MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Multistep Bayesian Strategy in Coin-tossing Games and Its Application to Asset Trading Games in Continuous Time

Kei TAKEUCHI, Masayuki KUMON and Akimichi TAKEMURA

METR 2008–08 February 2008

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.keisu.t.u-tokyo.ac.jp/research/techrep/index.html

Multistep Bayesian strategy in coin-tossing games and its application to asset trading games in continuous time

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

February, 2008

Abstract

We study multistep Bayesian betting strategies in coin-tossing games in the framework of game-theoretic probability of Shafer and Vovk (2001). We show that by a countable mixture of these strategies, a gambler or an investor can exploit arbitrary patterns of deviations of nature's moves from independent Bernoulli trials. We then apply our scheme to asset trading games in continuous time and derive the exponential growth rate of the investor's capital when the variation exponent of the asset price path deviates from two.

Keywords and phrases: Beta-binomial distribution, Hölder exponent, Kullback divergence, randomness, risk neutral probability, universal prior.

1 Introduction

The field of game-theoretic probability and finance established by Shafer and Vovk [12] has been rapidly developing in many directions. The present authors have been contributing to this exciting new field by focusing mainly on explicit strategies of the gambler and the growth rate of his capital ([14], [15], [6], [7], [8], [13], [5], [16]). Following the terminology of Shafer and Vovk [12], we refer to the gambler as Skeptic or Investor (Section 4) and refer to nature as Reality or Market (Section 4).

In this paper we extend the results of [7] and [16] by considering multistep Bayesian strategies. In [7] we considered a class of Bayesian strategies for Skeptic in coin-tossing games, strategies that were based only on the past average of Reality's moves. We proved the important fact that if Skeptic uses a Bayesian strategy and Reality violates the strong law of large numbers (SLLN), then the exponential growth rate of the Skeptic's capital process is very accurately 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. Furthermore in [16] we applied Bayesian strategies for coin-tossing games to asset trading games in continuous time. If we discretize a continuous-time game by an equi-spaced grid in the state space (i.e. the vertical axis), then the continuous time game can be approximated by an embedded coin-tossing game. Thus the results of discrete time coin-tossing game can be applied to continuous-time games. In particular we gave a proof of "$\sqrt{\mathrm{d}t}$-effect", i.e., Investor can force Market to choose a price path with variation exponent equal to two, within an arbitrary small constant.

More generally, discretization of continuous-time game in [16] is based on the requirement that Investor choose a countable number of discrete stopping times against a continuous path chosen by Market. This approach allows us to formulate and study continuous-time games in game-theoretic probability within the conventional theory of analysis, whereas in the book by Shafer and Vovk continuous-time games were formulated as limits of discrete time games using nonstandard analysis. Vovk has taken up this formulation and is currently rapidly developing it in [17], [18], [19]. By these works of Vovk it has now become clear that many measure-theoretic results on continuous-time stochastic processes can be more directly derived in the framework of game-theoretic probability. It should be emphasized that the game-theoretic approach is advantageous because no probabilistic assumptions on the paths are imposed a priori. Instead, a stochastic behavior of Market results from the protocol of the game. This has a far-reaching conceptual implications for the emergence of probability.

In the coin-tossing games Reality may deviate from independent Bernoulli trials in a subtle way without violating the strong law of large numbers. For example, Reality could choose a deterministic sequence where heads and tails alternate. Then SLLN holds for this kind of paths and they can not be prevented by Bayesian strategies in [7]. We could use "contrarian" strategies ([5]) based on the past average of Reality's moves for preventing this kind of paths. However a more natural approach is to model autocorrelations between successive moves of Reality. Note that the deterministic sequence of alternating heads and tails can be regarded as a sequence with the first-order autocorrelation of -1. These considerations naturally lead to multistep Bayesian strategies of the present paper. Then by a countable mixture of these strategies Skeptic can detect and exploit arbitrary patterns of deviations from the sequence of independent Bernoulli trials.

Our results have close relations to the ones in many fields. Universal source coding has been extensively studied in information theory (e.g. Han and Kobayashi [4]) and the equivalence of source coding and betting is discussed in Cover and Thomas [1]. Various notions of randomness have been studied from the viewpoint of Kolmogorov complexity (e.g. Li and Vitányi [10], Lambalgen [9]) and there exists an extensive literature on algorithmic theory of randomness. See two forthcoming books on algorithmic randomness by Downey and Hirschfeldt [2] and by Nies [11]. We discuss these relations in Section 5.

The organization of this paper is as follows. In Section 2 we set up notations for the coin-tossing game and give some preliminary results. In Section 3 we consider two types of Reality's moves which suggest deviations from independent Bernoulli trials. The first is a block type pattern and the second is a Markovian type pattern. We construct Skeptic's Bayesian strategies which can exploit these non-randomnesses. We show that it is asymptotically always advantageous to exploit higher order patterns in Reality's moves. In Section 4 we consider asset trading games in continuous time and investigate the consequences of high-frequency block type strategies and Markovian strategies. We derive the exponential growth rate of Investor's capital process when Market chooses a path with variation exponent not equal to two. Finally in Section 5 we discuss some aspects of our results and their implications for related fields.

2 Preliminaries on coin-tossing games and Bayesian strategies

In this section we summarize preliminary results on coin-tossing games and capital process of Bayesian strategies ([12], [7]). We also discuss a one-to-one correspondence between the set of probability distributions on the set of Reality's paths and the set of Skeptic's strategies. Finally we note the convexity of the Kullback divergence with respect to its first argument.

In this paper we consider the coin-tossing game in the following form. In the protocol the success probability $0 < \rho < 1$ is given.

Coin-Tossing Game
Protocol:

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots$:
  Skeptic announces $M_n \in \mathbb{R}$.
  Reality announces $x_n \in \{0, 1\}$.
  $\mathcal{K}_n = \mathcal{K}_{n-1} + M_n(x_n - \rho)$.
END FOR

If we write $M_n = M_n^1 - M_n^0$, then $M_n(x_n - \rho)$ can be rewritten as

$$M_n(x_n - \rho) = M_n^1(x_n - \rho) + M_n^0\big((1 - x_n) - (1 - \rho)\big). \tag{1}$$

In this case we say that Skeptic bets M_n^1 on $x_n = 1$ and M_n^0 on $x_n = 0$. Although (1) is a redundant expression, generalizations to multistep protocols in the next section can be more transparently understood in this form. A path $\xi = x_1 x_2 \ldots$ is an infinite sequence of

Reality's moves and the sample space $\Xi = \{\xi\} = \{0,1\}^\infty$ is the set of paths. $\xi^n = x_1 \dots x_n$ denotes the partial path of Reality's moves up to round n. Throughout this paper we use the notation

$$s_n = x_1 + \cdots + x_n, \qquad \bar{x}_n = \frac{s_n}{n}.$$

Skeptic's strategy $\mathcal{P}$ is a set of functions $\mathcal{P} : \xi^{n-1} \mapsto M_n$ which determines Skeptic's move M_n at round n based on Reality's moves up to the previous round $\xi^{n-1} = x_1 \dots x_{n-1}$. Given a strategy $\mathcal{P}$

$$\mathcal{K}_n^{\mathcal{P}}(\xi) = \mathcal{K}_0 + \sum_{i=1}^{n} M_i(\xi^{i-1})(x_i - \rho)$$

denotes Skeptic's capital process when he uses $\mathcal{P}$, starting with the initial capital of $\mathcal{K}_0 = 1$. Following the terminology in [17] we call $\mathcal{P}$ *prudent* if $\mathcal{K}_n^{\mathcal{P}}(\xi) \geq 0$ for all ξ and n, i.e., Skeptic's capital is never negative irrespective of the moves of Reality. We also say that Skeptic observes his collateral duty if he uses a prudent strategy. In this paper we require that Skeptic's strategies are prudent.

We consider probability distributions on the set of paths $\Xi = \{0,1\}^\infty$. In our framework a probability distribution Q on Ξ is just a collection of consistent discrete probability distributions $Q = \{Q_n; n \geq 1\}$, where Q_n is a discrete probability distribution on $\Xi_n = \{0,1\}^n$ satisfying the consistency condition

$$Q_n(\xi^n) = Q_{n+1}(\xi^n 0) + Q_{n+1}(\xi^n 1), \qquad \forall n, \forall \xi^n. \tag{2}$$

Note that we do not need measure-theoretic extension of Q to a probability measure on a σ-field of Ξ. Under the distribution Q, the conditional probability of $x_n = 1$ given $\xi^{n-1} = x_1 \dots x_{n-1}$ with $Q_{n-1}(\xi^{n-1}) > 0$ is written as

$$p_n^Q = p_n^Q(\xi^{n-1}) = \frac{Q_n(\xi^{n-1}1)}{Q_{n-1}(\xi^{n-1})}. \tag{3}$$

If $Q_{n-1}(\xi^{n-1}) = 0$, then p_n^Q is not defined. We call the probability distribution of i.i.d. Bernoulli trials with success probability ρ the *risk neutral measure* of the coin-tossing game.

Given a probability distribution Q, define a strategy $\mathcal{P} = \mathcal{P}_Q$ by

$$\mathcal{P}_Q : \xi^{n-1} \mapsto M_n = \mathcal{K}_{n-1} \frac{p_n^Q - \rho}{\rho(1-\rho)}. \tag{4}$$

A motivation of this definition is given in Appendix. If we write $M_n = M_n^1 - M_n^0$ as in (1), then M_n^1, M_n^0 are given as

$$M_n^1 = \mathcal{K}_{n-1} \frac{p_n^Q}{\rho}, \quad M_n^0 = \mathcal{K}_{n-1} \frac{1 - p_n^Q}{1 - \rho}. \tag{5}$$

The capital process of $\mathcal{P}_Q$ is explicitly written as follows ([7, Theorem 4.1]).

$$\mathcal{K}_n^{\mathcal{P}_Q}(\xi^n) = \frac{Q(\xi^n)}{\rho^{s_n}(1-\rho)^{n-s_n}}. \tag{6}$$

This is the likelihood ratio of Q to the risk neutral measure at the realized path ξ^n up to round n. In Appendix, we establish a one-to-one correspondence between the set of probability distributions on Ξ and the set of prudent strategies. Therefore capital process of a prudent strategy can always be expressed as (6).

In particular if we employ a beta-binomial distribution

$$Q(\xi^n) = \frac{\Gamma(a+b)\Gamma(a+s_n)\Gamma(b+n-s_n)}{\Gamma(a+b+n)\Gamma(a)\Gamma(b)} \qquad \text{with } \ p_n^Q = \frac{a+s_{n-1}}{a+b+n-1}, \tag{7}$$

where $a, b > 0$ are hyperparameters of the prior distribution, then by a simple application of Stirling's formula, the resulting capital process denoted by $\mathcal{K}_n^0$ behaves as

$$\log \mathcal{K}_n^0 = nD(\bar{x}_n \| \rho) - O(\log n), \tag{8}$$

where

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

is the Kullback divergence between two scalar probabilities p and q. Hence if $\bar{x}_n$ deviates from ρ, then $D(\bar{x}_n\|\rho)$ gives the average exponential growth rate of the capital process. We call $D(\bar{x}_n\|\rho)$ the main growth rate (or simply the growth rate) of the log capital.

In the subsequent sections we often use the convexity of Kullback divergence for probability vectors with respect to its first argument. Let $\boldsymbol{p} = \{p_j\}_{j=1}^k$ and $\boldsymbol{q} = \{q_j\}_{j=1}^k$ be probability vectors and let

$$D(\boldsymbol{p}\|\boldsymbol{q}) = \sum_{j=1}^k p_i \log \frac{p_j}{q_j}$$

denote the Kullback divergence between $\boldsymbol{p}$ and $\boldsymbol{q}$. Let $\boldsymbol{p}_i = \{p_{ij}\}_{j=1}^k$, $i = 1, 2$, be probability vectors and for $0 < \lambda < 1$ let $\bar{\boldsymbol{p}} = \lambda \boldsymbol{p}_1 + (1-\lambda)\boldsymbol{p}_2 = \{\bar{p}_j\}_{j=1}^k$. Then the following relation is easily obtained

$$\lambda D(\boldsymbol{p}_1\|\boldsymbol{q}) + (1-\lambda)D(\boldsymbol{p}_2\|\boldsymbol{q}) - D(\bar{\boldsymbol{p}}\|\boldsymbol{q}) = \lambda D(\boldsymbol{p}_1\|\bar{\boldsymbol{p}}) + (1-\lambda)D(\boldsymbol{p}_2\|\bar{\boldsymbol{p}}) \geq 0. \tag{9}$$

The left-hand side can also be written as

$$\lambda D(\boldsymbol{p}_1\|\boldsymbol{q}) + (1-\lambda)D(\boldsymbol{p}_2\|\boldsymbol{q}) - D(\bar{\boldsymbol{p}}\|\boldsymbol{q}) = \sum_{j=1}^k \bar{p}_j D(\frac{\lambda p_{1j}}{\bar{p}_j}\|\lambda) \geq 0.$$

3 Priors for higher order patterns and multistep strategies

As discussed in Section 1 there may be some deviating patterns from independent Bernoulli trials in the Reality's moves $x_1x_2\ldots$, although the path $\xi = x_1x_2\ldots$ satisfies SLLN $\lim_{n\to\infty} \bar{x}_n = \rho$. The strategy considered in [7] is based only on s_n and it can not exploit these patterns. Skeptic can increase his capital by strategies exploiting patterns not reflected in $\bar{x}_n$. In the following we investigate two types of such non-randomness or higher order patterns. The first is the block type pattern and the second is the Markovian type pattern. We give multistep Bayesian strategies which effectively exploit these patterns.

3.1 Block patterns

For clarity of presentation, we first consider sequence of pairs (i.e. blocks of length 2) and later generalize the results to blocks of arbitrary length.

Consider the sequence of pairs $(x_1x_2)(x_3x_4)\dots(x_{2n-1}x_{2n})$ of Reality's moves and denote the number of the pairs $(11),(10),(01),(00)$ among the first n blocks by m_n^{11}, m_n^{10}, m_n^{01}, m_n^{00}, respectively. If the sequence is random, i.e., the Reality's moves are i.i.d. Bernoulli trials with success probability ρ, then we will have

$$\lim_{n\to\infty}\frac{m_n^{ij}}{n}=\rho^{ij},\quad i,j=0,1,$$

where $\rho^{11}=(\rho)^2$, $\rho^{10}=\rho^{01}=\rho(1-\rho)$, $\rho^{00}=(1-\rho)^2$. We construct a strategy for which $\limsup_n \mathcal{K}_n=\infty$ whenever $\limsup_n |m_n^{ij}/n-\rho^{ij}|>0$ at least for one (i,j).

For this purpose, at the $(2n-1)$-th round ($n=1,2,\dots$) Skeptic chooses four amounts $M_n^{11},M_n^{10},M_n^{01},M_n^{00}$ and bet them on $(x_{2n-1}x_{2n})=(11),(10),(01),(00)$, respectively. Then we have

$$\mathcal{K}_{2n}=\mathcal{K}_{2n-2}+\sum_{i,j\in\{0,1\}}M_n^{ij}(z_n^{ij}-\rho^{ij}),$$

where

$$z_n^{11}=x_{2n-1}\times x_{2n}=\begin{cases}1, & \text{if } (x_{2n-1}x_{2n})=(11)\\ 0, & \text{otherwise}\end{cases}$$

and other z_n^{ij}, $i,j=0,1$, are defined similarly. Thus we define a derived capital process $\mathcal{K}_n^*=\mathcal{K}_{2n}$ with the protocol

$$\mathcal{K}_n^*=\mathcal{K}_{n-1}^*+\sum_{i,j\in\{0,1\}}M_n^{ij}(z_n^{ij}-\rho^{ij}),\quad n=1,2,\dots,\qquad (\mathcal{K}_0^*=1).$$

As a natural generalization of the beta-binomial distribution treated in [7], let us take the Dirichlet-multinomial distribution as $Q(\xi^{2n})=Q(z_1\dots z_n)$. Then the corresponding strategy is given by (cf. (5) and (7))

$$M_n^{ij}=\mathcal{K}_{n-1}^*\frac{m_{n-1}^{ij}+c^{ij}}{\rho^{ij}(n-1+c)},\quad i,j=0,1,$$

where c^{ij}'s are positive hyperparameters of the Dirichlet prior and $\sum_{i,j\in\{0.1\}}c^{ij}=c$. The capital process $\mathcal{K}_n^*=\mathcal{K}_{2n}^{\mathcal{P}_Q}$ for this strategy is given by

$$\begin{aligned}\mathcal{K}_n^*=\frac{Q(\xi^{2n})}{\prod_{i,j}(\rho^{ij})^{m_n^{ij}}}&=\frac{\prod_{i,j}\Gamma(m_n^{ij}+c^{ij})/\Gamma(c^{ij})}{\big(\Gamma(n+c)/\Gamma(c)\big)\prod_{i,j}(\rho^{ij})^{m_n^{ij}}}\\ &=\frac{\Gamma(c)\prod_{i,j}\Gamma(m_n^{ij}+c^{ij})}{\prod_{i,j}\Gamma(c^{ij})\Gamma(n+c)\prod_{i,j}(\rho^{ij})^{m_n^{ij}}},\end{aligned}$$

where in the products i, j range over $\{0, 1\}$.

We evaluate the asymptotic behavior of this capital. Denote $m_n^{ij}/n = \hat{p}_n^{ij}$, $i, j = 0, 1$. Then as in (8) we have

$$\log \mathcal{K}_n^* = n \sum_{i,j \in \{0,1\}} \hat{p}_n^{ij} \log \frac{\hat{p}_n^{ij}}{\rho^{ij}} - O(\log n).$$

Hence for even n the original capital process $\mathcal{K}_n^{\mathcal{P}_Q} = \mathcal{K}_{n/2}^*$ is written as

$$\log \mathcal{K}_n^{\mathcal{P}_Q} = \log \mathcal{K}_{n/2}^* = \frac{n}{2} D\big(\{\hat{p}_n^{ij}\} \parallel \{\rho^{ij}\}\big) - O(\log n).$$

Now if we neglect the pairwise block patterns and apply the strategy of [7] based on s_n only, then corresponding capital process $\mathcal{K}_n^0$ behaves as (8). We compare $\mathcal{K}_n^{\mathcal{P}_Q}$ and $\mathcal{K}_n^0$. By (9) with $\lambda = 1/2$ we have

$$\log \mathcal{K}_n^{\mathcal{P}_Q} - \log \mathcal{K}_n^0 = \frac{n}{2} D\big((\hat{p}_n^{11}, \hat{p}_n^{10}, \hat{p}_n^{01}, \hat{p}_n^{00}) \parallel (\hat{\rho}_n^{11}, \hat{\rho}_n^{10}, \hat{\rho}_n^{01}, \hat{\rho}_n^{00})\big) - O(\log n), \tag{10}$$

where $\hat{\rho}_n^{11} = (\bar{x}_n)^2, \hat{\rho}_n^{10} = \hat{\rho}_n^{01} = \bar{x}_n(1 - \bar{x}_n), \hat{\rho}_n^{00} = (1 - \bar{x}_n)^2$. We see that $\mathcal{K}_n^{\mathcal{P}_Q}$ exploits the pairwise block type non-randomness more effectively than $\mathcal{K}_n^0$ by the amount half the Kullback divergence given in the right-hand side of (10).

So far we have only considered even n. This is sufficient for analyzing the asymptotic behavior of $\mathcal{K}_n^{\mathcal{P}_Q}$. For completeness we discuss $\mathcal{K}_n^{\mathcal{P}_Q}$ for odd n. We can decompose Skeptic's bet M_n^{ij} on $z_n^{ij} = 1$ to the $(2n-1)$-th round and the $2n$-th round as follows.

1) at the $(2n-1)$-th round, Skeptic bets $\rho M_n^{11} + (1-\rho) M_n^{10}$ on $x_{2n-1} = 1$ and $\rho M_n^{01} + (1-\rho) M_n^{00}$ on $x_{2n-1} = 0$,

2a) if $x_{2n-1} = 1$, then at the $2n$-th round he bets M_n^{11} on $x_{2n} = 1$ and M_n^{10} on $x_{2n} = 0$,

2b) if $x_{2n-1} = 0$, then at the $2n$-th round he bets M_n^{01} on $x_{2n} = 1$ and M_n^{00} on $x_{2n} = 0$.

Denote $m_n^{i+} = \sum_{j=0}^1 m_n^{ij}$ and $c^{i+} = \sum_{j=0}^1 c^{ij}$, $i = 0, 1$. Then the capital at an odd round $\mathcal{K}_{2n+1}^{\mathcal{P}_Q}$ is written as follows.

$$\begin{aligned}\mathcal{K}_{2n+1}^{\mathcal{P}_Q} &= \mathcal{K}_{2n}^{\mathcal{P}_Q} \times \begin{cases} (m_n^{1+} + c^{1+})/(\rho(n-1+c)), & \text{if } x_{2n+1} = 1 \\ (m_n^{0+} + c^{0+})/((1-\rho)(n-1+c)), & \text{if } x_{2n+1} = 0 \end{cases} \\ &= \mathcal{K}_{2n}^{\mathcal{P}_Q} \times \frac{\Gamma(c)\Gamma(m_n^{1+} + c^{1+} + x_{2n+1})\Gamma(m_n^{0+} + c^{0+} + 1 - x_{2n+1})}{\rho^{x_{2n+1}}(1-\rho)^{1-x_{2n+1}}\Gamma(n+c)\Gamma(m_n^{1+} + c^{1+})\Gamma(m_n^{0+} + c^{0+})}.\end{aligned}$$

We can construct a similar strategy for the sequence of pairs $(x_2x_3)\dots(x_{2n}x_{2n+1})$ and can also combine these two strategies by splitting the initial capital into two equal parts and applying the corresponding strategy for each of them. Let $\mathcal{K}^B = \mathcal{K}^{\mathcal{P}_Q}$ denote the capital process considered so far and let $\tilde{\mathcal{K}}^B$ denote the similar capital process based on $(x_2x_3)(x_4x_5)\dots$. Then the capital process of the combined strategy is written as

$\mathcal{K} = (1/2)(\mathcal{K}^B + \tilde{\mathcal{K}}^B)$. Then $\limsup_n \mathcal{K}_n = \infty$, if the relative frequency of consecutive pairs $(x_{2n-1}x_{2n})$, $(x_{2n}x_{2n+1})$, taking the value (ij) does not converge to $\{\rho^{ij}/2\}$ for some (ij).

We have discussed blocks of length two for notational simplicity. The above derivation can be extended to n blocks of consecutive k-tuples, and hereafter we outline the procedure. Let $\boldsymbol{x}_1^k, \boldsymbol{x}_2^k, \dots, \boldsymbol{x}_n^k$ be the first n blocks of k-tuple with

$$\boldsymbol{x}_m^k = (x_{k(m-1)+1}, x_{k(m-1)+2}, \dots, x_{km}),\ m = 1, \dots, n,$$

and let $m_n^{\boldsymbol{\epsilon}_k}$, $\boldsymbol{\epsilon}_k = \epsilon_1 \dots \epsilon_k$, $\epsilon_i = 1$ or 0, denote the number of the consecutive k-tuple $\boldsymbol{\epsilon}_k$ among the first n blocks. Also denote

$$\rho^{\boldsymbol{\epsilon}_k} = \rho^{\sum_{i=1}^k \epsilon_i}(1-\rho)^{k-\sum_{i=1}^k \epsilon_i}.$$

If the sequence is random, we will have $\lim_{n\to\infty} m_n^{\boldsymbol{\epsilon}_k}/n = \rho^{\boldsymbol{\epsilon}_k}$ for all $\boldsymbol{\epsilon}_k \in \{0,1\}^k$.

At the $k(n-1)+1$-st round ($n = 1, 2, \dots$), Skeptic chooses 2^k amounts $M_n^{\boldsymbol{\epsilon}_k}$ and bet them on $\boldsymbol{x}_n^k = \boldsymbol{\epsilon}_k$. Then we have

$$\mathcal{K}_{kn} = \mathcal{K}_{k(n-1)} + \sum_{\boldsymbol{\epsilon}_k \in \{0,1\}^k} M_n^{\boldsymbol{\epsilon}_k}(z_n^{\boldsymbol{\epsilon}_k} - \rho^{\boldsymbol{\epsilon}_k}),$$

where

$$z_n^{\boldsymbol{\epsilon}_k} = \begin{cases} 1, & \text{if } \boldsymbol{x}_n^k = \boldsymbol{\epsilon}_k \\ 0, & \text{otherwise.} \end{cases}$$

Thus the derived game $\mathcal{K}_n^* = \mathcal{K}_{kn}$ is defined with the protocol

$$\mathcal{K}_n^* = \mathcal{K}_{n-1}^* + \sum_{\boldsymbol{\epsilon}_k \in \{0,1\}^k} M_n^{\boldsymbol{\epsilon}_k}(z_n^{\boldsymbol{\epsilon}_k} - \rho^{\boldsymbol{\epsilon}_k}),\ n = 1, 2, \dots, \qquad (\mathcal{K}_0^* = 1),$$

where $\sum_{\boldsymbol{\epsilon}_k \in \{0,1\}^k} z_n^{\boldsymbol{\epsilon}_k} = 1$.

If we take the Dirichlet-multinomial distribution as $Q(\xi^{kn}) = Q(z_1^{\boldsymbol{\epsilon}_k} \dots z_n^{\boldsymbol{\epsilon}_k})$, then corresponding strategy is given by

$$M_n^{\boldsymbol{\epsilon}_k} = \mathcal{K}_{n-1}^* \frac{m_{n-1}^{\boldsymbol{\epsilon}_k} + c^{\boldsymbol{\epsilon}_k}}{\rho^{\boldsymbol{\epsilon}_k}(n-1+c)},$$

where $\forall c^{\boldsymbol{\epsilon}_k} > 0$, $\sum_{\boldsymbol{\epsilon}_k \in \{0,1\}^k} c^{\boldsymbol{\epsilon}_k} = c$, are the hyperparameters of the Dirichlet prior. The capital process for this strategy is

$$\mathcal{K}_{kn}^{\mathcal{P}_Q} = \mathcal{K}_n^* = \frac{Q(\xi^{kn})}{\prod_{\boldsymbol{\epsilon}_k} (\rho^{\boldsymbol{\epsilon}_k})^{m_n^{\boldsymbol{\epsilon}_k}}} = \frac{\Gamma(c) \prod_{\boldsymbol{\epsilon}_k} \Gamma(m_n^{\boldsymbol{\epsilon}_k} + c^{\boldsymbol{\epsilon}_k})}{\prod_{\boldsymbol{\epsilon}_k} \Gamma(c^{\boldsymbol{\epsilon}_k}) \Gamma(n+c) \prod_{\boldsymbol{\epsilon}_k} (\rho^{\boldsymbol{\epsilon}_k})^{m_n^{\boldsymbol{\epsilon}_k}}}.$$

We evaluate the asymptotic behavior of this capital. Denote $m_n^{\boldsymbol{\epsilon}_k}/n = \hat{p}_n^{\boldsymbol{\epsilon}_k}$. Then

$$\log \mathcal{K}_n^* = nD\big(\{\hat{p}_n^{\boldsymbol{\epsilon}_k}\} \,\big\|\, \{\rho^{\boldsymbol{\epsilon}_k}\}\big) - O(\log n) = n \sum_{\boldsymbol{\epsilon}_k \in \{0,1\}^k} \hat{p}_n^{\boldsymbol{\epsilon}_k} \log \frac{\hat{p}_n^{\boldsymbol{\epsilon}_k}}{\rho^{\boldsymbol{\epsilon}_k}} - O(\log n). \tag{11}$$

Hence in the context of the original game, for n which is a multiple of k, we have

$$\log \mathcal{K}_n^{\mathcal{P}_Q} = \log \mathcal{K}_{n/k}^* = \frac{n}{k} D\big(\{\hat{p}_n^{\boldsymbol{\epsilon}_k}\} \parallel \{\rho^{\boldsymbol{\epsilon}_k}\}\big) - O(\log n).$$

Similar strategies can be constructed for the n blocks of the k-tuples with shift a

$$\boldsymbol{x}_{m+a}^k = (x_{k(m-1)+1+a}, x_{k(m-1)+2+a}, \dots, x_{km+a}), \quad 1 \le m \le n, \ a = 1, \dots, k-1.$$

We next combine k kinds of these strategies by splitting the initial capital into k equal parts and applying the corresponding strategy for each of them. Let $\mathcal{K}_n$ denote the resulting capital process. Then $\limsup_n \mathcal{K}_n = \infty$, if for any $\boldsymbol{\epsilon}_k$ the relative frequency of any of $\boldsymbol{x}_m^k, \boldsymbol{x}_{m+1}^k, \dots, \ \boldsymbol{x}_{m+k-1}^k \ \ (1 \le m \le n)$ taking the value $\boldsymbol{\epsilon}_k$ does not converge to $\rho^{\boldsymbol{\epsilon}_k}/k$.

It is of interest to compare the growth rates of $\mathcal{K}_n^{\mathcal{P}_Q}$ for different block lengths k. Note that block pattern strategy of length k also depends on the shift a. If the empirical distribution of $\{\boldsymbol{\epsilon}_k\}$ is different for a, then the main growth rate of a block strategy depends also on a. For simplicity we mainly consider the case that the empirical distribution $F_{n,a}$ of $\{\boldsymbol{x}_{m+a}^k\}_{m=1}^n$ is the same for different shifts a, in the sense that the total variation distance between $F_{n,a}$ and $F_{n,a'}$ converges to 0 for all $a \neq a'$. We call this case "homogeneous with respect to the shifts". We show in Section 3.3 that the main growth rate of (11) is non-decreasing in k under the assumption of homogeneity with respect to the shifts.

3.2 Markovian patterns

We can construct another strategy which exploits non-randomness in the moves of Reality. Such a procedure can be given as

$$M_1 = 0, \quad M_n = \begin{cases} M_n^+, & \text{if } x_{n-1} = 1 \\ M_n^-, & \text{if } x_{n-1} = 0 \end{cases} \quad n = 2, 3, \dots,$$

where M_n^+ and M_n^- can have different values. This is a first-order Markovian strategy, which incorporates the information on the previous move x_{n-1} of Reality. As in the previous section, for clarity of presentation we first consider the first-order Markovian strategy and then extend it to higher-order Markovian strategies.

Let $q_n^1 = s_n$ and $q_n^0 = n - s_n$. We also denote the numbers of pairs $(x_{i-1}x_i) = (11), (10), (01), (00)$, $i = 2, \dots, n$ by q_n^{11}, q_n^{10}, q_n^{01}, q_n^{00}, respectively. (For $n = 1$, let $0 = q_1^{11} = q_1^{10} = q_1^{01} = q_1^{00}$.) We take the beta-binomial distribution with parameters $a, b > 0$ for $Q(x_i \mid x_{i-1} = 1)$ and $Q(x_i \mid x_{i-1} = 0)$, $i \ge 2$. (The initial distribution is taken as $Q(1) = \rho = 1 - Q(0)$.) Then the corresponding strategy for $i \ge 2$ is given by

$$M_i^+ = \mathcal{K}_{i-1} \frac{p_i^+ - \rho}{\rho(1-\rho)}, \quad M_i^- = \mathcal{K}_{i-1} \frac{p_i^- - \rho}{\rho(1-\rho)},$$

with

$$p_i^+ = \frac{q_{i-1}^{11}+a}{q_{i-1}^{11}+q_{i-1}^{10}+a+b}, \quad p_i^- = \frac{q_{i-1}^{01}+a}{q_{i-1}^{01}+q_{i-1}^{00}+a+b}.$$

For $i=1$, we let $M_1^+ = M_1^- = 0$. The capital $\mathcal{K}_n = \mathcal{K}_n^{\mathcal{P}_Q}$ for this strategy is given by

$$\begin{aligned}\mathcal{K}_n &= \frac{\Gamma(a+b)\Gamma(q_n^{11}+a)\Gamma(q_n^{10}+b)}{\Gamma(a)\Gamma(b)\Gamma(q_n^{11}+q_n^{10}+a+b)\rho^{q_n^{11}}(1-\rho)^{q_n^{10}}} \times \frac{\Gamma(a+b)\Gamma(q_n^{01}+a)\Gamma(q_n^{00}+b)}{\Gamma(a)\Gamma(b)\Gamma(q_n^{01}+q_n^{00}+a+b)\rho^{q_n^{01}}(1-\rho)^{q_n^{00}}} \\ &= \frac{\Gamma(a+b)^2\Gamma(q_n^{11}+a)\Gamma(q_n^{10}+b)\Gamma(q_n^{01}+a)\Gamma(q_n^{00}+b)}{\Gamma(a)^2\Gamma(b)^2\Gamma(q_n^{11}+q_n^{10}+a+b)\Gamma(q_n^{01}+q_n^{00}+a+b)\rho^{q_n^{11}+q_n^{01}}(1-\rho)^{q_n^{10}+q_n^{00}}}.\end{aligned}$$

We evaluate the asymptotic behavior of this capital process. Write $q_n^1/n = \hat{p}_n = \bar{x}_n$, $q_n^{11}/q_n^1 = r_n^1$, and $q_n^{01}/q_n^0 = r_n^0$. Then as in (8) we have

$$\log \mathcal{K}_n = n\hat{p}_n D(r_n^1\|\rho) + n(1-\hat{p}_n)D(r_n^0\|\rho) - O(\log n). \tag{12}$$

Hence if either $\limsup_n |r_n^1-\rho|>0$ or $\limsup_n |r_n^0-\rho|>0$ then $\limsup_n \mathcal{K}_n = \infty$.

Now we compare (12) to the capital process (8) of the strategy based on s_n only. By counting the number of pairs we have $q_n^{11}+q_n^{01} = q_n^1 - x_1$ and

$$\hat{p}_n r_n^1 + (1-\hat{p}_n)r_n^0 = \frac{q_n^{11}}{n} + \frac{q_n^{01}}{n} = \hat{p}_n - \frac{x_1}{n} = \hat{p}_n - O(1/n).$$

Then by (9) with $\lambda = \hat{p}_n$ we have

$$\begin{aligned}\log \mathcal{K}_n - \log \mathcal{K}_n^0 &= n\Big[\hat{p}_n D(r_n^1\|\rho) + (1-\hat{p}_n)D(r_n^0\|\rho) - D(\hat{p}_n\|\rho)\Big] - O(\log n) \\ &= n\Big[\hat{p}_n D(r_n^1\|\hat{p}_n) + (1-\hat{p}_n)D(r_n^0\|\hat{p}_n)\Big] - O(\log n).\end{aligned} \tag{13}$$

We again see that $\log \mathcal{K}_n$ exploits the first-order Markovian non-randomness more effectively than $\log \mathcal{K}_n^0$ by the amount given above.

The above first-order procedure can be extended to the k-th order procedure based on sequence $\tilde{\boldsymbol{x}}_{n-1}^k = x_{n-k}\dots x_{n-1}$ of length k preceding x_n. We hereafter outline the procedure.

Let $q_n^{\boldsymbol{\epsilon}_k}$ denote the number of consecutive k-tuple $\boldsymbol{\epsilon}_k$ in $\xi^n = x_1\dots x_n$. For $(k+1)$-tuples $\boldsymbol{\epsilon}_k 1 = \epsilon_1\dots\epsilon_k 1$ and $\boldsymbol{\epsilon}_k 0 = \epsilon_1,\dots,\epsilon_k 0$ we similarly define $q_n^{\boldsymbol{\epsilon}_k 1}$ and $q_n^{\boldsymbol{\epsilon}_k 0}$. We take the beta-binomial distribution for $Q(x_i|\tilde{\boldsymbol{x}}_{i-1}^k = \boldsymbol{\epsilon}_k)$, $i \ge k+1$. (The initial distribution $Q(\xi^k)$ up to round k is taken as the risk neutral measure.) The corresponding strategy is

$$M_i = 0, \quad 1\le i\le k, \quad M_i = M_i^{\boldsymbol{\epsilon}_k} = \mathcal{K}_{i-1}\frac{p_i^{\boldsymbol{\epsilon}_k}-\rho}{\rho(1-\rho)}, \quad \text{if } \tilde{\boldsymbol{x}}_{i-1}^k = \boldsymbol{\epsilon}_k, \quad i\ge k+1,$$

with

$$p_i^{\boldsymbol{\epsilon}_k} = \frac{q_{i-1}^{\boldsymbol{\epsilon}_k 1}+a}{q_{i-1}^{\boldsymbol{\epsilon}_k 1}+q_{i-1}^{\boldsymbol{\epsilon}_k 0}+a+b}, \quad a,b>0.$$

The capital process $\mathcal{K}_n = \mathcal{K}_n^{\mathcal{P}_Q}$ for this strategy is written as

$$\begin{aligned}\mathcal{K}_n &= \prod_{\boldsymbol{\epsilon}_k\in\{0,1\}^k} \frac{\Gamma(a+b)\Gamma(q_n^{\boldsymbol{\epsilon}_k 1}+a)\Gamma(q_n^{\boldsymbol{\epsilon}_k 0}+b)}{\Gamma(a)\Gamma(b)\Gamma(q_n^{\boldsymbol{\epsilon}_k}+a+b)\rho^{q_n^{\boldsymbol{\epsilon}_k 1}}(1-\rho)^{q_n^{\boldsymbol{\epsilon}_k 0}}} \\ &= \frac{\Gamma(a+b)^{2^k}}{\Gamma(a)^{2^k}\Gamma(b)^{2^k}} \prod_{\boldsymbol{\epsilon}_k\in\{0,1\}^k} \frac{\Gamma(q_n^{\boldsymbol{\epsilon}_k 1}+a)\Gamma(q_n^{\boldsymbol{\epsilon}_k 0}+b)}{\Gamma(q_n^{\boldsymbol{\epsilon}_k}+a+b)\rho^{q_n^{\boldsymbol{\epsilon}_k 1}}(1-\rho)^{q_n^{\boldsymbol{\epsilon}_k 0}}}.\end{aligned}$$

We evaluate the asymptotic behavior of this capital. Denote $\hat{p}_n^{M,\boldsymbol{\epsilon}_k} = q_n^{\boldsymbol{\epsilon}_k}/n$, $r_n^{\boldsymbol{\epsilon}_k} = q_n^{\boldsymbol{\epsilon}_k 1}/q_n^{\boldsymbol{\epsilon}_k}$. Note that $\hat{p}_n^{M,\boldsymbol{\epsilon}_k}$ differs from $\hat{p}_n^{\boldsymbol{\epsilon}_k}$ in (11), because the latter only looks at the relative frequency of $\boldsymbol{\epsilon}_k$ among non-overlapping blocks of length k. As $n \to \infty$, we have

$$\log \mathcal{K}_n = n \sum_{\boldsymbol{\epsilon}_k\in\{0,1\}^k} \hat{p}_n^{M,\boldsymbol{\epsilon}_k} D\big(r_n^{\boldsymbol{\epsilon}_k} \| \rho\big) - O(\log n). \tag{14}$$

Hence if $\limsup_n |r^{\boldsymbol{\epsilon}_k} - \rho| > 0$ for any $\boldsymbol{\epsilon}_k$, then $\limsup_n \mathcal{K}_n = \infty$. In the next section we show that the first term on the right-hand side of (14) is non-decreasing in k.

3.3 Relations between block strategies and Markovian strategies

In (10) and (13) we saw that the block strategy of length two and the first-order Markovian strategy has a better main growth rate than the strategy based s_n only. In this section we give results on the comparison of main growth rates for general k, which is the block size for block strategies and the order for Markovian strategies.

For Markovian strategies we show that a larger k gives a better growth rate. Concerning block strategies we show that the same result holds under the assumption of homogeneity with respect to shifts and furthermore that Markovian strategy of order $k-1$ gives a better growth rate than the block strategy of length k.

We first consider Markovian strategies. Let $\mathcal{K}_n^{M,k}$ denote the capital process of k-th order Markovian strategy. In (14), for a given $\boldsymbol{\epsilon}_{k-1} = \epsilon_2 \dots \epsilon_k$ consider the sum of two terms involving $1\boldsymbol{\epsilon}_{k-1} = 1\epsilon_2 \dots \epsilon_k$ and $0\boldsymbol{\epsilon}_{k-1} = 0\epsilon_2 \dots \epsilon_k$. We have

$$\hat{p}_n^{M,1\boldsymbol{\epsilon}_{k-1}} + \hat{p}_n^{M,0\boldsymbol{\epsilon}_{k-1}} = \hat{p}_n^{M,\boldsymbol{\epsilon}_{k-1}} + O(1/n),$$

where $O(1/n)$ is due to the counting problem at the end of the sequence ξ^n. Also

$$\hat{p}_n^{M,1\boldsymbol{\epsilon}_{k-1}} r_n^{1\boldsymbol{\epsilon}_{k-1}} + \hat{p}_n^{M,0\boldsymbol{\epsilon}_{k-1}} r_n^{0\boldsymbol{\epsilon}_{k-1}} = \hat{p}_n^{M,\boldsymbol{\epsilon}_{k-1}} r_n^{\boldsymbol{\epsilon}_{k-1}} + O(1/n).$$

Then by (9)

$$\begin{aligned}&\hat{p}_n^{M,1\boldsymbol{\epsilon}_{k-1}} D(r_n^{1\boldsymbol{\epsilon}_{k-1}} \| \rho) + \hat{p}_n^{M,0\boldsymbol{\epsilon}_{k-1}} D(r_n^{0\boldsymbol{\epsilon}_{k-1}} \| \rho) - \hat{p}_n^{M,\boldsymbol{\epsilon}_{k-1}} D(r_n^{1\boldsymbol{\epsilon}_{k-1}} \| \rho) \\ &\qquad = \hat{p}_n^{M,1\boldsymbol{\epsilon}_{k-1}} D(r_n^{1\boldsymbol{\epsilon}_{k-1}} \| r_n^{\boldsymbol{\epsilon}_{k-1}}) + \hat{p}_n^{M,0\boldsymbol{\epsilon}_{k-1}} D(r_n^{0\boldsymbol{\epsilon}_{k-1}} \| r_n^{\boldsymbol{\epsilon}_{k-1}}) + O(1/n).\end{aligned} \tag{15}$$

Summing up over $\boldsymbol{\epsilon}_{k-1} \in \{0,1\}^{k-1}$ we have

$$\log \mathcal{K}_n^{M,k} - \log \mathcal{K}_n^{M,k-1}$$
$$= \sum_{\boldsymbol{\epsilon}_{k-1}\in\{0,1\}^{k-1}} \left[\hat{p}_n^{M,1\boldsymbol{\epsilon}_{k-1}} D(r_n^{1\boldsymbol{\epsilon}_{k-1}} \| r_n^{\boldsymbol{\epsilon}_{k-1}}) + \hat{p}_n^{M,0\boldsymbol{\epsilon}_{k-1}} D(r_n^{0\boldsymbol{\epsilon}_{k-1}} \| r_n^{\boldsymbol{\epsilon}_{k-1}})\right] - O(\log n). \quad (16)$$

Therefore the growth rate of the Markovian strategy of order k is larger than that of order $k-1$ by the amount shown on the right-hand side.

Next we consider block patterns. For $j = 0, 1$, define

$$\hat{p}_n^{j|\boldsymbol{\epsilon}_{k-1}} = \frac{\hat{p}_n^{\boldsymbol{\epsilon}_{k-1}j}}{\hat{p}_n^{\boldsymbol{\epsilon}_{k-1}}}, \qquad \rho^{j|\epsilon_{k-1}} = \begin{cases} \rho, & \text{if } j = 1, \\ 1-\rho, & \text{if } j = 0. \end{cases}$$

Then

$$\sum_{\boldsymbol{\epsilon}_k\in\{0,1\}^k} \hat{p}_n^{\boldsymbol{\epsilon}_k} \log \frac{\hat{p}_n^{\boldsymbol{\epsilon}_k}}{\rho^{\boldsymbol{\epsilon}_k}} = \sum_{\boldsymbol{\epsilon}_k\in\{0,1\}^{k-1}} \hat{p}_n^{\boldsymbol{\epsilon}_{k-1}} \sum_{j=0}^{1} \hat{p}_n^{j|\boldsymbol{\epsilon}_{k-1}} \log \frac{\hat{p}_n^{j|\boldsymbol{\epsilon}_{k-1}}}{\rho^{j|\boldsymbol{\epsilon}_{k-1}}} + \sum_{\boldsymbol{\epsilon}_k\in\{0,1\}^{k-1}} \hat{p}_n^{\boldsymbol{\epsilon}_{k-1}} \log \frac{\hat{p}_n^{\boldsymbol{\epsilon}_{k-1}}}{\rho^{\boldsymbol{\epsilon}_{k-1}}}.$$

Now under the the assumption of homogeneity with respect to shifts, the relative frequency $\hat{p}_n^{\boldsymbol{\epsilon}_k}$ of $\boldsymbol{\epsilon}_k$ is the same for different shifts a and this implies that $\hat{p}_n^{\boldsymbol{\epsilon}_{k-1}} = \hat{p}_n^{M,\boldsymbol{\epsilon}_{k-1}} + o(1)$ and $\hat{p}_n^{1|\boldsymbol{\epsilon}_{k-1}} = r_n^{\boldsymbol{\epsilon}_{k-1}} + o(1)$. Therefore

$$D\big(\{\hat{p}_n^{\boldsymbol{\epsilon}_k}\} \,\big\|\, \{\rho^{\boldsymbol{\epsilon}_k}\}\big) = \sum_{\boldsymbol{\epsilon}_k\in\{0,1\}^k} \hat{p}_n^{M,\boldsymbol{\epsilon}_{k-1}} D\big(r_n^{\boldsymbol{\epsilon}_{k-1}} \| \rho\big) + D\big(\{\hat{p}_n^{\boldsymbol{\epsilon}_{k-1}}\} \,\big\|\, \{\rho^{\boldsymbol{\epsilon}_{k-1}}\}\big) + o(1).$$

By induction on k we obtain

$$k \log \mathcal{K}_n^{B,k} = \log \mathcal{K}_n^{M,k-1} + (k-1)\log \mathcal{K}_n^{B,k-1} + o(n) = \sum_{i=0}^{k-1} \log \mathcal{K}_n^{M,i} + o(n),$$

or

$$\log \mathcal{K}_n^{B,k} = \frac{1}{k}\sum_{i=0}^{k-1} \log \mathcal{K}_n^{M,i} + o(n), \quad (17)$$

where $\mathcal{K}_n^{M,0} = \mathcal{K}_n^0$ is the capital process of the strategy based on s_n only. In (16) we saw that the growth rate of Markovian strategy is non-decreasing in k. It follows that under the assumption of homogeneity with respect to shifts the growth rate of the block strategy is also non-decreasing in k and furthermore the growth rate of the Markovian strategy of order $k-1$ is better than that of the block strategy of length k.

We should note that if the homogeneity does not hold, then the growth rate of the block strategy of length k might be better than that of the Markovian strategy of order $k-1$. If we divide the initial capital into k equal parts corresponding to each shift, then the combined capital process is the arithmetic average of the capital process for different shifts. The growth rate of the combined capital equals the maximum of capital processes for different shifts and the maximum might be better than that of the Markovian strategy or order $k-1$. The question of homogeneity comes up again in consideration of the asset trading game in continuous time.

3.4 Universal Bayesian Skeptic by mixture of priors

When we incorporate the strategies developed in the previous subsections, we get a strategy which can exploit any block or Markovian patterns of any length deviating from independent Bernoulli trials.

Let $\mathcal{P}^{B,k}$ be the Bayesian strategy which exploits the k-th order block patterns constructed in Section 3.1, and let $\mathcal{P}^{M,k}$ be the Bayesian strategy which exploits the k-th order Markovian patterns constructed in Section 3.2. At first we divide the initial capital $\mathcal{K}_0 = 1$ into equal two parts $\mathcal{K}_{0B} = 1/2$ and $\mathcal{K}_{0M} = 1/2$, and further divide $\mathcal{K}_{0B} = 1/2$ into countably infinite accounts with positive initial capitals $c_{B1}, c_{B2}, \dots,$ $\sum_{k=1}^{\infty} c_{Bk} = 1/2$, and also divide $\mathcal{K}_{0M} = 1/2$ similarly as $c_{M1}, c_{M2}, \dots,$ $\sum_{k=1}^{\infty} c_{Mk} = 1/2$. We apply the strategy $\mathcal{P}^{B,k}$ to the k-th account c_{Bk} and apply the strategy $\mathcal{P}^{M,k}$ to the k-th account c_{Mk}, respectively. The resulting "universal" strategy

$$\mathcal{P}^* = \mathcal{P}^*_B + \mathcal{P}^*_M = \sum_{k=1}^{\infty} c_{Bk}\mathcal{P}^{B,k} + \sum_{k=1}^{\infty} c_{Mk}\mathcal{P}^{M,k}$$

can exploit any block or Markovian pattern of any length.

We shall formulate and state this fact within the framework of measure-theoretic probability in order to clarify the connection to the universal source coding in information theory. For simplicity of statement we consider the coin-tossing game with $\rho = 1/2$ and use the base two logarithm. Let $\{p^{\boldsymbol{\epsilon_k}}\}$ denote the k-dimensional probability distribution of a random vector $(X_1, \dots, X_k) \in \{0,1\}^k$. Then

$$\frac{1}{\log 2} D(\{p^{\boldsymbol{\epsilon_k}}\} \| \{\rho^{\boldsymbol{\epsilon_k}}\}) = \sum_{\boldsymbol{\epsilon_k} \in \{0,1\}^k} p^{\boldsymbol{\epsilon_k}} \log_2 \frac{p^{\boldsymbol{\epsilon_k}}}{2^{-k}} = k - H(X_1, \dots, X_k),$$

where $H(X_1, \dots, X_k)$ denotes the entropy of $\{p^{\boldsymbol{\epsilon_k}}\}$. For an infinite sequence $X_1, X_2, \dots$ of stationary and ergodic 0-1 random variables, the entropy $H(\mathcal{X}) = H(X_1, X_2, \dots)$ is defined as $H(\mathcal{X}) = \lim_k (1/k) H(X_1, \dots, H_k)$. Under the assumption of stationarity and ergodicity, each k-dimensional empirical distribution converges to the probability distribution $\{p^{\boldsymbol{\epsilon_k}}\}$ almost surely. Furthermore in the previous subsection we saw that larger block sizes achieve better growth rates. Therefore, arguing as in Chapter 13 of [1] we have the following proposition.

Proposition 3.1. *If $X_1, X_2, \dots,$ are stationary and ergodic sequence of 0-1 random variables, then*

$$\frac{1}{n} \log_2 \mathcal{K}_n^{\mathcal{P}^*} \to 1 - H(\mathcal{X}), \quad a.s. \qquad (n \to \infty).$$

Thus we can say that $\mathcal{K}_n^{\mathcal{P}^*} \to \infty$ achieves the optimal rate in the sense of universal source coding in information theory.

4 Application to asset trading games in continuous time

In this section we apply the results on the block strategy of length two and the first-order Markovian strategy of the previous section to an asset trading game in continuous time considered in [16]. It should be noted that our interest here is to derive explicit growth rates of our strategies applied to the asset trading game, rather than a rigorous treatment of forcing of the variation exponent of two. Therefore in our derivation we proceed with informal definitions and convenient regularity conditions.

In Section 4.1 we summarize the setup of an asset trading game. In Section 4.2 we obtain growth rates of our strategies for the asset trading game when the variation exponent of the asset price path deviates from two.

4.1 Preliminaries on asset trading games in continuous time

Here we summarize preliminary facts on the asset trading game as formulated in [16]. Our framework in [16] is now much generalized in the recent papers of Vovk ([17],[18],[19]).

Suppose that there is a financial asset which is traded in continuous time. Let $S(t)$ denote the price of the unit amount of the asset at time t. We assume that $S(t)$ is positive and a continuous function of t. The price path $S(\cdot)$ is chosen by a player "Market", which is the same as Reality in the coin-tossing game. "Investor" enters the market at time $t = t_0 = 0$ with the initial capital of $\mathcal{K}(0) = 1$. He decides discrete time points $0 = t_0 < t_1 < t_2 < \cdots$ to trade the financial asset. The trading time t_i and the amount M_i of the asset Investor holds for the interval $[t_i, t_{i+1})$ can depend on the path of $S(t)$ up to time t_i.

The basic fact on the behavior of $S(t)$ is the "$\sqrt{\mathrm{d}t}$-effect" ([20], [16]), which asserts that infinitesimal increments $|\mathrm{d}S(t)|$ of the price path have to be of the order $O(\sqrt{\mathrm{d}t})$, in the sense that otherwise Investor can make arbitrarily large profit without risking bankruptcy. When $|\mathrm{d}S(t)| = O((\mathrm{d}t)^H)$, then $H \in (0, 1]$ is called the Hölder exponent or the Hurst index of $S(t)$. $1/H$ is called the variation exponent. Then the game-theoretic statement of $\sqrt{\mathrm{d}t}$-effect is that Investor can force the variation exponent of two.

We consider "limit order" strategy of Investor. Let $\delta > 0$ be a constant. Investor determines the trading times $t_1, t_2, \dots$ as follows. After t_i is determined, let t_{i+1} be the first time after t_i when either

$$\frac{S(t_{i+1})}{S(t_i)} = 1 + \delta \quad \text{or} \quad = \frac{1}{1+\delta} \tag{18}$$

happens. This procedure leads to a discrete time coin-tossing game embedded into the asset trading game as follows. Let

$$x_n = \frac{(1+\delta)S(t_{n+1}) - S(t_n)}{\delta(2+\delta)S(t_n)} = \begin{cases} 1, & \text{if } S(t_{n+1}) = S(t_n)(1+\delta), \\ 0, & \text{if } S(t_{n+1}) = S(t_n)/(1+\delta) \end{cases}$$

and

$$\rho = \rho_\delta = \frac{1}{2+\delta}.$$

Also write $\tilde{\mathcal{K}}_n = \mathcal{K}(t_{n+1})$. Clearly x_n can be thought as "heads" or "tails" chosen by Market after the time t_i. More formally we define the following protocol of an embedded discrete time coin-tossing game.

EMBEDDED DISCRETE TIME COIN-TOSSING GAME
Protocol:

$\tilde{\mathcal{K}}_0 := 1$.
FOR $n = 1, 2, \dots$:
  Investor announces $\nu_n \in \mathbb{R}$.
  Market announces $x_n \in \{0, 1\}$.
  $\tilde{\mathcal{K}}_n = \tilde{\mathcal{K}}_{n-1}(1 + \nu_n(x_n - \rho))$.
END FOR

This embedded discrete time game allows us to apply results on coin-tossing games to the asset trading game in continuous time. The amount M_i of the asset held during $[t_i, t_{i+1})$ is determined by our strategies in Section 3. From now on we fix the time interval $[0, T]$ of the asset trading game. Then the total number rounds $n = n(\delta)$ played in the embedded coin-tossing game is finite. For a given path, $n(\delta)$ is increased by letting $\delta \downarrow 0$ in (18). $n(\delta)$ diverges to infinity as $\delta \downarrow 0$, unless $S(t)$ is constant on $[0, T]$. We call a strategy with a small δ a high-frequency strategy. In [16] we applied high-frequency Bayesian strategy of [7] to the embedded discrete time game and proved that Investor can force the variation exponent of two, within an arbitrary small constant.

However the Bayesian strategy of [7] is based only on s_n and does not take the higher order patterns of the increments of $S(t)$ into account. In the previous section we saw that multistep Bayesian strategies can effectively exploit higher order patterns when Reality's moves are not random. Therefore it is of interest to investigate how fast Investor can increase his capital by a high-frequency multistep Bayesian strategy, if the variation exponent of $S(t)$ deviates from two.

4.2 Growth rates of block strategy of length two and first-order Markov strategy for asset trading game

In this section we derive growth rates of high-frequency block strategy of length two and first-order Markov strategy in the embedded coin-tossing game. Our results are stated in two propositions at the end of this section.

Write $\eta = \log(1 + \delta)$. Then $\eta \downarrow 0$ is equivalent to $\delta \downarrow 0$. We decrease η to zero as

$$\eta_k = 2^{-k}, \ k = 1, 2, \dots.$$

The advantage of taking this sequence of η_k is that the equi-spaced grids for $\log S(t)$ are completely nested in k and we can establish some important relations between the empirical distributions of block patterns for different k (see Lemma 4.1 below). We call the embedded coin-tossing game with η_k the k-th embedded coin-tossing game. Let $n(\eta_k)$ denote the total number of rounds of the k-th embedded coin-tossing game. For notational convenience we sometimes write n or n_k instead of $n(\eta_k)$. As in Section 3.1, let $m^{ij}_{n_k}$, $i, j = 0, 1$, denote the number of pairs (ij) among $(x_1x_2)(x_3x_4)\dots$ in the k-th embedded coin-tossing game. For the shift of one, let $\tilde{m}^{ij}_{n_k}$ denote the number of pairs (ij) among $(x_2x_3)(x_4x_5)\dots$ in the k-th embedded coin-tossing game. We define $q^i_{n_k}, q^{ij}_{n_k}$ as in Section 3.2.

We now give a preliminary consideration on the behavior of counts $n(\eta_k)$, $q^i_{n_k}$, $q^{ij}_{n_k}$, $m^{ij}_{n_k}$, $\tilde{m}^{ij}_{n_k}$ for different k. At this point it is helpful to consider properties of fractional Brownian motion (Chapter 4 of [3]). Let $\{B_H(t)\}$ denote the fractional Brownian motion of Hurst index H. $B_H(t)$ corresponds to $\log S(t)$ in the asset trading game. $B_H(t)$ is a typical stochastic process with $|\mathrm{d}B_H(t)| = O((\mathrm{d}t)^H)$. $\{B_H(t)\}$ is self-similar, i.e., for every $a > 0$ the distribution of $\{B_H(at)\}$ coincides with that of $\{a^H B_H(t)\}$. This implies that making the grid finer as $\eta_k \to \eta_k/2$ is equivalent (in distribution) to increasing T as $T \to 2^{1/H}T$. This suggests that when Market chooses a path $S(t)$ with a fixed exponent H, then

$$n_{k+1} \simeq 2^{1/H} n_k. \tag{19}$$

Furthermore $\{B_H(t)\}$ has stationary increments, i.e., the distribution of the increments of $\{B_H(t)\}$ are invariant with respect to arbitrary time shift. This corresponds to our assumption of homogeneity with respect to the shifts in Section 3.3. Under the homogeneity assumption we expect

$$2m^{ij}_{n_k} \simeq 2\tilde{m}^{ij}_{n_k} \simeq q^{ij}_{n_k}. \tag{20}$$

Also note the following trivial combinatorial relation for any n:

$$q^{11}_n + q^{01}_n = q^1_n - x_1, \ q^{11}_n + q^{10}_n = q^1_n - x_n \quad \text{and} \quad |q^{01}_n - q^{10}_n| = |x_n - x_1| \le 1.$$

From game-theoretic viewpoint Investor can force $H = 1/2$. If Market chooses a path with $H \neq 1/2$, then the notion of forcing can not be applied and there is no guarantee that (19) and (20) hold. However even for $H \neq 1/2$ we use (19), (20) as convenient regularity conditions for evaluating the growth rates of our strategies in view of the properties of the fractional Brownian motion. In Proposition 4.1 and Proposition 4.2 below we obtain the growth rate of the first-order Markovian strategy and block strategy of length two under these conditions. In these propositions the approximate equalities in (19) and (20) are understood in the sense that the ratios two sides converge to 1.

Now we state the following crucial combinatorial fact.

Lemma 4.1. *For each path $S(t)$ and for each k*

$$m^{11}_{n_k} = q^1_{n_{k-1}}, \qquad m^{00}_{n_k} = q^0_{n_{k-1}}.$$

Proof. Consider two nested equi-spaced grids with intervals η_{k-1} and $\eta_k = \eta_{k-1}/2$. For the grid with the interval η_{k-1}, consider a step, where the price is going upward from t_i to t_{i+1} i.e. $\log S(t_{i+1}) = \log S(t_i) + \eta_{k-1}$ in (18). It is obvious that this upward step corresponds exactly to two consecutive upward steps for the the grid with the interval η_k. Therefore $m^{11}_{n_k} = q^1_{n_{k-1}}$. By counting downward steps, we similarly obtain $m^{00}_{n_k} = q^0_{n_{k-1}}$. □

We now derive the growth rate of the first-order Markovian strategy. Define

$$TV(\eta_k, T) = \sum_{i=1}^{n_k} |\log S(t_i) - \log S(t_{i-1})| = n_k\eta_k = (q^1_{n_k} + q^0_{n_k})\eta_k,$$
$$L(\eta_k, T) = \log S(t_{n_k}) - \log S(0) = (q^1_{n_k} - q^0_{n_k})\eta_k.$$

Under (19), $n_{k+1}\eta_{k+1} \simeq 2^{1/H-1} n_k\eta_k$. Therefore $n_{k+1}\eta_{k+1} \to \infty$ as $k \to \infty$ for the case $H < 1$. On the other hand $S(t_{n_k}) \to S(T)$ as $k \to \infty$ and $L(\eta_k, T) \to \log S(T) - \log S(0)$. Therefore for each path, $L(\eta_k, T)/TV(\eta_k, T) \to 0$ as $k \to \infty$ and this implies that

$$\frac{q^1_{n_k}}{n_k} \to \frac{1}{2} \qquad (k \to \infty). \tag{21}$$

Also note that $1/2 = \lim_{\delta\to 0} \rho_\delta$.

Furthermore by Lemma 4.1, under (19)

$$q^{11}_{n_k} \simeq 2m^{11}_{n_k} = 2q^1_{n_{k-1}} \simeq n_{k-1}.$$

Therefore

$$r^1_{n_k} = \frac{q^{11}_{n_k}}{q^1_{n_k}} \simeq \frac{n_{k-1}}{n_k/2} \simeq \frac{1}{2^{1/H-1}}.$$

Similarly $r^0_{n_k} = q^{01}_{n_k}/q^0_{n_k} \simeq 1 - 1/2^{1/H-1}$. Let $\mathcal{K}^M_{n_k}$ denote the capital of the first-order Markovian strategy at the end of the k-th embedded coin-tossing game. By (12) we obtained the following proposition.

Proposition 4.1. *Suppose that Market chooses a path such that*

$$1 = \lim_{k\to\infty} \frac{n_{k+1}}{2^{1/H} n_k} = \lim_{k\to\infty} \frac{2m^{ij}_{n_k}}{q^{ij}_{n_k}} = \lim_{k\to\infty} \frac{2\tilde{m}^{ij}_{n_k}}{q^{ij}_{n_k}}, \quad i, j = 0, 1.$$

Then

$$\lim_{k\to\infty} \frac{1}{n_k} \log \mathcal{K}^M_{n_k} = D\Big(\frac{1}{2^{1/H-1}} \| \frac{1}{2}\Big).$$

We now consider the block strategy of length two. Let $\mathcal{K}^B_{n_k}$ denote the capital of block strategy at the end of the k-th embedded coin-tossing game. By (17) we know that $\log \mathcal{K}^B_{n_k}$ is the average of $\log \mathcal{K}^{M,1}_{n_k}$ and $\log \mathcal{K}^{M,0}_{n_k}$. However by (21) the growth rate of $\mathcal{K}^{M,0}_{n_k}$ is zero. Therefore we have the following result.

Proposition 4.2. *Under the same assumption as in Proposition 4.1*

$$\lim_{k\to\infty} \frac{1}{n_k} \log \mathcal{K}^B_{n_k} = \frac{1}{2} D\Big(\frac{1}{2^{1/H-1}} \| \frac{1}{2}\Big).$$

Therefore the growth rate of the block strategy is half of rate of the Markovian strategy. $\mathcal{K}_{n_k}^{M,0}$ is the capital process of the strategy based only on the past average of Reality's moves considered in [16], whose growth rate is zero. It is of interest to note that despite this zero growth rate, the strategy in [16] was sufficient to force the variation exponent of two of the Market's path. This suggests that looking for a simple strategy for forcing certain event and looking for a more aggressive strategy with a better growth rate need different considerations.

5 Discussions

In this paper we studied multistep Bayesian strategies for coin-tossing games. Our general conclusion was that asymptotically we obtain better growth rates by incorporating larger and larger block sizes for block strategies or longer orders for Markovian strategies. However this conclusion has to be taken with the following cautions. When the main growth term expressed in terms of the Kullback divergence is close to zero, we have to compare this to the term of order $O(\log n)$. Generally the term of order $O(\log n)$ can be understood as a penalty term for larger models, i.e., for using strategies incorporating larger blocks. Therefore if the coin-tossing game is played only a finite number of rounds, or Reality does not deviate too much from the independent Bernoulli trials, then it might be advantageous to use shorter block sizes. This is essentially the same tradeoff as in statistical model selection based on various information criteria. It is of great interest to consider selecting among strategies or dynamically adjusting weights for them.

For convenience we made the assumption of homogeneity for block strategies in 3.3 and in Section 4. We initially thought that homogeneity can be "forced" on Reality by appropriate strategies of Skeptic. However, when Reality deviates from independent Bernoulli trials, the game-theoretic notion of forcing can not be applied. Intuitively it seems that Skeptic can further exploit patterns in Reality's moves when the homogeneity with respect to shifts does not hold. However at present it seems difficult to formulate results in this direction.

In Section 3.4 we considered an infinite countable mixture of block strategies and Markovian strategies. Using this countable mixture, Skeptic can asymptotically exploit any deviation of Reality's moves from independent Bernoulli trials. We pointed out that the idea of the universal source coding in information theory is similar. Our result is also very closely connected to results in algorithmic theory of randomness. We can think of each component strategy as a test of randomness of Reality's moves. In algorithmic randomness there are strong computability restrictions on the allowed sample spaces. In the game-theoretic approach we do not have to worry about computability and by appropriate discretization it is now possible to discuss the randomness of continuous paths.

In Section 4.2 we only considered block strategies of length two and first-order Markovian strategies in the embedded coin-tossing game. We could obtain the explicit expression for the growth rate in Proposition 4.1 and Proposition 4.2 because of the combinatorial fact of Lemma 4.1. It is of interest to investigate growth rates of higher-order Markovian strategies in the asset trading game.

For measure-theoretic stochastic processes, the regularity conditions assumed in Propositions 4.1 and 4.2 are basically law of large numbers, and we expect that they hold for fractional Brownian motions. However the trading times in (18) are stopping times and the fractional Brownian motion for $H \neq 1/2$ is not a Markov process. Therefore it is not easy to prove that the regularity conditions hold for fractional Brownian motions.

A Equivalence of Bayesian strategy and prudent strategy in coin-tossing games

Here we establish a one-to-one correspondence between Skeptic's prudent strategy and a probability distribution on the set of paths Ξ in the coin-tossing game.

For one direction suppose that Skeptic models Reality's moves by a probability distribution Q. Write $\alpha_n = M_n/\mathcal{K}_{n-1}$. Given $\mathcal{K}_{n-1}$ assume that Skeptic tries to maximize the conditional expected value of $\log \mathcal{K}_n$. It is equivalent to maximizing

$$p_n \log(1 + \alpha_n(1-\rho)) + (1-p_n)\log(1-\alpha_n\rho) \tag{22}$$

with respect to α_n, where $p_n = p_n^Q$ is given in (3). The maximizing value of α_n is uniquely given as

$$\alpha_n = \frac{p_n - \rho}{\rho(1-\rho)}.$$

With this α_n,

$$\mathcal{K}_n = \begin{cases} \mathcal{K}_{n-1} p_n/\rho, & \text{if } x_n = 1 \\ \mathcal{K}_{n-1}(1-p_n)/(1-\rho), & \text{if } x_n = 0. \end{cases}$$

Note that $\mathcal{K}_n = 0$ if either $p_n = 0$ and $x_n = 1$ or $p_n = 1$ and $x_n = 0$. In this case Skeptic can not play any more. For other cases he can keep playing the game. It should be noted that this is consistent with the definition of conditional probability in (3), namely, Skeptic can continue the game if and only if (3) is defined. We have shown that a probability distribution Q leads to the strategy given in (4).

For another direction let $\mathcal{P}$ be a prudent strategy of Skeptic. Starting with the initial capital of $\mathcal{K}_0 = 1$, define

$$Q_1(1) = \rho + M_1\rho(1-\rho) = \rho(1 + M_1(1-\rho))$$
$$Q_1(0) = 1 - \rho - M_1\rho(1-\rho) = (1-\rho)(1 - M_1\rho)$$

Then $Q_1(0)$ and $Q_1(1)$ are non-negative and $1 = Q_1(0) + Q_1(1)$. For the case $\mathcal{K}_{n-1}(\xi^{n-1}) > 0$ recursively define

$$Q_n(\xi^{n-1}1) = \rho Q_{n-1}(\xi^{n-1})\Big(1 + \frac{M_n(\xi^{n-1})}{\mathcal{K}_{n-1}(\xi^{n-1})}(1-\rho)\Big)$$
$$Q_n(\xi^{n-1}0) = (1-\rho) Q_{n-1}(\xi^{n-1})\Big(1 - \frac{M_n(\xi^{n-1})}{\mathcal{K}_{n-1}(\xi^{n-1})}\rho\Big).$$

These are non-negative and satisfy the consistency condition (2). If $\mathcal{K}_{n-1}(\xi^{n-1}) = 0$, then define $0 = Q_n(\xi^{n-1}1) = Q_n(\xi^{n-1}0)$, which is also consistent. By this procedure a Skeptic's prudent strategy leads to a probability distribution $\mathcal{P} \mapsto Q$.

By construction it is obvious that this map is the inverse map to (4) and therefore there exists a one-to-one correspondence between the set of probability distributions and the set of Skeptic's strategies satisfying the collateral duty.

Finally we state the following Bayesian optimality result, which follows easily from the maximization in (22)

Proposition A.1. *Let Q be a probability distribution on Ξ and let $\mathcal{P}$ be the strategy corresponding to Q. For any other strategy $\tilde{\mathcal{P}}$*

$$E^Q(\log \mathcal{K}_n^{\mathcal{P}_Q}) \geq E^Q(\log \mathcal{K}_n^{\tilde{\mathcal{P}}}).$$

References

[1] T.M. Cover and J.A. Thomas. *Elements of Information Theory*, 2nd ed., Wiley, New York, 2006.

[2] Rod Downey and Denis R. Hirschfeldt. *Algorithmic Randomness and Complexity.* To be published by Springer, 2008.

[3] Paul Embrechts and Makoto Maejima. *Selfsimilar Processes.* Princeton University Press, New Jersey, 2002.

[4] Te Sun Han and Kingo Kobayashi. *Mathematics of Information and Coding.* Translations of mathematical monographs, v.203, American Mathematical Society, Providence, RI, 2002.

[5] Yasunori Horikoshi and Akimichi Takemura. Implications of contrarian and one-sided strategies for the fair-coin game. *Stochastic Processes and their Applications*, `doi:10.1016/j.spa.2007.11.007`, 2007.

[6] Masayuki Kumon and Akimichi Takemura. On a simple strategy weakly forcing the strong law of large numbers in the bounded forecasting game. *Annals of the Institute of Statistical Mathematics*, `doi:10.1007/s10463-007-0125-5`, 2007.

[7] Masayuki Kumon, Akimichi Takemura and Kei Takeuchi. Capital process and optimality properties of a Bayesian Skeptic in coin-tossing games. `arXiv:math/0510662v1`. To appear in *Stochastic Analysis and Applications*, 2008.

[8] Masayuki Kumon, Akimichi Takemura and Kei Takeuchi. Game-theoretic versions of strong law of large numbers for unbounded variables. *Stochastics*, **79**, No.5, 449–468, 2007.

[9] Michiel van Lambalgen. Von Mises' definition of random sequences reconsidered. *The Journal of Symbolic Logic*, **52**, 1987.

[10] Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications.* Springer, New York, 1997.

[11] André Nies. *Computability and Randomness.* To be published by Oxford University Press. 2008.

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

[13] Akimichi Takemura and Taiji Suzuki. Game theoretic derivation of discrete distributions and discrete pricing formulas. *Journal of the Japan Statistical Society*, **37**, 87–104, 2007.

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

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

[16] Kei Takeuchi, Masayuki Kumon and Akimichi Takemura. A new formulation of asset trading games in continuous time with essential forcing of variation exponent. `arXiv:0708.0275v1`, 2007.

[17] Vladimir Vovk. Continuous-time trading and emergence of randomness. `arXiv:0712.1275v2`, 2007.

[18] Vladimir Vovk. Continuous-time trading and emergence of volatility. `arXiv:0712.1483v2`, 2007.

[19] Vladimir Vovk. Game-theoretic Brownian motion. `arXiv:0801.1309v1`, 2008.

[20] Vladimir Vovk and Glenn Shafer. A game-theoretic explanation of the $\sqrt{dt}$ effect. Working Paper No.5. 2003. Available at `http://www.probabilityandfinance.com`