St. Petersburg portfolio games

László Györfi * and Péter Kevei **

Department of Computer Science and Information Theory
Budapest University of Technology and Economics
Magyar Tudósok körútja 2.,Budapest, Hungary, H-1117
gyorfi@szit.bme.hu
Analysis and Stochastics Research Group
Hungarian Academy of Sciences
Aradi vértanúk tere 1, Szeged, Hungary, H-6720
kevei@math.u-szeged.hu

Abstract. We investigate the performance of the constantly rebalanced portfolios, when the random vectors of the market process $\{\mathbf{X}_i\}$ are independent, and each of them distributed as $(X^{(1)}, X^{(2)}, \dots, X^{(d)}, 1)$, $d \geq 1$, where $X^{(1)}, X^{(2)}, \dots, X^{(d)}$ are nonnegative iid random variables. Under general conditions we show that the optimal strategy is the uniform: $(1/d, \dots, 1/d, 0)$, at least for d large enough. In case of St. Petersburg components we compute the average growth rate and the optimal strategy for $d = 1, 2$. In order to make the problem non-trivial, a commission factor is introduced and tuned to result in zero growth rate on any individual St. Petersburg components. One of the interesting observations made is that a combination of two components of zero growth can result in a strictly positive growth. For $d \geq 3$ we prove that the uniform strategy is the best, and we obtain tight asymptotic results for the growth rate.

1 Constantly rebalanced portfolio

Consider a hypothetical investor who can access d financial instruments (asset, bond, cash, return of a game, etc.), and who can rebalance his wealth in each round according to a portfolio vector $\mathbf{b} = (b^{(1)}, \dots, b^{(d)})$. The j-th component $b^{(j)}$ of $\mathbf{b}$ denotes the proportion of the investor's capital invested in financial instrument j. We assume that the portfolio vector $\mathbf{b}$ has nonnegative components and sum up to 1. The nonnegativity assumption means that short selling is not allowed, while the latter condition means that our investor does not consume nor deposit new cash into his portfolio, but reinvests it in each round. The set

* The first author acknowledges the support of the Computer and Automation Research Institute of the Hungarian Academy of Sciences.
** The work was supported in part by the Hungarian Scientific Research Fund, Grant T-048360.

of portfolio vectors is denoted by

$$\Delta_d = \left\{ \mathbf{b} = (b^{(1)}, \ldots, b^{(d)});\ b^{(j)} \geq 0,\ \sum_{j=1}^{d} b^{(j)} = 1 \right\}.$$

The behavior of the market is given by the sequence of return vectors $\{\mathbf{x}_n\}$, $\mathbf{x}_n = (x_n^{(1)}, \ldots, x_n^{(d)})$, such that the j-th component $x_n^{(j)}$ of the return vector $\mathbf{x}_n$ denotes the amount obtained after investing a unit capital in the j-th financial instrument on the n-th round.

Let S_0 denote the investor's initial capital. Then at the beginning of the first round $S_0 b_1^{(j)}$ is invested into financial instrument j, and it results in return $S_0 b_1^{(j)} x_1^{(j)}$, therefore at the end of the first round the investor's wealth becomes

$$S_1 = S_0 \sum_{j=1}^{d} b_1^{(j)} x_1^{(j)} = S_0 \langle \mathbf{b}_1 \,,\, \mathbf{x}_1 \rangle \,,$$

where $\langle \cdot \,,\, \cdot \rangle$ denotes inner product. For the second round $\mathbf{b}_2$ is the new portfolio and S_1 is the new initial capital, so

$$S_2 = S_1 \cdot \langle \mathbf{b}_2 \,,\, \mathbf{x}_2 \rangle = S_0 \cdot \langle \mathbf{b}_1 \,,\, \mathbf{x}_1 \rangle \cdot \langle \mathbf{b}_2 \,,\, \mathbf{x}_2 \rangle \,.$$

By induction, for the round n the initial capital is S_{n-1}, therefore

$$S_n = S_{n-1} \langle \mathbf{b}_n \,,\, \mathbf{x}_n \rangle = S_0 \prod_{i=1}^{n} \langle \mathbf{b}_i \,,\, \mathbf{x}_i \rangle \,. \tag{1}$$

Of course the problem is to find the optimal investment strategy for a long run period, that is to maximize S_n in some sense. The best strategy depends on the optimality criteria. A naive attitude is to maximize the expected return in each round. This leads to the risky strategy to invest all the money into the financial instrument j, with $\mathbb{E}X_n^{(j)} = \max\{\mathbb{E}X_n^{(i)} \,:\, i = 1, 2, \ldots, n\}$, where $\mathbf{X}_n = (X_n^{(1)}, X_n^{(2)}, \ldots, X_n^{(d)})$ is the market vector in the n-th round. Since the random variable $X_n^{(j)}$ can be 0 with positive probability, repeated application of this strategy lead to quick bankrupt. The underlying phenomena is the simple fact that $\mathbb{E}(S_n)$ may increase exponentially, while $S_n \to 0$ almost surely. A more delicate optimality criterion was introduced by Breiman [3]: in each round we maximize the expectation $\mathbb{E} \ln\langle \mathbf{b}, \mathbf{X}_n \rangle$ for $\mathbf{b} \in \Delta_d$. This is the so-called *log-optimal portfolio*, which is optimal under general conditions [3].

If the market process $\{\mathbf{X}_i\}$ is memoryless, i.e., it is a sequence of independent and identically distributed (i.i.d.) random return vectors then the log-optimal portfolio vector is the same in each round:

$$\mathbf{b}^* := \operatorname*{arg\,max}_{\mathbf{b} \in \Delta_d} \mathbb{E}\{\ln \langle \mathbf{b} \,,\, \mathbf{X}_1 \rangle\}.$$

In case of constantly rebalanced portfolio (CRP) we fix a portfolio vector $\mathbf{b} \in \Delta_d$. In this special case, according to (1) we get $S_n = S_0 \prod_{i=1}^n \langle \mathbf{b}\, ,\, \mathbf{x}_i \rangle$, and so the average growth rate of this portfolio selection is

$$\frac{1}{n} \ln S_n = \frac{1}{n} \ln S_0 + \frac{1}{n} \sum_{i=1}^n \ln \langle \mathbf{b}\, ,\, \mathbf{x}_i \rangle \, ,$$

therefore without loss of generality we can assume in the sequel that the initial capital $S_0 = 1$.

The optimality of $\mathbf{b}^*$ means that if $S_n^* = S_n(\mathbf{b}^*)$ denotes the capital after round n achieved by a log-optimum portfolio strategy $\mathbf{b}^*$, then for any portfolio strategy $\mathbf{b}$ with finite $\mathbb{E}\{\ln \langle \mathbf{b}\, ,\, \mathbf{X}_1 \rangle\}$ and with capital $S_n = S_n(\mathbf{b})$ and for any memoryless market process $\{\mathbf{X}_n\}_1^\infty$,

$$\lim_{n\to\infty} \frac{1}{n} \ln S_n \le \lim_{n\to\infty} \frac{1}{n} \ln S_n^* \quad \text{almost surely}$$

and maximal asymptotic average growth rate is

$$\lim_{n\to\infty} \frac{1}{n} \ln S_n^* = W^* := \mathbb{E}\{\ln \langle \mathbf{b}^*\, ,\, \mathbf{X}_1 \rangle\} \quad \text{almost surely.}$$

The proof of the optimality is a simple consequence of the strong law of large numbers. Introduce the notation

$$W(\mathbf{b}) = \mathbb{E}\{\ln \langle \mathbf{b}\, ,\, \mathbf{X}_1 \rangle\}.$$

Then the strong law of large numbers implies that

$$\begin{aligned} \frac{1}{n} \ln S_n &= \frac{1}{n} \sum_{i=1}^n \ln \langle \mathbf{b}\, ,\, \mathbf{X}_i \rangle \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}\{\ln \langle \mathbf{b}\, ,\, \mathbf{X}_i \rangle\} + \frac{1}{n} \sum_{i=1}^n \left(\ln \langle \mathbf{b}\, ,\, \mathbf{X}_i \rangle - \mathbb{E}\{\ln \langle \mathbf{b}\, ,\, \mathbf{X}_i \rangle\}\right) \\ &= W(\mathbf{b}) + \frac{1}{n} \sum_{i=1}^n \left(\ln \langle \mathbf{b}\, ,\, \mathbf{X}_i \rangle - \mathbb{E}\{\ln \langle \mathbf{b}\, ,\, \mathbf{X}_i \rangle\}\right) \\ &\to W(\mathbf{b}) \quad \text{almost surely.} \end{aligned}$$

Similarly,

$$\lim_{n\to\infty} \frac{1}{n} \ln S_n^* = W(\mathbf{b}^*) = \max_{\mathbf{b}} W(\mathbf{b}) \quad \text{almost surely.}$$

In connection with CRP in a more general setup we refer to Kelly [8] and Breiman [3].

In the following we assume that the i.i.d. random vectors $\{\mathbf{X}_i\}$, have the general form $\mathbf{X} = (X^{(1)}, X^{(2)}, \ldots, X^{(d)}, X^{(d+1)})$, where $X^{(1)}, X^{(2)}, \ldots, X^{(d)}$ are nonnegative i.i.d. random variables and $X^{(d+1)}$ is the cash, that is $X^{(d+1)} \equiv 1$, and $d \geq 1$. Then the concavity of the logarithm, and the symmetry of the first d components immediately imply that the log-optimal portfolio has the form $\mathbf{b} = (b, b, \dots, b, 1 - db)$, where of course $0 \leq b \leq 1/d$. When does $b = 1/d$ correspond to the optimal strategy; that is when should we play with all our money? In our special case W has the form

$$W(b) = \mathbb{E}\left\{ \ln \left(b \sum_{i=1}^{d} X^{(i)} + 1 - bd \right) \right\}.$$

Let denote $Z_d = \sum_{i=1}^{d} X^{(i)}$. Interchanging the order of integration and differentiation, we obtain

$$\frac{d}{db} W(b) = \mathbb{E}\left\{ \frac{d}{db} \ln \left(b \sum_{i=1}^{d} X^{(i)} + 1 - bd \right) \right\} = \mathbb{E}\left\{ \frac{Z_d - d}{bZ_d + 1 - bd} \right\}.$$

For $b = 0$ we have $W'(0) = \mathbb{E}(Z_d) - d$, which is nonnegative if and only if $\mathbb{E}(X^{(1)}) \geq 1$. This implies the intuitively clear statement that we should risk at all, if and only if the expectation of the game is not less than one. Otherwise the optimal strategy is to take all your wealth in cash. The function $W(\cdot)$ is concave, therefore the maximum is in $b = 1/d$ if $W'(1/d) \geq 0$, which means that

$$\mathbb{E}\left\{ \frac{d}{Z_d} \right\} \leq 1\,. \tag{2}$$

According to the strong law of large numbers $d/Z_d \to 1/\mathbb{E}(X^{(1)})$ a.s. as $d \to \infty$, thus under some additional assumptions for the underlying variables $\mathbb{E}(d/Z_d) \to 1/\mathbb{E}(X^{(1)})$, as $d \to \infty$. Therefore if $\mathbb{E}(X^{(1)}) > 1$, then for d large enough the optimal strategy is $(1/d, \dots, 1/d, 0)$.

In the latter computations we tacitly assumed some regularity conditions, that is we can interchange the order of differentiation and integration, and that we can take the L^1-limit instead of almost sure limit. One can show that these conditions are satisfied if the underlying random variables have strictly positive infimum. We skip the technical details.

2 St.Petersburg game

2.1 Iterated St.Petersburg game

Consider the simple St.Petersburg game, where the player invests 1$ and a fair coin is tossed until a tail first appears, ending the game. If the first tail appears in step k then the the payoff X is 2^k and the probability of this event is 2^{-k}:

$$\mathbb{P}\{X = 2^k\} = 2^{-k}. \tag{3}$$

The distribution function of the gain is

$$F(x) = \mathbb{P}\{X \leq x\} = \begin{cases} 0, & \text{if } x < 2\,, \\ 1 - \frac{1}{2^{\lfloor \log_2 x \rfloor}} = 1 - \frac{2^{\{\log_2 x\}}}{x}, & \text{if } x \geq 2\,, \end{cases} \tag{4}$$

where $\lfloor x \rfloor$ is the usual integer part of x and $\{x\}$ stands for the fractional part.

Since $\mathbb{E}\{X\} = \infty$, this game has delicate properties (cf. Aumann [1], Bernoulli [2], Haigh [7], and Samuelson [10]). In the literature, usually the repeated St. Petersburg game (called iterated St.Petersburg game, too) means multi-period game such that it is a sequence of simple St.Petersburg games, where in each round the player invests 1\$. Let X_n denote the payoff for the n-th simple game. Assume that the sequence $\{X_n\}_{n=1}^{\infty}$ is i.i.d. After n rounds the player's gain in the repeated game is $\bar{S}_n = \sum_{i=1}^{n} X_i$, then

$$\lim_{n\to\infty} \frac{\bar{S}_n}{n \log_2 n} = 1$$

in probability, where $\log_2$ denotes the logarithm with base 2 (cf. Feller [6]). Moreover,

$$\liminf_{n\to\infty} \frac{\bar{S}_n}{n \log_2 n} = 1$$

a.s. and

$$\limsup_{n\to\infty} \frac{\bar{S}_n}{n \log_2 n} = \infty$$

a.s. (cf. Chow and Robbins [4]). Introducing the notation for the largest payoff

$$X_n^* = \max_{1 \le i \le n} X_i$$

and for the sum with the largest payoff withheld

$$S_n^* = \sum_{i=1}^{n} X_i - X_n^* = \bar{S}_n - X_n^*,$$

one has that

$$\lim_{n\to\infty} \frac{S_n^*}{n \log_2 n} = 1$$

a.s. (cf. Csörgő and Simons [5]).

2.2 Sequential St.Petersburg game

According to the previous results $\bar{S}_n \approx n \log_2 n$. Next we introduce the sequential St.Petersburg game, having exponential growth. The sequential St.Petersburg game means that the player starts with initial capital $S_0 = 1\$$, and there is a sequence of simple St.Petersburg games, and for each simple game the player reinvests his capital. If $S_{n-1}^{(c)}$ is the capital after the $(n-1)$-th simple game then the invested capital is $S_{n-1}^{(c)}(1-c)$, while $S_{n-1}^{(c)} c$ is the proportional cost of the simple game with commission factor $0 < c < 1$. It means that after the n-th round the capital is

$$S_n^{(c)} = S_{n-1}^{(c)}(1-c)X_n = S_0(1-c)^n \prod_{i=1}^{n} X_i = (1-c)^n \prod_{i=1}^{n} X_i.$$

Because of its multiplicative definition, $S_n^{(c)}$ has exponential trend:

$$S_n^{(c)} = 2^{nW_n^{(c)}} \approx 2^{nW^{(c)}},$$

with average growth rate

$$W_n^{(c)} := \frac{1}{n}\log_2 S_n^{(c)}$$

and with asymptotic average growth rate

$$W^{(c)} := \lim_{n\to\infty} \frac{1}{n}\log_2 S_n^{(c)}.$$

Let's calculate the the asymptotic average growth rate. Because of

$$W_n^{(c)} = \frac{1}{n}\log_2 S_n^{(c)} = \frac{1}{n}\left(n\log_2(1-c) + \sum_{i=1}^{n}\log_2 X_i\right),$$

the strong law of large numbers implies that

$$W^{(c)} = \log_2(1-c) + \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\log_2 X_i = \log_2(1-c) + \mathbb{E}\{\log_2 X_1\}$$

a.s., so $W^{(c)}$ can be calculated via expected log-utility (cf. Kenneth [9]). A commission factor c is called fair if

$$W^{(c)} = 0,$$

so the growth rate of the sequential game is 0. Let's calculate the fair c:

$$\log_2(1-c) = -\mathbb{E}\{\log_2 X_1\} = -\sum_{k=1}^{\infty} k\cdot 2^{-k} = -2,$$

i.e., $c = 3/4$.

2.3 Portfolio game with one or two St.Petersburg components

Consider the portfolio game, where a fraction of the capital is invested in simple fair St.Petersburg games and the rest is kept in cash, i.e., it is a CRP problem with the return vector

$$\mathbf{X} = (X^{(1)},\dots,X^{(d)},X^{(d+1)}) = (X_1',\dots,X_d',1)$$

($d \geq 1$) such that the first d i.i.d. components of the return vector $\mathbf{X}$ are of the form

$$\mathbb{P}\{X' = 2^{k-2}\} = 2^{-k}, \tag{5}$$

($k \geq 1$), while the last component is the cash. The main aim is to calculate the largest growth rate W_d^*.

Proposition 1. *We have that* $W_1^* = 0.149$ *and* $W_2^* = 0.289$.

Proof. For $d = 1$, fix a portfolio vector $\mathbf{b} = (b, 1-b)$, with $0 \le b \le 1$. The asymptotic average growth rate of this portfolio game is

$$W(b) = \mathbb{E}\{\log_2 \langle \mathbf{b}\,, \mathbf{X}\rangle\} = \mathbb{E}\{\log_2(bX' + 1 - b)\} = \mathbb{E}\{\log_2(b(X/4 - 1) + 1)\}.$$

The function $\log_2$ is concave, therefore $W(b)$ is concave, too, so $W(0) = 0$ (keep everything in cash) and $W(1) = 0$ (the simple game is fair) imply that for all $0 < b < 1$, $W(b) > 0$. Let's calculate $\max_b W(b)$. We have that

$$\begin{aligned} W(b) &= \sum_{k=1}^{\infty} \log_2(b(2^k/4 - 1) + 1) \cdot 2^{-k} \\ &= \log_2(1 - b/2) \cdot 2^{-1} + \sum_{k=3}^{\infty} \log_2(b(2^{k-2} - 1) + 1) \cdot 2^{-k}. \end{aligned}$$

Figure 1 shows the curve of the average growth rate of the portfolio game. The function $W(\cdot)$ attains its maximum at $b = 0.385$, that is

$$\mathbf{b}^* = (0.385, 0.615)\,,$$

where the growth rate is $W_1^* = W(0.385) = 0.149$. It means that if for each round of the game one reinvests 38.5% of his capital such that the real investment is 9.6%, while the cost is 28.9%, then the growth rate is approximately 11%, i.e., the portfolio game with two components of zero growth rate (fair St.Petersburg game and cash) can result in growth rate of 10.9%.

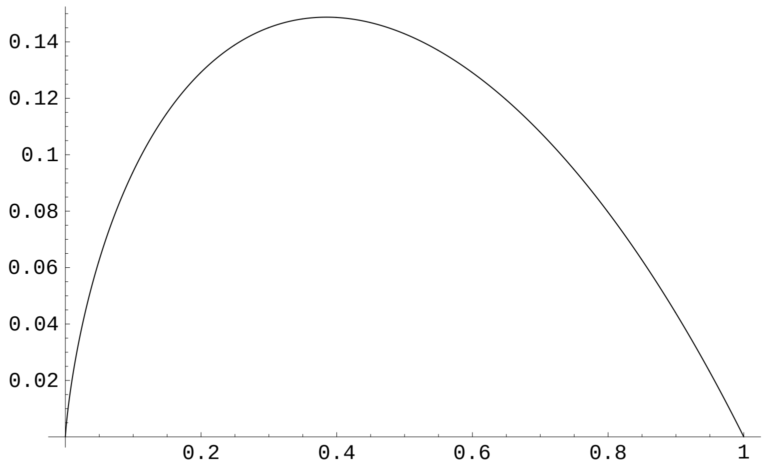

Figure 1. The growth rate for one St.Petersburg component

Consider next $d = 2$. At the end of Section 1 we proved that the log-optimal portfolio vector has the form $\mathbf{b} = (b, b, 1-2b)$, with $0 \le b \le 1/2$. The asymptotic average growth rate of this portfolio game is

$$\begin{aligned} W(b) &= \mathbb{E}\{\log_2 \langle \mathbf{b}\,,\, \mathbf{X}\rangle\} = \mathbb{E}\{\log_2(bX_1' + bX_2' + 1 - 2b)\} \\ &= \mathbb{E}\{\log_2(b((X_1 + X_2)/4 - 2) + 1)\}. \end{aligned}$$

Figure 2 shows the curve of the average growth rate of the portfolio game. Numerically we can determine that the maximum is taken at $b = 0.364$, so

$$\mathbf{b}^* = (0.364, 0.364, 0.272)\,,$$

where the growth rate is $W_2^* = W(0.364) = 0.289$. ∎

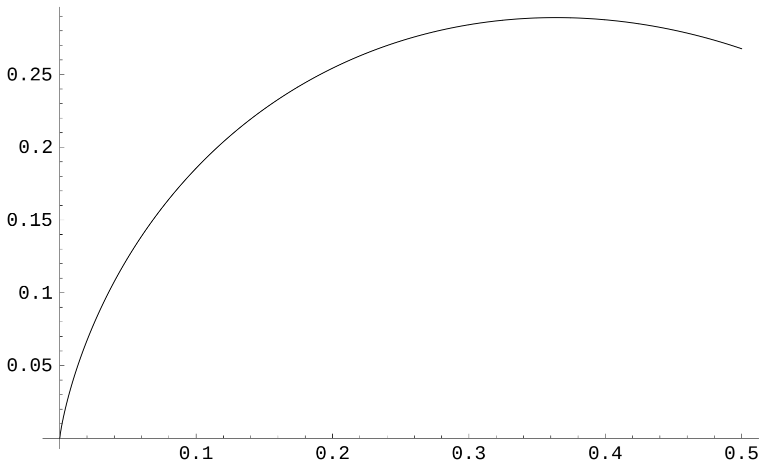

Figure 2. The growth rate for two St.Petersburg component

2.4 Portfolio game with at least three St.Petersburg components

Consider the portfolio game with $d \ge 3$ St.Petersburg components. We saw that the log-optimal portfolio has the form $\mathbf{b} = (b, \ldots, b, 1 - db)$ with $b \le 1/d$.

Proposition 2. *For $d \ge 3$, we have that*

$$\mathbf{b}^* = (1/d, \ldots, 1/d, 0).$$

Proof. Using the notations at the end of Section 1, we have to prove the inequality

$$\frac{d}{db} W(1/d) \ge 0\,.$$

According to (2) this is equivalent with

$$1 \geq \mathbb{E}\left\{\frac{d}{X_1' + \cdots + X_d'}\right\}.$$

For $d = 3, 4, 5$, numerically one can check this inequality. One has to prove the proposition for $d \geq 6$, which means that

$$1 \geq \mathbb{E}\left\{\frac{1}{\frac{1}{d}\sum_{i=1}^{d} X_i'}\right\}. \tag{6}$$

We use induction. Assume that (6) holds until $d-1$. Choose the integers $d_1 \geq 3$ and $d_2 \geq 3$ such that $d = d_1 + d_2$. Then

$$\begin{aligned}\frac{1}{\frac{1}{d}\sum_{i=1}^{d} X_i'} &= \frac{1}{\frac{1}{d}\sum_{i=1}^{d_1} X_i' + \frac{1}{d}\sum_{i=d_1+1}^{d} X_i'} \\ &= \frac{1}{\frac{d_1}{d}\frac{1}{d_1}\sum_{i=1}^{d_1} X_i' + \frac{d_2}{d}\frac{1}{d_2}\sum_{i=d_1+1}^{d} X_i'},\end{aligned}$$

therefore the Jensen inequality implies that

$$\frac{1}{\frac{1}{d}\sum_{i=1}^{d} X_i'} \leq \frac{d_1}{d}\frac{1}{\frac{1}{d_1}\sum_{i=1}^{d_1} X_i'} + \frac{d_2}{d}\frac{1}{\frac{1}{d_2}\sum_{i=d_1+1}^{d} X_i'},$$

and so

$$\begin{aligned}\mathbb{E}\left\{\frac{1}{\frac{1}{d}\sum_{i=1}^{d} X_i'}\right\} &\leq \mathbb{E}\left\{\frac{d_1}{d}\frac{1}{\frac{1}{d_1}\sum_{i=1}^{d_1} X_i'} + \frac{d_2}{d}\frac{1}{\frac{1}{d_2}\sum_{i=d_1+1}^{d} X_i'}\right\} \\ &= \frac{d_1}{d}\mathbb{E}\left\{\frac{1}{\frac{1}{d_1}\sum_{i=1}^{d_1} X_i'}\right\} + \frac{d_2}{d}\mathbb{E}\left\{\frac{1}{\frac{1}{d_2}\sum_{i=1}^{d_2} X_i'}\right\} \\ &\leq \frac{d_1}{d} + \frac{d_2}{d} = 1,\end{aligned}$$

where the last inequality follows from the assumption of the induction. ∎

2.5 Portfolio game with many St.Petersburg components

For $d \geq 3$, the best portfolio is the uniform portfolio with asymptotic average growth rate

$$W_d^* = \mathbb{E}\left\{\log_2\left(\frac{1}{d}\sum_{i=1}^{d} X_i'\right)\right\} = \mathbb{E}\left\{\log_2\left(\frac{1}{4d}\sum_{i=1}^{d} X_i\right)\right\}.$$

First we compute this growth rate numerically for small values of d, then we determine the exact asymptotic growth rate for $d \to \infty$.

For $d \geq 2$ arbitrary, by (3) we may write

$$\mathbb{E}\left\{\log_2\left(\sum_{i=1}^d X_i\right)\right\} = \sum_{i_1,i_2,\ldots,i_d=1}^{\infty} \frac{\log_2\left(2^{i_1}+2^{i_2}+\cdots+2^{i_d}\right)}{2^{i_1+i_2+\cdots+i_d}}.$$

Straightforward calculation shows that for $d \leq 8$, summing from 1 to 20 in each index independently, that is taking only 20^d terms, the error is less then $1/1000$. Here are the first few values:

d	1	2	3	4	5	6	7	8
W_d^*	0.149	0.289	0.421	0.526	0.606	0.669	0.721	0.765

Notice that W_1^* and W_2^* come from Section 2.3.

Now we return to the asymptotic results.

Theorem 1. *For the asymptotic behavior of the average growth rate we have*

$$-\frac{0.8}{\ln 2}\frac{1}{\log_2 d} \leq W_d^* - \log_2\log_2 d + 2 \leq \frac{\log_2\log_2 d + 4}{\ln 2 \log_2 d}.$$

Proof. Because of

$$W_d^* = \mathbb{E}\left\{\log_2\left(\frac{1}{4d}\sum_{i=1}^d X_i\right)\right\} = \mathbb{E}\left\{\log_2\left(\frac{\sum_{i=1}^d X_i}{d\log_2 d}\right)\right\} + \log_2\log_2 d - 2,$$

we have to show that

$$-\frac{0.8}{\ln 2}\frac{1}{\log_2 d} \leq \mathbb{E}\left\{\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d}\right\} \leq \frac{\log_2\log_2 d + 4}{\ln 2\log_2 d}.$$

Concerning the upper bound in the theorem, use the decomposition

$$\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d} = \log_2 \frac{\sum_{i=1}^d \tilde{X}_i}{d\log_2 d} + \log_2 \frac{\sum_{i=1}^d X_i}{\sum_{i=1}^d \tilde{X}_i},$$

where

$$\tilde{X}_i = \begin{cases} X_i, & \text{if } X_i \leq d\log_2 d, \\ d\log_2 d, & \text{otherwise.} \end{cases}$$

We prove that

$$\mathbb{E}\left\{\log_2 \frac{\sum_{i=1}^d \tilde{X}_i}{d\log_2 d}\right\} \leq \frac{\log_2\log_2 d + 2}{\ln 2\log_2 d}, \tag{7}$$

and

$$0 \leq \mathbb{E}\left\{\log_2 \frac{\sum_{i=1}^d X_i}{\sum_{i=1}^d \tilde{X}_i}\right\} \leq \frac{2}{\ln 2\log_2 d}. \tag{8}$$

For (8), we have that

$$\begin{aligned}
\mathbb{P}\left\{\log_2 \frac{\sum_{i=1}^d X_i}{\sum_{i=1}^d \tilde{X}_i} \geq x\right\} &= \mathbb{P}\left\{\frac{\sum_{i=1}^d X_i}{\sum_{i=1}^d \tilde{X}_i} \geq 2^x\right\} \\
&\leq \mathbb{P}\{\exists\, i \leq d \ : \ X_i \geq 2^x \tilde{X}_i\} \\
&= \mathbb{P}\{\exists\, i \leq d \ : \ X_i \geq 2^x \min\{X_i, d\log_2 d\}\} \\
&= \mathbb{P}\{\exists\, i \leq d \ : \ X_i \geq 2^x d\log_2 d\} \\
&\leq d\,\mathbb{P}\{X \geq 2^x d\log_2 d\} \\
&\leq d\frac{2}{2^x d\log_2 d}\,,
\end{aligned}$$

where we used that $\mathbb{P}\{X \geq x\} \leq 2/x$, which is an immediate consequence of (4). Therefore

$$\begin{aligned}
\mathbb{E}\left\{\log_2 \frac{\sum_{i=1}^d X_i}{\sum_{i=1}^d \tilde{X}_i}\right\} &= \int_0^\infty \mathbb{P}\left\{\log_2 \frac{\sum_{i=1}^d X_i}{\sum_{i=1}^d \tilde{X}_i} \geq x\right\} dx \\
&\leq \int_0^\infty \frac{2}{2^x \log_2 d} dx = \frac{2}{\ln 2 \log_2 d},
\end{aligned}$$

and the proof of (8) is finished. For (7), put $l = \lfloor \log_2(d\log_2 d)\rfloor$. Then for the expectation of the truncated variable we have

$$\begin{aligned}
\mathbb{E}(\tilde{X}_1) &= \sum_{k=1}^{l} 2^k \frac{1}{2^k} + d\log_2 d \sum_{k=l+1}^{\infty} \frac{1}{2^k} \\
&= l + d\log_2 d \frac{1}{2^{l+1}} 2 = l + \frac{d\log_2 d}{2^{\lfloor \log_2(d\log_2 d)\rfloor}} \leq l + 2\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}\left\{\log_2 \frac{\sum_{i=1}^d \tilde{X}_i}{d\log_2 d}\right\} &= \frac{1}{\ln 2}\mathbb{E}\left\{\ln \frac{\sum_{i=1}^d \tilde{X}_i}{d\log_2 d}\right\} \\
&\leq \frac{1}{\ln 2}\mathbb{E}\left\{\frac{\sum_{i=1}^d \tilde{X}_i}{d\log_2 d} - 1\right\} \\
&= \frac{1}{\ln 2}\left(\frac{\mathbb{E}\{\tilde{X}_1\}}{\log_2 d} - 1\right) \\
&\leq \frac{1}{\ln 2}\left(\frac{l+2}{\log_2 d} - 1\right) \\
&= \frac{1}{\ln 2}\left(\frac{\lfloor \log_2(d\log_2 d)\rfloor + 2}{\log_2 d} - 1\right) \\
&\leq \frac{1}{\ln 2}\left(\frac{\log_2 d + \log_2\log_2 d + 2}{\log_2 d} - 1\right) \\
&= \frac{1}{\ln 2}\frac{\log_2\log_2 d + 2}{\log_2 d}.
\end{aligned}$$

Concerning the lower bound in the theorem, consider the decomposition

$$\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d} = \left(\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d}\right)^+ - \left(\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d}\right)^- .$$

On the one hand for arbitrary $\varepsilon > 0$, we have that

$$\begin{aligned}
\mathbb{P}\left\{\frac{\sum_{i=1}^d X_i}{d\log_2 d} \le 2^x\right\} &\le \mathbb{P}\{ \text{ for all } i \le d,\ X_i \le 2^x d\log_2 d\} \\
&= \mathbb{P}\left\{X \le 2^x d\log_2 d\right\}^d \\
&\le \left(1 - \frac{1}{2^x d\log_2 d}\right)^d \\
&\le e^{-\frac{1}{2^x \log_2 d}} \\
&\le 1 - \frac{1-\varepsilon}{2^x \log_2 d},
\end{aligned}$$

for d large enough, where we used the inequality $e^{-z} \le 1-(1-\varepsilon)z$, which holds for $z \le -\ln(1-\varepsilon)$. Thus

$$\mathbb{P}\left\{\frac{\sum_{i=1}^d X_i}{d\log_2 d} > 2^x\right\} \ge \frac{1-\varepsilon}{2^x \log_2 d},$$

which implies that

$$\begin{aligned}
\mathbb{E}\left\{\left(\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d}\right)^+\right\} &= \int_0^\infty \mathbb{P}\left\{\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d} > x\right\} dx \\
&= \int_0^\infty \mathbb{P}\left\{\frac{\sum_{i=1}^d X_i}{d\log_2 d} > 2^x\right\} dx \\
&\ge \int_0^\infty \frac{1-\varepsilon}{2^x \log_2 d} dx \\
&= \frac{1}{\log_2 d} \frac{1-\varepsilon}{\ln 2}.
\end{aligned}$$

Since ε is arbitrary we obtain

$$\mathbb{E}\left\{\left(\log_2 \frac{\sum_{i=1}^d X_i}{d\log_2 d}\right)^+\right\} \ge \frac{1}{\log_2 d} \frac{1}{\ln 2}.$$

For the estimation of the negative part we use an other truncation method. Now we cut the variable at d, so put

$$\hat{X}_i = \begin{cases} X_i, \text{ if } X_i \le d\,, \\ d, \quad \text{otherwise.} \end{cases}$$

Introduce also the notations $\hat{S}_d = \sum_{i=1}^d \hat{X}_i$ and $c_d = \mathbb{E}(\hat{X}_1)/\log_2 d$. Similar computations as before show that

$$\mathbb{E}(\hat{X}_1) = \lfloor \log_2 d \rfloor + \frac{d}{2^{\lfloor \log_2 d \rfloor}} = \log_2 d + 2^{\{\log_2 d\}} - \{\log_2 d\} \quad \text{and}$$

$$\mathbb{E}\left[\hat{X}_1^2\right] \le 2\left(2^{\lfloor \log_2 d \rfloor} - 1\right) + \frac{d^2}{2^{\lfloor \log_2 d \rfloor}} = d\left(2^{1-\{\log_2 d\}} + 2^{\{\log_2 d\}}\right) \le 3d\,,$$

where we used that $2\sqrt{2} \le 2^{1-y} + 2^y \le 3$ for $y \in [0,1]$; this can be proved easily. Simple analysis shows again that $0.9 \le 2^y - y \le 1$ for $y \in [0,1]$, and so for $c_d - 1$ we obtain

$$\frac{0.9}{\log_2 d} < c_d - 1 < \frac{1}{\log_2 d}\,.$$

Since $\sum_{i=1}^d X_i \ge \sum_{i=1}^d \hat{X}_i$ we have that

$$\mathbb{E}\left\{\left(\log_2 \frac{\sum_{i=1}^d X_i}{d \log_2 d}\right)^-\right\} \le \mathbb{E}\left\{\left(\log_2 \frac{\sum_{i=1}^d \hat{X}_i}{d \log_2 d}\right)^-\right\}.$$

Noticing that

$$\log_2 \frac{\hat{S}_d}{d \log_2 d} > \log_2 \frac{2d}{d \log_2 d} = 1 - \log_2 \log_2 d\,,$$

we obtain

$$\mathbb{E}\left\{\left(\log_2 \frac{\hat{S}_d}{d \log_2 d}\right)^-\right\} = \int_{-\log_2 \log_2 d}^{0} \mathbb{P}\left\{\log_2 \frac{\hat{S}_d}{d \log_2 d} \le x\right\} dx\,,$$

thus we have to estimate the tail probabilities of $\hat{S}_d$.

According to Bernstein's inequality, for $x < 0$ we have

$$\begin{aligned}
\mathbb{P}\left\{\log_2 \frac{\hat{S}_d}{d \log_2 d} \le x\right\} &= \mathbb{P}\left\{\frac{\hat{S}_d - \mathbb{E}(\hat{S}_d)}{d \log_2 d} \le 2^x - \frac{\mathbb{E}(\hat{S}_d)}{d \log_2 d}\right\} \\
&= \mathbb{P}\left\{\frac{\hat{S}_d - \mathbb{E}(\hat{S}_d)}{d \log_2 d} \le 2^x - c_d\right\} \\
&\le \exp\left\{-\frac{d^2 \log_2^2 d\,(c_d - 2^x)^2}{2\left(d\,\mathbb{E}[(\hat{X})^2] + \frac{d^2 \log_2 d\,(c_d - 2^x)}{3}\right)}\right\} \\
&\le \exp\left\{-\frac{\log_2^2 d\,(c_d - 2^x)^2}{6 + \frac{2}{3}\log_2 d\,(c_d - 2^x)}\right\}.
\end{aligned}$$

Let $\gamma > 0$ be fixed, we define it later. For $x < -\gamma$ and d large enough the last upper bound $\le d^{-(1-2^{-\gamma})^2}$, therefore

$$\int_{-\log_2 \log_2 d}^{-\gamma} \mathbb{P}\left\{\log_2 \frac{\hat{S}_d}{d \log_2 d} \le x\right\} dx \le \frac{\log_2 \log_2 d}{d^{(1-2^{-\gamma})^2}}\,.$$

We give an estimation for the integral on $[-\gamma, 0]$:

$$\int_{-\gamma}^{0} \mathbb{P}\left\{\log_2 \frac{\hat{S}_d}{d\log_2 d} \le x\right\} dx \le \int_0^{\gamma} \exp\left\{-\frac{\log_2^2 d\,(c_d - 2^{-x})^2}{6 + \frac{2}{3}\log_2 d\,(c_d - 2^{-x})}\right\} dx$$
$$= \frac{1}{\ln 2}\int_0^{\gamma \ln 2} \exp\left\{-\frac{\log_2^2 d\,(c_d - e^{-x})^2}{6 + \frac{2}{3}\log_2 d\,(c_d - e^{-x})}\right\} dx\,.$$

For arbitrarily fixed $\varepsilon > 0$ we choose $\gamma > 0$ such that $1 - x \le e^{-x} \le 1 - (1-\varepsilon)x$, for $0 \le x \le \gamma \ln 2$. Using also our estimations for $c_d - 1$ we may write

$$\exp\left\{-\frac{\log_2^2 d\,(c_d - e^{-x})^2}{6 + \frac{2}{3}\log_2 d\,(c_d - e^{-x})}\right\} \le \exp\left\{-\frac{\log_2^2 d\,(0.9/\log_2 d + (1-\varepsilon)x)^2}{6 + \frac{2}{3}\log_2 d\,(1/\log_2 d + x)}\right\}$$

and continuing the estimation of the integral we have

$$\le \frac{1}{\ln 2}\int_0^{\gamma\ln 2} \exp\left\{-\frac{\log_2^2 d\,(0.9/\log_2 d + (1-\varepsilon)x)^2}{6 + \frac{2}{3}\log_2 d\,(1/\log_2 d + x)}\right\} dx$$
$$= \frac{1}{\ln 2}\frac{1}{\log_2 d}\int_0^{\log_2 d\,\gamma \ln 2} \exp\left\{-\frac{(0.9 + (1-\varepsilon)x)^2}{6 + \frac{2}{3}(1+x)}\right\} dx$$
$$\le \frac{1}{\ln 2}\frac{1}{\log_2 d}\int_0^{\infty} \exp\left\{-\frac{(0.9 + (1-\varepsilon)x)^2}{6 + \frac{2}{3}(1+x)}\right\} dx$$
$$\le \frac{1.7}{\ln 2}\frac{1}{\log_2 d}\,,$$

where the last inequality holds if ε is small enough.

Summarizing, we have

$$\mathbb{E}\left\{\left(\log_2 \frac{\sum_{i=1}^d \hat{X}_i}{d\log_2 d}\right)^-\right\} = \int_{-\log_2\log_2 d}^{0} \mathbb{P}\left\{\log_2 \frac{\hat{S}_d}{d\log_2 d} \le x\right\} dx$$
$$\le \frac{\log_2\log_2 d}{d^{(1-2^{-\gamma})^2}} + \frac{1.7}{\ln 2}\frac{1}{\log_2 d}$$
$$\le \frac{1.8}{\ln 2}\frac{1}{\log_2 d}\,,$$

for d large enough. Together with the estimation of the positive part this proves our theorem. ∎

References

1. Aumann, R. J. The St. Petersburg paradox: A discussion of some recent comments. *Journal of Economic Theory*, 14:443–445, 1977.
2. Bernoulli, D. Originally published in 1738; translated by L. Sommer. Exposition of a new theory on the measurement of risk. *Econometrica*, 22:22–36, 1954.

3. Breiman, L. Optimal gambling systems for favorable games. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, 1:65–78, Univ. California Press, Berkeley, 1961.
4. Chow, Y. S. and Robbins, H. On sums of independent random variables with infinite moments and "fair" games. *Proc. Nat. Acad. Sci. USA*, 47:330–335, 1961.
5. Csörgő, S. and Simons, G. A strong law of large numbers for trimmed sums, with applications to generalized St. Petersburg games. *Statistics and Probability Letters*, 26:65–73, 1996.
6. Feller, W. Note on the law of large numbers and "fair" games. *Ann. Math. Statist.*, 16:301–304, 1945.
7. Haigh, J. *Taking Chances.* Oxford,UK: Oxford University Press, 1999.
8. Kelly, J. L. A new interpretation of information rate. *Bell System Technical Journal*, 35:917–926, 1956.
9. Kenneth, A. J. The use of unbounded utility functions in expected-utility maximization: Response. *Quarterly Journal of Economics*, 88:136–138, 1974.
10. Samuelson, P. The St. Petersburg paradox as a divergent double limit. *International Economic Review*, 1:31–37, 1960.