 1/\sqrt{n}$ (admitting $\pm 1/n$ rounding error).

We decompose the increment of the log capital as

$$\log \mathcal{K}_{n+N} - \log \mathcal{K}_n = \sum_{h=0}^{M-1} (\log \mathcal{K}_{n_{h+1}} - \log \mathcal{K}_{n_h}),$$

and show that for each h except for small orders

$$\log \mathcal{K}_{n_{h+1}} - \log \mathcal{K}_{n_h} \simeq n_{h+1} D\Big(\frac{1+\bar{x}_{n_{h+1}}}{2} \,\Big\|\, \frac{1}{2}\Big) - n_h D\Big(\frac{1+\bar{x}_{n_h}}{2} \,\Big\|\, \frac{1}{2}\Big).$$

Then (16) follows by summing up these terms.

For notational convenience, we put $\epsilon_n = \pm 1/\sqrt{n}$, and without loss of generality consider the first step $h=0$ with $\bar{x}_{n_0} = l$ and $\bar{x}_{n_1} = l + \epsilon_n$.
From

$$\log(1+l+\epsilon_n) = \log(1+l) + \frac{\epsilon_n}{1+l} + O(\epsilon_n^2),$$
$$\log(1-l-\epsilon_n) = \log(1-l) - \frac{\epsilon_n}{1-l} + O(\epsilon_n^2),$$

we have

$$\begin{aligned} D\Big(\frac{1+l+\epsilon_n}{2} \,\Big\|\, \frac{1}{2}\Big) &= \frac{1+l+\epsilon_n}{2}\log(1+l+\epsilon_n) + \frac{1-l-\epsilon_n}{2}\log(1-l-\epsilon_n) \\ &= D\Big(\frac{1+l}{2} \,\Big\|\, \frac{1}{2}\Big) + \frac{\epsilon_n}{2}\log(1+l) - \frac{\epsilon_n}{2}\log(1-l) + O(\epsilon_n^2). \end{aligned}$$

Therefore

$$\begin{aligned} & n_1 D\Big(\frac{1+\bar{x}_{n_1}}{2} \,\Big\|\, \frac{1}{2}\Big) - n_0 D\Big(\frac{1+\bar{x}_{n_0}}{2} \,\Big\|\, \frac{1}{2}\Big) \\ &\quad = (n_1 - n_0) D\Big(\frac{1+l}{2} \,\Big\|\, \frac{1}{2}\Big) + n_1\Big[\frac{\epsilon_n}{2}\log(1+l) - \frac{\epsilon_n}{2}\log(1-l)\Big] + O(n_1\epsilon_n^2) \\ &\quad = (n_1 - n_0)\Big[\frac{1+l}{2}\log(1+l) + \frac{1-l}{2}\log(1-l)\Big] \\ &\qquad\qquad + n_1\Big[\frac{\epsilon_n}{2}\log(1+l) - \frac{\epsilon_n}{2}\log(1-l)\Big] + O(n_1\epsilon_n^2) \\ &\quad = \frac{1}{2}[(n_1-n_0)(1+l) + n_1\epsilon_n]\log(1+l) \\ &\qquad\qquad + \frac{1}{2}[(n_1-n_0)(1-l) - n_1\epsilon_n]\log(1-l) + O(n_1\epsilon_n^2). \end{aligned}$$

When there are n_u times of $x_m = +1$'s and n_d times of $x_m = -1$'s in the time interval $n_1 - n_0$,

$$n_1 - n_0 = n_u + n_d,$$
$$n_1(l+\epsilon_n) - n_0 l = n_u - n_d.$$

Therefore

$$\frac{1}{2}[(n_1-n_0)(1+l)+n_1\epsilon_n]=n_u,\quad \frac{1}{2}[(n_1-n_0)(1-l)-n_1\epsilon_n]=n_d,$$

and

$$n_1 D\Big(\frac{1+\bar{x}_{n_1}}{2}\ \Big\|\ \frac{1}{2}\Big)-n_0 D\Big(\frac{1+\bar{x}_{n_0}}{2}\ \Big\|\ \frac{1}{2}\Big)=n_u\log(1+l)+n_d\log(1-l)+O(n_1\epsilon_n^2).$$

Since $\bar{x}_m$, $n_0\le m<n_1$, moves in the interval $(l-1/\sqrt{n}, l+1/\sqrt{n})$, it follows that

$$\begin{aligned} n_u\log(1+l)+n_d\log(1-l) &= \sum_{m=n_0+1}^{n_1}\log(1+\bar{x}_{m-1}x_m)+O(n_1\epsilon_n)\\ &= \log\mathcal{K}_{n_1}-\log\mathcal{K}_{n_0}+O(n_1\epsilon_n), \end{aligned}$$

and thus

$$n_1 D\Big(\frac{1+\bar{x}_{n_1}}{2}\ \Big\|\ \frac{1}{2}\Big)-n_0 D\Big(\frac{1+\bar{x}_{n_0}}{2}\ \Big\|\ \frac{1}{2}\Big)=\log\mathcal{K}_{n_1}-\log\mathcal{K}_{n_0}+O(n_1\epsilon_n)+O(n_1\epsilon_n^2).$$

By summing up this relation from $h=0$ to $M-1$, we have (16) as desired. This completes the proof. □

We consider the consequence of this theorem when $\bar{x}_n$ wanders from $l\simeq l_\infty$ to $u\simeq u_\infty$ and then back to l, in a cycle. Suppose that $\bar{x}_n=l$ goes up to $\bar{x}_{n+N_u}=u$ and goes down to $\bar{x}_{n+N_l}=l$. Then by adding $\log\mathcal{K}_{n+N_u}-\log\mathcal{K}_n$ and $\log\mathcal{K}_{n+N_l}-\log\mathcal{K}_{n+N_u}$, we obtain

$$\log\mathcal{K}_{n+N_l}-\log\mathcal{K}_n=N_l D\Big(\frac{1+l}{2}\ \Big\|\ \frac{1}{2}\Big)+o(N_l). \tag{17}$$

Similarly when $\bar{x}_n=u$ goes down to $\bar{x}_{n+N_l}=l$ and goes up to $\bar{x}_{n+N_u}=u$, we have

$$\log\mathcal{K}_{n+N_u}-\log\mathcal{K}_n=N_u D\Big(\frac{1+u}{2}\ \Big\|\ \frac{1}{2}\Big)+o(N_u). \tag{18}$$

From (17) and (18), we can say that if Reality violates SLLN, log capital of the strategy (1) with $c=1$ increases at the rate $D(\frac{1+l}{2}\|\frac{1}{2})$ whenever $\bar{x}_n$ comes to l_∞, and increases at the rate $D(\frac{1+u}{2}\|\frac{1}{2})$ whenever $\bar{x}_n$ comes to u_∞.

Theorem 4.1 is now a simple consequence of Theorem 4.2. Note that for sufficiently large $n_1<n_2$, (16) is written as

$$\log\mathcal{K}_{n_2}-\log\mathcal{K}_{n_1}=n_2 D\Big(\frac{1+\bar{x}_{n_2}}{2}\ \Big\|\ \frac{1}{2}\Big)-n_1 D\Big(\frac{1+\bar{x}_{n_1}}{2}\ \Big\|\ \frac{1}{2}\Big)+o(n_2). \tag{19}$$

Under the assumption $l_\infty<u_\infty$, for sufficiently large m, $\bar{x}_m$ wanders between l_∞ and u_∞. Starting from a sufficiently large (but fixed) n_0 with $\bar{x}_{n_0}\simeq l_\infty$, consider successive time points when $\bar{x}_m$ returns to l_∞ before finally reaching $\bar{x}_n$. Adding the increments in (17) and finally the increment in (19) we have

$$\log\mathcal{K}_n-\log\mathcal{K}_{n_0}=nD\Big(\frac{1+\bar{x}_n}{2}\ \Big\|\ \frac{1}{2}\Big)-n_0 D\Big(\frac{1+l}{2}\ \Big\|\ \frac{1}{2}\Big)+o(n).$$

Theorem 4.1 is proved by dividing both sides by n and letting $n\to\infty$.

A Summary of mixture ϵ-strategy in Chapter 3 of Shafer and Vovk (2001)

Here we summarize the mixture ϵ-strategy in Chapter 3 of Shafer and Vovk (2001) for the bounded forecasting game in such a way that its resource requirement (computational time and memory) becomes explicit.

For a single fixed ϵ-strategy $\mathcal{P}^\epsilon$, which sets $M_n = \epsilon \mathcal{K}_{n-1}$, the capital process is given as $\mathcal{K}_n^{\mathcal{P}^\epsilon}(x_1 \dots x_n) = \prod_{i=1}^n (1 + \epsilon x_i)$. Shafer and Vovk combine $\mathcal{P}^\epsilon$ for many values of ϵ. Let $1/2 \geq \epsilon_1 > \epsilon_2 > \cdots > 0$ be a sequence of positive numbers converging to 0. Write $\epsilon_{-k} = -\epsilon_k$, $k = 1, 2, \dots$. Furthermore let $\{p_i\}_{i=0,\pm 1,\pm 2,\dots}$ be a probability distribution on the set of integers $\mathbb{Z}$, such that $p_0 = 0$ and $p_{-k} = p_k$ (symmetric). Actually p_k is the initial amount (out of 1 dollar) put into the account with the strategy $\mathcal{P}^{\epsilon_k}$, $k = \pm 1, \pm 2, \dots$. For example we could take

$$\epsilon_k = \frac{1}{2^{|k|}}, \quad p_k = \frac{1}{2^{|k|+1}}, \ k = \pm 1, \pm 2, \dots, \tag{20}$$

as is done in Chapter 3 of Shafer and Vovk (2001). However it is more convenient here to leave ϵ_k and p_k to be general. Then the mixture strategy weakly forcing SLLN is given by the weighted average of the strategies $\mathcal{P}^{\epsilon_k}$, $k = \pm 1, \pm 2, \dots$, with respect to the weights $\{p_k\}_{k=0,\pm 1,\pm 2,\dots}$, namely

$$\mathcal{P}^* = \sum_{k=-\infty}^{\infty} p_k \mathcal{P}^{\epsilon_k}.$$

Consider the value M_n^* of $\mathcal{P}^*$. It is written as

$$M_n^* = \sum_{k=-\infty}^{\infty} p_k \epsilon_k \prod_{i=1}^{n-1} (1 + \epsilon_k x_i).$$

Now introduce the elementary symmetric functions $e_{n,0}, e_{n,1}, \dots, e_{n,n}$ of the numbers $x_1, \dots, x_n$ as

$$e_{n,0} = 1, \ e_{n,1} = \sum_{i=1}^n x_i, \ e_{n,2} = \sum_{1 \leq i < j \leq n} x_i x_j, \ \dots, \ e_{n,n} = \prod_{i=1}^n x_i.$$

Recall that the set of values $\{x_1, \dots, x_n\}$ and the set of values $\{e_{n,1}, \dots, e_{n,n}\}$ are in one-to-one relationship. Therefore computing all the values of the elementary symmetric functions of $\{x_1, \dots, x_n\}$ is computationally equivalent to keep all the values of $\{x_1, \dots, x_n\}$. Using elementary symmetric functions, M_n^* is written as

$$M_n^* = \sum_{k=-\infty}^{\infty} p_k \sum_{i=0}^{n-1} \epsilon_k^{i+1} e_{n-1,i} = \sum_{i=0}^{n-1} e_{n-1,i} \sum_{k=-\infty}^{\infty} p_k \epsilon_k^{i+1}. \tag{21}$$

Let F denote the discrete probability distribution on $[-1/2, 1/2]$ with

$$F(\{\epsilon_k\}) = p_k, \ k = \pm 1, \pm 2, \dots,$$

and let

$$\mu_m = E_F(X^m) = \sum_{k=-\infty}^{\infty} p_k \epsilon_k^m = \int_{-1/2}^{1/2} x^m F(dx)$$

denote the m-th moment of F. Then (21) is written as

$$M_n^* = \sum_{i=0}^{n-1} \mu_{i+1} e_{n-1,i}. \tag{22}$$

At this point it becomes clear that we can take an arbitrary probability distribution F on $[-1/2, 1/2]$ provided that F is not degenerate at 0 and 0 is the point of support, i.e., for all $\epsilon > 0$

$$F(\epsilon) - F(-\epsilon) > 0.$$

M_n^* in (22) for such an F is a strategy weakly forcing SLLN. Now the simplest F seems to be the uniform distribution on $[-1/2, 1/2]$. Actually (20) is in a sense close to the uniform distribution. Then

$$\mu_m = \int_{-1/2}^{1/2} x^m dx = \begin{cases} \frac{1}{m+1}\frac{1}{2^m} & m : \text{even} \\ 0 & m : \text{odd}. \end{cases}$$

In this case there is virtually no computational resource is needed to compute μ_m, since it is explicitly given. Furthermore the strategy weakly forcing SLLN based on the uniform distribution is given by

$$M_n^* = \sum_{\substack{i=0 \\ i:\text{odd}}}^{n-1} \frac{1}{i+2}\frac{1}{2^{i+1}} e_{n-1,i}.$$

In order to compute the M_n^* we need all the values of the elementary symmetric functions $e_{n-1,i}$, $i = 1, \dots, n-1$, which is equivalent to keeping all the values of $\{x_1, \dots, x_{n-1}\}$ as discussed above. Therefore the mixture ϵ-strategy needs a memory of size proportional to n at time n, whereas our strategy in (1) only needs to keep track of the values of $\bar{x}_{n-1}$ and $\mathcal{K}_{n-1}$.

References

[1] Kazuoki Azuma. Weighted sums of certain dependent random variables. *Tôhoku Math. Journ.*, **19**, 357–367. 1967.

[2] George Bennett. Probability inequalities for the sum of independent random variables. *Journal of American Statistical Association*, **57**, 33–45, 1962.

[3] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications.* 2nd ed., Springer, New York, 1998.

[4] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, **58**, 13–30, 1963.

[5] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!*. Wiley, New York, 2001.

[6] Kei Takeuchi. *Kake no suuri to kinyu kogaku* (Mathematics of betting and financial engineering). Saiensusha, Tokyo, 2004. (in Japanese)

[7] Kei Takeuchi. On strategies in favourable betting games. 2004. Unpublished manuscript.

[8] Vladimir Vovk, Alex Gammerman, and Glenn Shafer. *Algorithmic Learning in a Random World*, Springer, New York, 2005.

[9] Vladimir Vovk, Akimichi Takemura, and Glenn Shafer. Defensive forecasting. in *Proceedings of the tenth international workshop on artificial intelligence and statistics.* R.G.Cowell and Z.Ghahramani editors, 365–372, 2005. (Available electronically at http://www.gatsby.ucl.ac.uk/aistats/)