
# Politecnico di Torino

**Master in Physics of Complex Systems**

**A.a. 2022/2023**

**Graduation session 07/2023**

## Learning strategies in Kelly's horse model

**Supervisors:**
prof. Matteo Marsili
prof. Andrea Pagnani

**Candidate:**
Pierluigi Proietti
300252

**Abstract**

We define Kelly's strategy for the horse race model, describing the concept of growth rate. We outline an adaptive strategy from Despons et al. (1), in which the gambler uses Bayesian inference to try and learn the win probabilities of the horses. We implement a modified version of this strategy, using our knowledge of the payoffs. Moreover, we use a Bayesian model selection approach to the problem, in the framework of Haimovici et al. (2). Lastly, we go back to the modified Bayesian inference approach and devise a strategy that aims at maximizing a linear combination of the expected value and the variance of the growth rate.

# Contents

# Introduction

In 1956 Kelly published a research based on the application of the concepts of information theory to gambling (3). Although these fields seem to be unrelated, the model he developed showed otherwise. Indeed, in the so called "Horse race model" there is a strong connection between the optimal growth rate of the player's wealth and the entropy of the underlying horses win probability distribution.
This study had a massive impact in finance, since the idea contributed to explain the usefulness of diversification in repeated investments (4). Over the years Kelly's strategy has been extended to the management of portfolio investments and has become an important tool in finance (5).
The approach differs from the one of standard portfolio theory, which treats investments as a single-period problem. In this case, one is left only with the problem of picking the best portfolio for a given level of risk, evaluating the probability and the outcome of a series of possible investments. In this framework, as a general principle, one should follow the proverb "don't put all your eggs in one basket", diversifying in order to avoid volatility: the implicit assumptions are that the investment horizon is known and that the investment is final. Understandably, this is not the case when one considers multi-period investment problems: the horse race is a special case of these kind of investments in the stock market (6).

Kelly's horse race model is a standard betting model where the player is asked to distribute his gamble on a batch of horses. Then the race takes place and one horse comes out victorious with a certain probability which is fixed in time. In the end, the gambler receives what he bet on the winning horse times its payoff, which is set before the race and remains also constant over time. The game is repeated indefinitely, therefore the gambler wishes to maximize the asymptotic behaviour of his capital or *growth rate*.
We will first describe the details of the model and then show the optimal solution in the standard case where both the win probabilities of the horses and their payoffs are known. In the case of no *track-take*, i.e. when the sum of the inverse of the payoffs is 1, Kelly himself demonstrated that the optimal growth rate is achieved betting *proportionally* to the win probabilities, thus disregarding the values of the payoffs.

We will focus on the case of unknown probabilities but known payoffs, which reflects a more realistic setting. Then, to play optimally, one needs to learn the probabilities and for this we will resort to Bayesian inference. The adaptation procedures we present allow to estimate the probabilities that are necessary to bet proportionally. First, we describe an estimator based on Laplace rule of succession, that is, calculated as the expected value of the posterior distribution of the probabilities, as in (1). We continue analysing an improved version of this estimator, obtained integrating as prior information the knowledge of the payoffs. On top of that, we build a slightly different strategy, based on the learning time estimate. Since it is very likely to observe a loss phase in the beginning, we wonder if a more cautious approach yields a better result.
Furthermore, we consider a Bayesian model selection approach, in the framework of Haimovici et al. (2). Instead of estimating from the beginning a number of parameter corresponding to the number of horses, to avoid fitting noise we produce only a small number of estimators. Horses with similar frequencies will see themselves being assigned the same probability parameter, thus being betted the same amount on. This procedure should in principle grant a better statistical relevance.
At last, we use a mean-variance approach, which is a standard method of portfolio theory (6). It consists in making the bet according to the estimator that maximises a utility function properly constructed, using the first two moments of the growth rate calculated on the posterior distribution of the $\mathbf{p}$ vector.
Clearly, the expected value measures the payoff of the race. Risk, instead, is considered through a measure of dispersion (variance). If offered several strategies with the same risk, a greedy player should prefer the one with the highest payoff. Similarly, among several strategies with the same expected value, a risk averse agent prefers the one with the lowest variance. Thus, a greedy and risk averse agent has preferences represented by a functional, which we will try to maximize.

There exist an interesting parallel, although with its limitations, between the natural selection of bacterial populations in varying environments and this model of financial investment, as pointed out by many studies (7)(8). Indeed, there is an analogy between the capital in Kelly's model, which is invested on different assets, and the bacteria population in the biological model, which is distributed among different phenotypes that may respond differently if subject to perturbations.
Kussell and Leibler managed to show that in bacterial growth in a fluctuating environment, organisms can adopt a *bet-hedging* strategy, investing their resources to produce a variety of phenotypes. In oscillating conditions bacteria must constantly adapt to survive and this may be achieved through sensing their surroundings. However, maintaining an active sensory machinery has a cost, therefore a mechanism of stochastic phenotype switching can be preferred: the result will be a heterogeneous population that has better chance at survival. The information is conveyed, rather than directly as in the sensing case, indirectly to the organism via natural selection,

but it grants nonetheless the whole system some knowledge of the environment. A broader research on the value of the information for populations in varying environment was conducted by Rivoire et al., who studied the the problem of treating a quantitative measure of information for populations in varying environment. In his work another Kelly's bound plays a crucial role. It is the equality between the maximum increase in growth rate caused by some side information $Y$ on the environment $X$ and the mutual information between $X$ and $Y$. The possibility of measuring the gathered information and how this reflects on the growth rate is another question that motivates this study.
An interesting extension of this work is the study of the case of where both the probabilities and the payoffs are unknown (to a certain degree). This reflects both Shannon work on communication channels, where it is possible to have transmission errors, and also a real scenario like population growth in a varying environment. The question to answer in this case would be what strategy is optimal when the actual returns the background provides are perceived in a distorted way or perceived poorly (or not at all).

# 1

# Definition of the model

Consider a race that involves M horses. The $i_{th}$ horse wins with probability $p_i$ and its payoff is $o_i$ for each unit invested. Let us call $\mathbf{r}$ the inverse odd vector, defining $R = \sum_x \frac{1}{o_x}$ and $r_i = \frac{1}{Ro_i}$ so that $\mathbf{r}$ is always normalized. The gambler distributes all of his wealth across the horses: let $b_i$ be the fraction of the gambler's wealth invested in horse $i$, where $b_i \geq 0 \ \forall i$ and $B = \sum_i b_i = 1$. Hence if horse i wins the race, the gambler will receive $b_i o_i$ whereas all the other bets will be lost. Thus, at the end of the race, the gambler will have multiplied his wealth by a factor $b_x o_x$ having horse $x$ won, which happens with probability $p_x$. Thus, it is possible to write:

$$W_{t+1} = \frac{b_{x_{t+1}}}{r_{x_{t+1}}} W_t = S_{x_{t+1}} W_t \tag{1.1}$$

Where $W_t$ is the wealth at time $t$ and $S_{x_{t+1}}$ is defined as the wealth relative, i.e. the factor by which the gambler's wealth grows if horse $x$ wins the race.

Since $x_t$ is a random variable, then also the wealth at the end of the race is a random variable, and the gambler wishes to maximize its value. The naive strategy consists in betting everything on the horse that has the maximum expected return (i.e., the one with the maximum $p_i o_i$). However, given that one is dealing with multiperiod investments, it is obvious that the player will eventually lose all its money.
Instead, looking closely at equation 1.1 one can notice that it is possible to write

$$\frac{W_n}{W_0} = \prod_{i=1}^{n} S_{x_i} \qquad\qquad \left(\log \frac{W_n}{W_0}\right)^{\frac{1}{n}} = \frac{1}{n} \sum_{i=1}^{n} \log S_{x_i} \overset{n\to\infty}{=} m \tag{1.2}$$

Where we defined $m = \mathbb{E}\left[\log S_{x_n}\right]$, which will later on be referred at as doubling rate or growth rate. The last equality follows from the law of large numbers since all $S_x$ are independent and identically distributed. This implies that $W_t$ grows asymptotically to $W_0 2^{mt}$. Thus the problem of maximizing the exponential growth

of $W_t$ is analogous to finding the strategy that maximizes the expected value of the utility function $U(x) = \log x$. Note that the reason why we proceed in this way has nothing to do with a choice based on finance principles of the utility function to attach to the gambler's money, but merely with the fact that it is the logarithm which is additive in repeated races and to which the law of large numbers applies (3). In this way, one is able to treat the problem similarly to a single-period problem: with an individual race perspective and considering this utility function one is guaranteed the maximum growth rate in the long-run[1].

### 1.0.1 The optimal doubling rate

The doubling rate is a function of $\mathbf{p}$ and $\mathbf{b}$. That is, $m$ depends on both the probability distribution and the portfolio allocation. To find the best strategy one has to maximize $m(\mathbf{p}, \mathbf{b})$ with respect to $\mathbf{p}$, constrained to $B = \sum_x b_x = 1$. In order to do this, we write a functional using Lagrange multipliers

$$\mathcal{L} = \sum_i p_i \log (b_i o_i) - \nu \left( \sum_i b_i - 1 \right) \tag{1.3}$$

$$\frac{\partial \mathcal{L}}{\partial b_x} = \frac{p_x}{b_x} - \nu = 0 \Rightarrow p_x = \nu b_x \tag{1.4}$$

$\forall\, x = 1, ..., M$. The initial constraint and $\mathbf{p}$ normalization imposes $\nu = 1$. Therefore one can conclude that $\mathbf{b} = \mathbf{p}$ is a stationary point. To check whether or not it is also a maximum one can observe the following

$$\begin{aligned} m(\mathbf{p}, \mathbf{b}) = \sum_i p_i \log(b_i o_i) = \sum_i p_i \log \left( \frac{p_i}{p_i} b_i o_i \right) = \\ = \sum_i p_i \log o_i - H(\mathbf{p}) - D_{KL}(\mathbf{p} \| \mathbf{b}) \leq \sum_i p_i \log o_i - H(\mathbf{p}) \end{aligned} \tag{1.5}$$

Where $D_{KL}(\mathbf{p}\|\mathbf{b})$ is the Kullback-Leibler divergence which is always $\geq 0$. The equality is satisfied if and only if $\mathbf{b} = \mathbf{p}$ since this is the only case where $D_{KL}(\mathbf{p}\|\mathbf{b}) = 0$. Therefore proportional betting is log-optimal for a sequence of independent horse races if the gambler reinvests all of his wealth at each step (i.e. $B = 1$), independently of R. We identify with $m^*(\mathbf{p})$ the doubling rate yielded by the best strategy. The counter-intuitive lesson that one can learn from this calculation is that in this particular case of $\sum_x b_x = 1$ the strategy does not depend on the payoffs.

[1]This paragraph is adapted from "Elements of Information Theory"(9)

In the special case when $R = 1$, i.e. there is no *track-take* from the bookmakers, one has that

$$m(\mathbf{p}, \mathbf{b}) = \sum_i p_i \log\left(\frac{p_i}{p_i}\frac{b_i}{r_i}\right) = D_{KL}(\mathbf{p}\|\mathbf{r}) - D_{KL}(\mathbf{p}\|\mathbf{b}) \tag{1.6}$$

From this equation it is possible to get another interpretation for the doubling rate, as the difference between the distance of the bookmaker estimate from $\mathbf{p}$ and the distance of the gambler's estimate from $\mathbf{p}$: the gambler can make money only if his estimate is better than the one of the bookmaker. Notice also that it is possible to employ another strategy, imposing $\mathbf{b} = \mathbf{r}$. This strategy will be referred as the null strategy because it yields a null growth rate and thus keeps the wealth constant.

### 1.0.2 The case B$\leq$ 1

In general the strategy of investing the whole capital every time is not always optimal. To see this, we consider the same minimization problem as before but with the constraint $\sum_i b_i = B \leq 1$ . The functional then becomes

$$\mathcal{L} = \sum_i p_i \log\left(1 - B + b_i o_i\right) - \nu\left(\sum_i b_i - B\right) \tag{1.7}$$

Following the same steps as before (we refer the reader to Appendix A for the calculations) one gets

$$b_x^* = \left[(1-B)R + B\right] p_x - \frac{1-B}{o_x} \tag{1.8}$$

$$m^*(\mathbf{p}) = \log\left(1 - B + \frac{B}{R}\right) + D_{KL}(\mathbf{p}\|\mathbf{r}) \tag{1.9}$$

Thus, it is possible to identify three cases, based on the value of R.

**Fair game, R=1**

One has that $m^*(\mathbf{p}) = D_{KL}(\mathbf{p}\|\mathbf{r})$, therefore the optimal strategy does not depend on B, as long as every $b_x^*$ is non negative. The choice $B = 1$ yields the best growth rate possible along with proportional gambling.

**Favourable game, R<1**

Here $m^*(\mathbf{p}) = \log\left[1 + B\left(\frac{1}{R} - 1\right)\right] + D_{KL}(\mathbf{p}\|\mathbf{r})$, which is an increasing function of $B$. Again, $B = 1$ and proportional betting are optimal.

**Unfavourable game, R>1**

In this case the doubling rate is a decreasing function of $B$. The optimal result is achieved when $B$ is equal to the smallest possible result that keeps every $b_x$ non negative

$$b_x^* = p_x R - \frac{1}{o_x} + B\left(p_x(1-R) + \frac{1}{o_x}\right) \geq 0 \tag{1.10}$$

$$B \geq \frac{1 - o_x p_x R}{o_x p_x (1-R) + 1} \tag{1.11}$$

This is a decreasing function of $o_x p_x$ so the relevant quantity is $\min_x o_x p_x = k/R$.

$$B^* = \frac{R(1-k)}{R(1-k)+k} \tag{1.12}$$

This solution yields $W^*(p) = D_{KL}(p||r) - \log\left(R(1-k)+k\right)$, which becomes negative if $R > \bar{R} = \frac{e^{D_{KL}(\mathbf{p}\|\mathbf{r})}-k}{1-k}$. Therefore, the best strategy is to play the modified strategy $\mathbf{b}^*$ if $1 < R < \bar{R}$ and to not play at all if $R > \bar{R}$.

# 2

# Learning to play

A common practice in finance is to use fractional Kelly's strategies, betting only a part of the available capital and not the whole sum even when the odds would suggest so. This is due to the fact that the pure Kelly is considered to be unsafe (5), whereas fractional Kelly reduces risk, although at the price of a non optimal growth rate. This observation suggests that there exists a trade-off between the risk the gambler is ready to take and the growth rate of his capital. This problem occurs particularly in the case where the probabilities $p_x$ are unknown. The gambler needs then to resort to some adaptation mechanisms to learn progressively the underlying distribution. Following the steps of Despons et al. (1), we devise a way to use Bayesian inference to learn the probabilities.
The method that will be presented qualifies as universal because no foreknowledge of the distribution $\mathbf{p}$ is needed; in economics such a strategy is known as universal portfolio. The relevance of this strategy resides in the fact that it performs asymptotically as well as the best constant balanced portfolios (10).
As stated previously, in the following it will be assumed that the payoffs are known and for simplicity we also consider the case where $R = 1$ and the gambler invests all of his capital at each step.

## 2.1 Bayesian inference Strategy

### 2.1.1 Laplace Strategy

Since proportional betting requires the gambler to know the probability vector, we start building an adaptive strategy to learn $\mathbf{p}$ using Bayesian inference. When no previous information on $\mathbf{p}$ is available, the prior distribution can be assumed uniform. To obtain an estimator for the probabilities one can therefore take the

expected value on the posterior distribution, which takes the form of

$$P[\mathbf{p}|\mathbf{n}^t] \propto \prod_{i=1}^{M} p_i^{n_i^t} \delta\left(\sum_x p_x - 1\right) \tag{2.1}$$

Where $\mathbf{n}^t$ is the vector that encodes the history of the races. This means that $n_i^t$ is equal to the times horse $i$ won the race up to time $t$ and that $\sum_x n_x^t = t$. Notice that in equation 2.1 the delta comes from the prior, imposing normalization, whereas the likelihood takes the form of a multinomial distribution.
The expectation value on this distribution produces Laplace's estimator

$$\hat{p}_x^{LAP,t+1} = \frac{n_x^t + 1}{t + M} \tag{2.2}$$

The strategy then consists in allocating a fraction $\hat{p}_x$ on horse $x$ and, after the race, updating the estimator according to the result. It is possible to check that Laplace estimator becomes unbiased for large $t$. To see this, as pointed out in (1), it is sufficient to calculate the Kullback-Leibler divergence between $\mathbf{p}$ and $\hat{\mathbf{p}}_x^{LAP}$ and show that it behaves as $\frac{M-1}{2t}$.
Although one could in principle consider the estimator given by the maximum of the posterior, it is possible to observe that Laplace estimator performs better specially when some of the $p_i$ are very small with respect to $\frac{1}{M}$ (1). Indeed, the estimator from equation 2.2 yields a non zero value even for unobserved events. Furthermore, at $t = 0$ the wealth allocation with this estimator is uniform, which is what one should expect with no prior information.

The average result obtained for the logarithm of the wealth (this is the self averaging quantity hence the one it is preferable to look at) by following Laplace's strategy has to be compared with the optimal one, given by proportional betting. Their difference is also referred as "Gambler's regret", since it quantifies the wealth lost by not playing accordingly to the real probabilities.
As shown before, the expected behaviour for $\log W$ when using Kelly's strategy is linear in time, with a slope equal to $D_{KL}(\mathbf{p}\|\mathbf{r})$; this result is matched numerically as shown in figure 2.1. Instead, for Laplace's strategy the same displays a *burn-in* phase, in which the gambler loses capital at short times because of his poor knowledge of the distribution $\mathbf{p}$. Then for large $t$ the optimal growth rate is recovered together with the linear behaviour.

In (1) it is presented a way to quantify the time up until one loses wealth if he plays following Laplace's strategy. This time, which we will refer to as $t^*$ or learning time, represents the time it takes to learn the probabilities better than the bookmaker. Qualitatively, one expects this time to depend on the distance between the inverse of the payoffs and the probabilities: if $\mathbf{r}$ is similar to $\mathbf{p}$, it will take

much more time to learn the probabilities well enough to get an actual gain than if the two distributions are far apart. Indeed, one can prove that $t^*_{LAP} \simeq \frac{M-1}{2D_{KL}(\mathbf{p}\|\mathbf{r})}$ (1). Note that this formula depends on $\mathbf{p}$, hence it is not possible to know the learning time exactly (otherwise it would be straightforward to wait up until $t^*$ before betting to have a better strategy), but an estimate will be given in section 2.1.3. In order to get a better understanding of the meaning of distance between $\mathbf{r}$ and $\mathbf{p}$, a particular definition for the inverse odd vector or rates will be used. We introduce the parameter $\varepsilon$ and a random vector $\boldsymbol{\eta}$ independent from $\mathbf{p}$ and also uniformly distributed, just like $\mathbf{p}$. One can then set

$$\mathbf{r} = (1-\varepsilon)\mathbf{p} + \varepsilon\boldsymbol{\eta} \tag{2.3}$$

This definition, although it is a strong assumption on the payoffs, will be used for the numerical simulations.

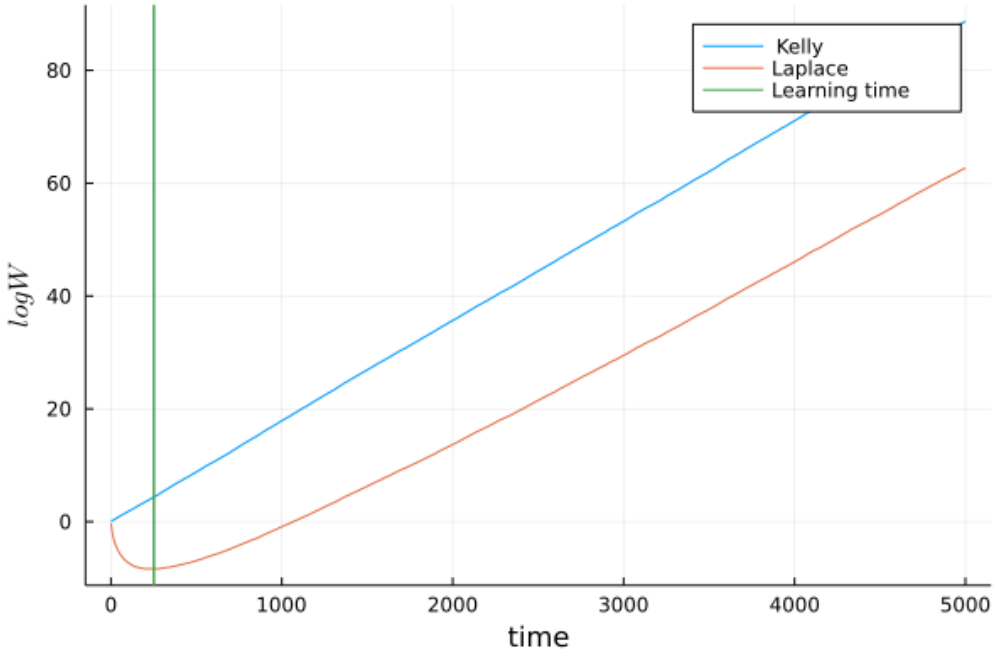


Figure 2.1. Comparison between Kelly's and Laplace strategy. The initial loss phase of Laplace strategy lasts up until the learning time. The time unit is one race, after which is possible to update the estimator. The parameters used are $\varepsilon = 0.3$, M=10 and the average are performed over 1000 trials

### 2.1.2 Modified Laplace Strategy

A question worth asking is how does the estimator change if one implements some prior information in equation 2.1. Using as a conjugate prior again a multinomial

distribution, the posterior becomes then

$$P[\mathbf{p}|\mathbf{n}^t] \propto \prod_{i=1}^{M} p_i^{n_i^t + a_i^\tau} \delta\left(\sum_x p_x - 1\right) \tag{2.4}$$

Where the elements $a_x^\tau$ encode the prior information. The estimator is therefore obtained by taking the expected value of $p_x$ over this distribution

$$\hat{p}_x^{LAP,t+1} = \frac{n_x^t + a_x^\tau + 1}{t + \tau + M} \tag{2.5}$$

Where $\tau = \sum_x a_x$. At this point, one should use the information in his possession to set the values of the $a_x$. At $t = 0$, the estimator, which is none other than $\mathbb{E}[p_x]$ calculated on equation 2.4, is $\frac{a_x^\tau + 1}{\tau + M}$; it is then reasonable to assert that $\mathbb{E}[p_x] = r_x$, since in principle there is ignorance about the payoff distribution. Note that this is not the correct assumption if one knows that the payoffs are defined as in equation 2.3. In this case, a better way of setting the $a_x$ is given in Appendix B. However, in both cases $\tau$ is left as a free parameter.
One has then $a_x^\tau + 1 = r_x(\tau + M)$ and thus

$$\hat{p}_x^{LAP,t} = \frac{n_x^t + r_x(\tau + M)}{t + \tau + M} \tag{2.6}$$

This is the same estimator one would get if $t + \tau + M$ race were observed, being that the first $\tau + M$ races yield an estimator equal to $r_x$. The parameter $\tau$ has then the meaning of the number of races the bookmaker saw before establishing the payoff using Bayesian inference.
The Modified Laplace strategy may yield a burn-in phase too, depending on how close $\mathbf{r}$ and $\mathbf{p}$ are, as shown in figure 2.6. If it is so, all conditions being equal, it allows to get a smaller $t^*$ than plain Laplace. Moreover, if this learning time is large enough for the Central Limit Theorem to work and if the assumption in equation 2.3 holds, it is possible to find a close formula for it

$$t^*_{MLAP} = \frac{M - 1}{2D_{KL}(\mathbf{p}\|\mathbf{r})} - 2\tau \tag{2.7}$$

Details about the calculation are given in Appendix C. It is worth noticing that a low $t^*$ does not necessary imply that the strategy is overall better. The quantity to consider is instead the gambler's regret.

$$\lim_{t\to\infty} \left(\log W_t^{KELLY} - \log W_t^{MLAP/LAP}\right) \tag{2.8}$$

Any strategy minimizing this difference will maximize the capital at large $t$. This is the reason why in the Modified Laplace method an high value of $\tau$ is not always

preferable to a low value although the first yields a lower $t^*$ than the second, as equation 2.7 shows. Rather, to determine the optimal fixed $\tau$ for each case numerical analysis was employed.

Even if it would seem like the Modified Laplace strategy is always preferable to plain Laplace, this is not the case when the inverse odd vector is distant enough from the probability vector, as shown in figure 2.6.

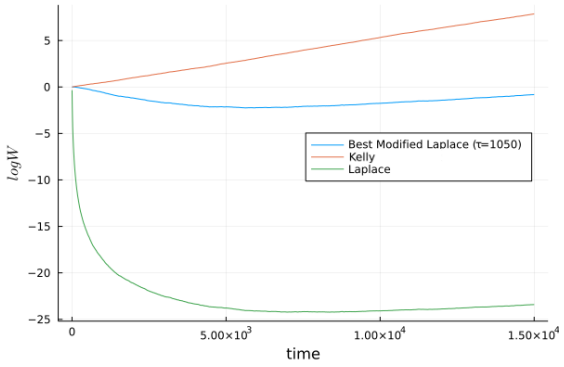


Figure 2.2. $\varepsilon = 0.01$

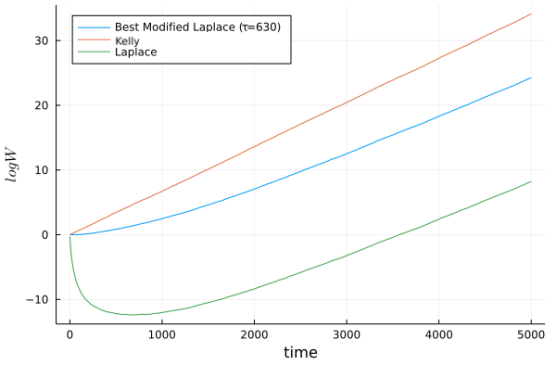


Figure 2.3. $\varepsilon = 0.1$

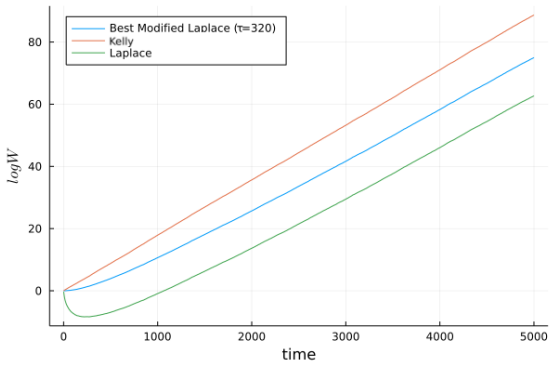


Figure 2.4. $\varepsilon = 0.3$

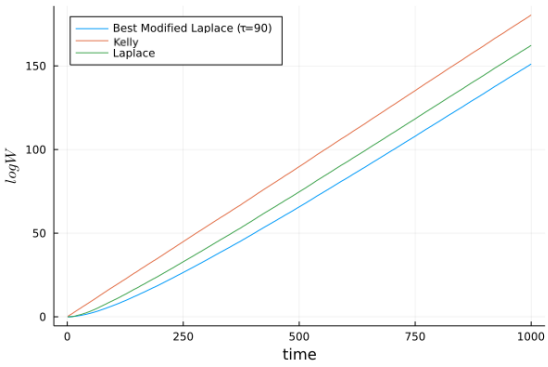


Figure 2.5. $\varepsilon = 0.5$

Figure 2.6. Numerical simulations comparison for different $\varepsilon$ values. ML shows a burn-in phase only in the case $\varepsilon = 0.01$ whereas L shows a burn-in phase always. In the case $\varepsilon = 0.5$ L is preferable to ML whereas in all other cases the opposite is true.

### 2.1.3 Learning time estimate

As stated previously, using

$$t^*_{LAP} \simeq \frac{M-1}{2D_{KL}(\mathbf{p}\|\mathbf{r})} \tag{2.9}$$

it is possible to get an estimate of the learning time in the Laplace method. If one also considers true equation 2.3 it is possible to derive

$$t^*_{MLAP} \simeq \frac{M-1}{2D_{KL}(\mathbf{p}\|\mathbf{r})} - 2\tau \tag{2.10}$$

It is then worth asking how well does perform a strategy in which the gambler calculates $t^*$, using his best knowledge of the $\mathbf{p}$, and waits up until that time before he starts playing. To see this, a numerical approach was adopted for both L and ML. Clearly, one expects this method to work only in the cases where one has no burn-in phase, since in that case betting immediately is surely advantageous and waiting to play is pointless. Of course, in principle one does not know if there is going to be such a phase, so also the cases of no burn-in phase for ML will be analyzed. It is possible to consider three situations: the case where L is preferable to ML and ML doesn't have a burn-in phase, the case where ML is preferable to L and ML doesn't have a burn-in phase and the case where ML is preferable to L and ML has a burn-in phase; note that L shows always a burn-in phase.

From figures 2.7, 2.8 and 2.9 it is possible to observe that for both L, ML the strategies that impose to wait, for large times, perform in the same way as the ones that establish to play from the start. Even if at small times the strategy of waiting may lead to better results, eventually the curve matches the result of the strategy where ones doesn't wait, both for ML and L. As figure 2.7 shows, this is not due to the fact that the $t^*$ are estimated poorly. Both on average but also on the single run, by the time that $t^*$ occurs the estimate is already pretty good. The analogy between the two approaches instead must be ascribed to the fact that by waiting, one secures no initial loss but also avoids all the highly profitable runs. It is however a safer way to proceed if one wishes to ensure in a single run a gain closer to the average.

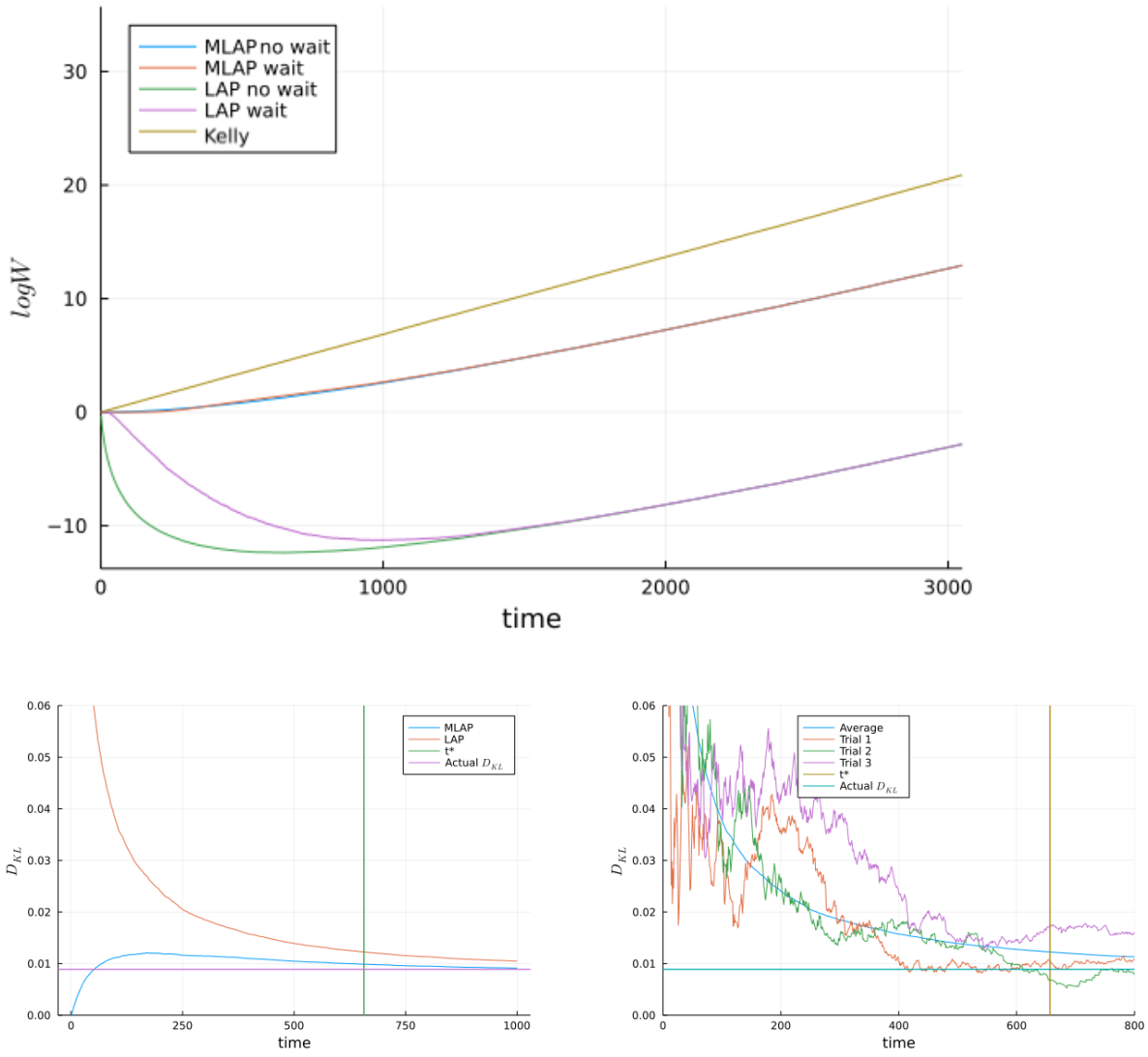


Figure 2.7. Numerical simulations for $\varepsilon = 0.1$. L shows a burn-in phase whereas ML doesn't; also, ML is preferable to L. The first graph represents the $\log W$ obtained with the different strategies which are: betting proportionally, using ML from the start, using ML after $\hat{t}^*_{MLAP}$ and analogously for L. The second graph shows the behaviour of the average estimated $D_{KL}$ using L estimator (orange) and using ML estimator (blue). In green it is highlighted $t^*_{LAP}$. The third graph shows a comparison between the $D_{KL}$ average and the $D_{KL}$ estimate in three different runs

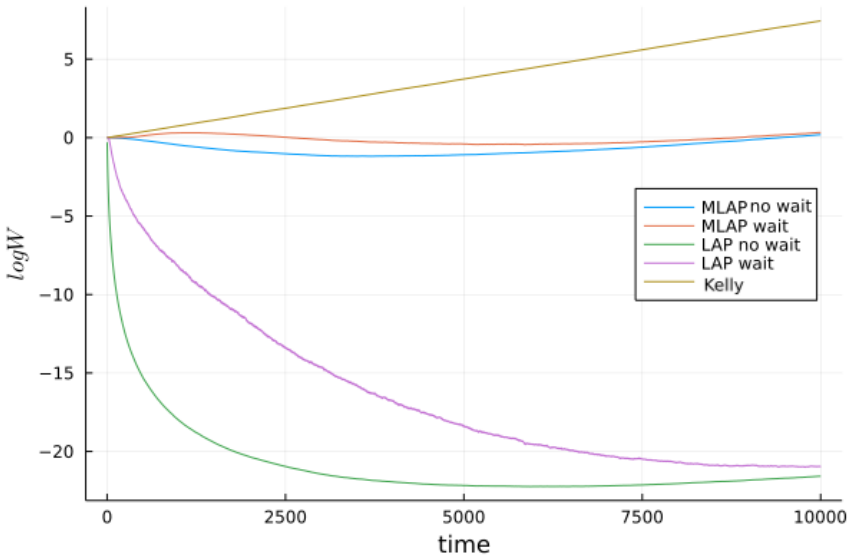


Figure 2.8. $\varepsilon = 0.03$

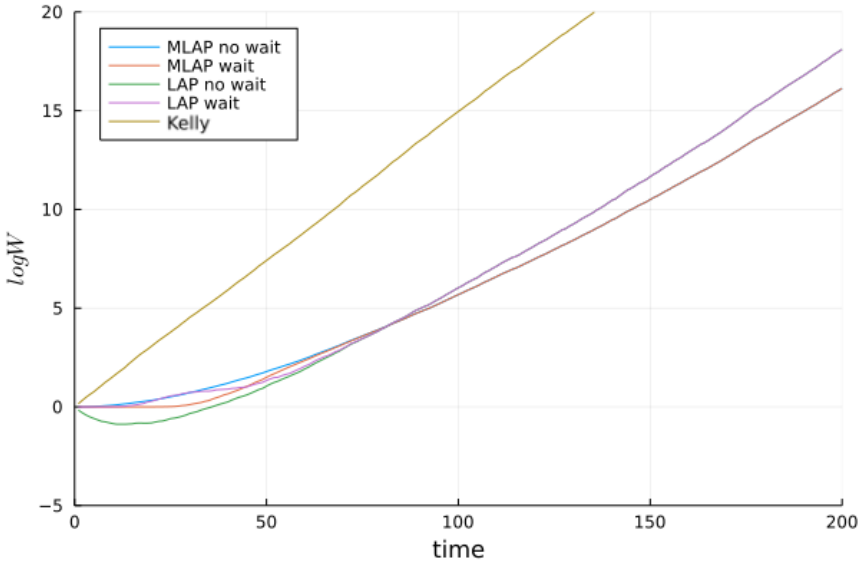


Figure 2.9. $\varepsilon = 0.5$

Figure 2.10. The two graphs represents the $\log W$ obtained with the different strategies, the same as above. ML shows a burn-in phase only in the case $\varepsilon = 0.03$ whereas L shows a burn-in phase always. In the case $\varepsilon = 0.5$ L is preferable to ML whereas in the other case the opposite is true.

## 2.2 Bayesian Model Selection Strategy

We now analyze a Bayesian model selection approach to the problem, using the framework of (2). The idea that motivates the need for such a method is that for small $t$ Laplace estimator can be largely affected by the result of the first few runs. To avoid the problem of fitting noise, it is then worth asking if, for a given sample size, all the states observed in the sample should be distinguished. Instead, only a partition of the sampled states can be resolved, but each of these partitions corresponds to a different model. Using Bayesian model selection it will be possible to identify the partition that maximizes the posterior. This argument should obviously only be applied in an under-sampling regime, whereas out of this case one expects to recover the standard estimator found previously.
Qualitatively one expects this classification of states to be finer and finer as time goes on, up until the sample has enough statistical relevance to justify the distinction of every state. In practice this means that in order for different outcomes to be assigned different probabilities, the frequency with which they occur in the sample must be sufficiently different.

### 2.2.1 Definition of the model

We define $\hat{s} = (s^{(1)}, ..., s^{(t)})$ as the dataset of $t$ observations of the state $s$ of the system up until time $t$, where $s^{(i)}$ is the configuration of the system at time $i$. The number of possible different states is equal to the possible outcomes so it's equal to $M$; this set will be referred as $\mathcal{S}$. We restrict to the case where $t$ is not too much larger than $M$ to fall in the case of under-sampling.
Then, we define a partition $\mathcal{Q}$ of the states as follows

$$\mathcal{Q} = (\mathcal{Q}_1, ..., \mathcal{Q}_N) \qquad \bigcup_{q=1}^{N} \mathcal{Q}_q = \mathcal{S} \tag{2.11}$$

Whereas before $M$ sets each one of cardinality one were considered, now with this model in a single set there can be more states: even if they have different frequencies in the analyzed sample, they may still be associated to the same set and therefore to the same probability parameter $\mu$. Calling $m_q = |\mathcal{Q}_q|$ i.e. the number of states in $\mathcal{Q}_q$, the normalization constraint is now $\sum_{q=1}^{N} m_q \mu_q = 1$. It is also possible to include states that have not yet been sampled. Notice that the most refined partition that is possible to obtain is one in which states with the same frequency have the same probability, just like for Laplace estimator. We will refer to this partition as the $\mathcal{K}$ partition.

As stated before, each partition identifies a model. Defining the frequency of each state $k_s = \sum_{i=1}^{t} \delta_{s,s^{(i)}}$ and the overall frequency of the set

$$K_q = \sum_{s \in \mathcal{Q}_q} k_s \qquad\qquad \sum_q K_q = t \tag{2.12}$$

one is able to write the likelihood of each partition for the dataset

$$P(\hat{s}|\mathcal{Q}) = \int d\vec{\mu} \prod_q \mu^{K_q} P_0^{(\mathcal{Q})}(\vec{\mu}) \tag{2.13}$$

where $P_0^{(\mathcal{Q})}(\vec{\mu})$ is the prior distribution of the parameters given the partition $\mathcal{Q}$

$$P_0^{(\mathcal{Q})}(\vec{\mu}) = \Gamma(aN) \prod_q \frac{m_q^a}{\Gamma(a)} \mu^{a-1} \delta \left( \sum_q m_q \mu_q - 1 \right) \tag{2.14}$$

Therefore one has

$$P(\hat{s}|\mathcal{Q}) = \frac{\Gamma(aN)}{\Gamma(t+aN)} \prod_q \frac{\Gamma(K_q + a)}{m_q^{K_q} \Gamma(a)} \tag{2.15}$$

Then, it is possible to make the choice of a uniform prior, since all the models are considered a-priori equally likely. Therefore, one gets the posterior probability $P(\mathcal{Q}|\hat{s})$. The optimal partition, which will be called $\mathcal{Q}^*$, is then given by the maximization of this posterior, which is analogous to the maximization of the likelihood

$$\mathcal{Q}^* = \arg\max_{\mathcal{Q}} P(\hat{s}|\mathcal{Q}) \tag{2.16}$$

Once this partition is found, the estimator of the probabilities is straightforward to compute. As seen previously in equation 2.1, one has a Dirichlet prior and a multinomial likelihood. The posterior is again

$$P_1^{(\mathcal{Q})}(\vec{\mu}|\hat{s}) = \Gamma(t+aN) \prod_q \frac{m_q^{K_q+a}}{\Gamma(K_q+a)} \mu^{K_q+a-1} \delta \left( \sum_q m_q \mu_q - 1 \right) \tag{2.17}$$

The expected value taken on this distribution yields

$$\mu_q = \frac{1}{m_q} \frac{K_q + a}{t + aN} \tag{2.18}$$

as an estimator for the probability of the state $s \in \mathcal{Q}_q$.

### 2.2.2 Algorithm to find $\mathcal{Q}^*$

To find more easily the optimal partition, one assumes that partitions that put together states with adjacent frequencies are preferable to those that merge states with non adjacent frequencies. Therefore, it is enough to consider models where all the states in the set $\mathcal{Q}_q$ have frequency which is larger than the frequency of any state belonging to the set $\mathcal{Q}_{q'}$ with $q' < q$. Then, it is possible to devise an algorithm to compute the optimal partition of the sample

- Start from the $\mathcal{Q} = \mathcal{K}$ partition
- For each $q = 1, ..., N_{\mathcal{Q}} - 1$ define a new partition obtained by merging the sets $\mathcal{Q}_q$ and $\mathcal{Q}_{q+1}$ and compute the changes in the log-likelihood
- If the largest of this changes is positive and happens merging $q$ and $q + 1$, update the partition merging the $q$ and the $q + 1$ sets; then repeat the previous steps
- If the largest of this changes is negative the current partition is the optimal one

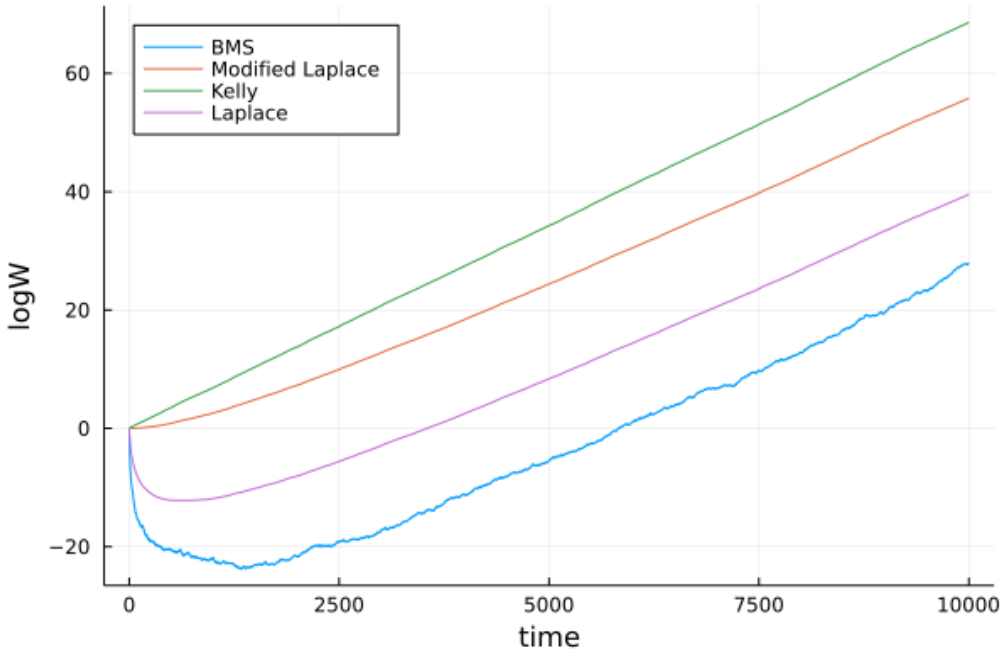


Figure 2.11. Comparison between Kelly's, Laplace, Modified Laplace and Bayesian Model selection strategy. The initial loss phase of BMS is the largest one and affects the large time value of the log-capital. Changing the parameter $\varepsilon$ does not influence this behaviour.

In figure 2.11 one can observe the performance of such strategy. It is clear that despite the model selected maximizes the likelihood, it does not grant the best gain for large times. This is due to the fact that learning the probabilities and maximizing the growth rate are two distinct concepts: being able to do well for the first one does not imply the same for the second one. The numerical simulations have been performed with $\varepsilon = 0.01$ and $M = 10$, although changing these values does not affect visibly the large time behaviour of the log-wealth.

## 2.3 Mean-variance approach

It is possible to device yet another strategy based on the Modified Laplace setting. Instead of being fixed, $\tau$ (from equation 2.6) can change over time. Indeed, one expects that an high value of $\tau$ is to be preferred for small $t$, whereas as time goes on and the probabilities are learnt more accurately, one wants to remove the portion that is bet proportionally to the inverse odd vector, so $\tau$ will be small.
If one bets following Modified Laplace, the doubling rate that will be obtained is

$$m(\mathbf{p}, \hat{\mathbf{p}}) = \sum_x p_x \log\left(\frac{\hat{p}_x}{r_x}\right) \tag{2.19}$$

Using the posterior 2.4 one gets

$$\mathbb{E}\left[m(\mathbf{p}, \hat{\mathbf{p}})\right] = \sum_x \hat{p}_x \log\left(\frac{\hat{p}_x}{r_x}\right) \tag{2.20}$$

One can try and maximise numerically with respect to $\tau$ the expected value of such quantity. This kind of analysis yields the result that the best $\tau$ is 0 at every time. However, if one compares $\mathbb{E}\left[m(\mathbf{p}, \hat{\mathbf{p}})\right]$ and $m(\mathbf{p}, \hat{\mathbf{p}})$ using $\tau = 0$ a substantial difference will be observed, as in figure 2.12: the expected value of the doubling rate is far from its actual value. Therefore, since it is clear that the wrong quantity is being maximized, another approach will be followed.

This strategy yields a bad result because it is too risky to bet proportionally when the probabilities haven't been learned yet. Thus, one may think to use a *mean-variance* approach (6): since the actual utility function is unknown, one approximates it using its first two moments.
Making use of such a method is justified by the fact that in classical finance risk is associated with the statistical uncertainty on the final outcome (11). Therefore, taking into account both the expected value and the variance has the meaning of maximizing the gain and simultaneously reducing the risk.

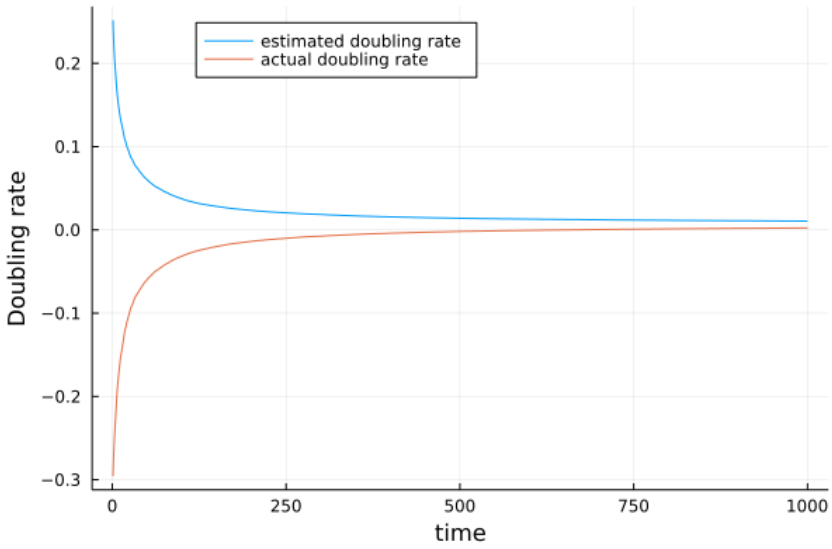


Figure 2.12. Comparison between the best estimated doubling rate and the actual one. The simulations have been performed with $M = 10$, $\varepsilon = 0.03$ and averaged over 500 trials.

It is thus reasonable to try and find the optimal $\tau$ maximizing a linear combination of the expected value of the doubling rate and its variance

$$M\text{-}V\left[m(\mathbf{p},\hat{\mathbf{p}},b)\right] = \mathbb{E}\left[m(\mathbf{p},\hat{\mathbf{p}})\right] - b\mathbb{V}\left[m(\mathbf{p},\hat{\mathbf{p}})\right] =$$
$$= \sum_x \hat{p}_x \log\left(\frac{\hat{p}_x}{r_x}\right) - \frac{b}{t+\tau+M+1}\left[\sum_x \hat{p}_x \log^2\left(\frac{\hat{p}_x}{r_x}\right) - \sum_{x,x'} \hat{p}_x\hat{p}_{x'} \log\left(\frac{\hat{p}_x}{r_x}\right)\log\left(\frac{\hat{p}_{x'}}{r_{x'}}\right)\right] \tag{2.21}$$

In this formula $b$ is a positive parameter since one tries to avoid risk.
Again, one tries to find the optimal $\tau$ numerically and at each step uses this result to bet as in equation 2.6. To avoid fitting noise, one introduces a memory parameter $\lambda$ so that

$$\tau_t^* = \lambda\tau_{t-1}^* + (1-\lambda)\max_\tau(M\text{-}V\left[m(\mathbf{p},\hat{\mathbf{p}}),b\right]) \tag{2.22}$$

This method has been studied for different values of $\varepsilon$, using the optimal parameters $b$ and $\lambda$ in each case. The optimal parameters were determined heuristically.

It is possible to notice from figure 2.13 that only in the first case, with $\varepsilon_1 = 0.02$, the strategy is better than the best Modified Laplace. Note that both the best ML strategy and the parameter $b$ and $\lambda$ have been determined heuristically since we have found no way to determine them analytically.

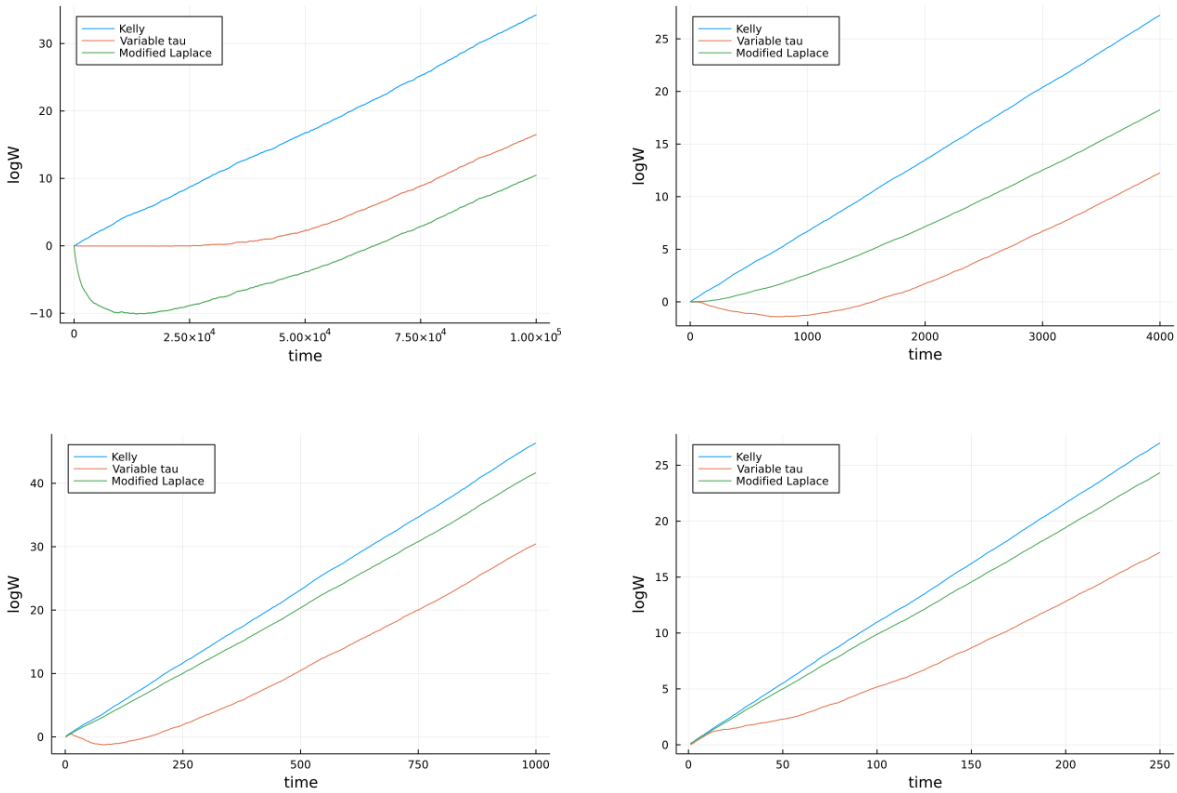


Figure 2.13. Numerical simulations for $\varepsilon_1 = 0.02$, $\varepsilon_2 = 0.1$, $\varepsilon_3 = 0.3$, $\varepsilon_4 = 0.5$. The strategy is preferable to the best ML only in the first case, i.e. when the bookmaker's knowledge of the probabilities is very accurate. The optimal parameter were $b_1 = 0.4$, $b_2 = 0.3$, $b_3 = 0.1$, $b_4 = 0.05$, $\lambda_1 = 0.5$, $\lambda_2 = 0.5$, $\lambda_3 = 0.5$, $\lambda_4 = 0.5$. The results have been obtained averaging over 2000 trials.

# 3

# Conclusions

So far we have described an adaptive version of Kelly's horse race model based on Bayesian inference and we have analysed the loss phase such strategy yields. Indeed, both Laplace and Modified Laplace methods are affected by a considerable decrease of the gambler's capital at the early stages, decrease which is reduced in the second method if the distribution of the odds is informative enough. The portion of the capital that is invested proportionally to the inverse odd vector depends on the parameter $\tau$. An high value of this parameter forces the gambler to play carefully at the beginning, avoiding big losses, but incurring in the risk of missing potential gains when Laplace's estimator starts being accurate. Thus, there exists a trade-off between playing safely and missing revenues and this guarantees the existence of a $\tau_{best}$. However, we were not able to compute analytically this optimal parameter and resorted to heuristic methods to determine it.
This trade-off appeared also in the learning time estimate method, in which the player waited until the estimated $t^*$, which is the time up until the gambler loses, and then played following Laplace/ Modified Laplace. Quite remarkably this strategy performs identically as plain Laplace and plain Modified Laplace. This indicates that these two strategies are indeed the same, since they make use of the same knowledge and operate in the same framework of Bayesian inference.
Subsequently we tried and built an adaptive strategy based on Bayesian Model selection, without any remarkable result. Although it should be an optimal approach to learn the probabilities, clearly the parameter grouping procedure causes inaccuracies that lead to a conspicuous loss. The same happens if one increases $M$ to higher values (more horses): being able to avoid fitting noise in an early phase does not compensate for the intrinsic capital drop of this method.
The mean variance approach is motivated by what is stated above: a big value of $\tau$ is to be preferred when the probabilities are still unknown but after the probabilities are learnt a small value of $\tau$ should be preferred. It seems logical to try and find a time dependent $\tau$, but in doing so one faces another problem, that is finding the optimal function to minimize in order to obtain this parameter. In order to overcome

this obstacle, we defined a functional that depends on the first two moments of the growth rate, at the cost of introducing a new parameter $b$. To avoid fitting noise too much, a memory effect was inserted, at the cost of another parameter $\lambda$. With this prescription, in the case of inverse odd vector highly faithful to the original probability vector, it was possible to get an improved result with respect to the best Modified Laplace. However, also in this case the maximization on $b$ and $\lambda$ was performed heuristically. Furthermore, we were not able to supply these two parameters with any physical meaning.

Ultimately, it is possible to observe that, although the problem faced was simple, we failed in finding an equally simple and elegant solution. We were unsuccessful in fulfilling the initial aim at quantifying the value of the information learned and its relation with the growth rate, as this concept seems to be elusive with this Bayesian inference framework.

# Appendix A

# General solution of Kelly's Game

The functional to perform the maximization on is

$$\mathcal{L} = \sum_i p_i \log\left(1 - B + b_i o_i\right) - \nu\left(\sum_i b_i - B\right) \tag{A.1}$$

Thus, for $x = 1, ..., M$, one has

$$\frac{\partial \mathcal{L}}{\partial b_x} = \frac{p_x o_x}{1 - B + b_x o_x} - \nu = 0 \tag{A.2}$$

$$p_x - \nu \frac{1 - B}{o_x} = \nu b_x \tag{A.3}$$

The equation for $\nu$ then becomes

$$1 - \nu(1 - B)R = \nu B$$
$$\nu = \frac{1}{B + R - BR}$$

Inserting back in A.3 one gets

$$b_x = \left[(1 - B)R + B\right] p_x - \frac{1 - B}{o_x} \tag{A.4}$$

Therefore the doubling rate becomes

$$m^*(\mathbf{p}) = \sum_i p_i \log\Big(\left[(1 - B)R + B\right] p_i o_i\Big) = \log\left(1 - B + \frac{B}{R}\right) + D_{KL}(\mathbf{p}\|\mathbf{r}) \tag{A.5}$$

# Appendix B

# Prior setting in Modified Laplace strategy

In order to assign the proper value to the $a_x$'s, we take into account the form of equation B.1. Using also the fact that the inverse odd vector is known, it is possible to write what follows.

$$\mathbf{r} = (1-\varepsilon)\mathbf{p} + \varepsilon\boldsymbol{\eta} \tag{B.1}$$

The prior on $\mathbf{p}$ can also be expressed as

$$P(\mathbf{p}) \propto \prod_k p_k^{a_k^\tau} \delta\left(\sum_i p_i - 1\right) \prod_k \theta(p_k)\theta(1-p_k) \tag{B.2}$$

Where $\theta(p_k)$ is the Heavyside step function. From here on we will refer to $\prod_k \theta(p_k)$ as $\theta(\mathbf{p})$. As done previously, we set

$$\frac{a_x + 1}{\tau + M} = \mathbb{E}[p_x] \tag{B.3}$$

The expected value has to be taken over the posterior B.4. The likelihood in this case is none other than the distribution of $\boldsymbol{\eta}$

$$P(\mathbf{p}|\mathbf{r}) \propto P(\mathbf{r}|\mathbf{p})P(\mathbf{p}) = P\left(\boldsymbol{\eta} = \frac{\mathbf{r}}{\varepsilon} - \frac{1-\varepsilon}{\varepsilon}\mathbf{p}\right) P(\mathbf{p}) \tag{B.4}$$

Since $\boldsymbol{\eta}$ has to be included in [0,1], the deltas in equation B.4 provide another condition for $\mathbf{p}$ (the first is given by the step functions in B.2). The posterior therefore becomes

$$\begin{aligned} P(\mathbf{p}|\mathbf{r}) &\propto \delta\left(\sum_i p_i - 1\right) \theta(\mathbf{p})\theta(1-\mathbf{p})\theta\left(\frac{\mathbf{r}}{\varepsilon} - \frac{1-\varepsilon}{\varepsilon}\mathbf{p}\right)\theta\left(1 - \frac{\mathbf{r}}{\varepsilon} + \frac{1-\varepsilon}{\varepsilon}\mathbf{p}\right) = \\ &= \delta\left(\sum_i p_i - 1\right) \theta(\mathbf{p})\theta\left(\mathbf{p} - \frac{\mathbf{r}-\varepsilon}{1-\varepsilon}\right)\theta(1-\mathbf{p})\theta\left(\frac{\mathbf{r}}{1-\varepsilon} - \mathbf{p}\right) \end{aligned} \tag{B.5}$$

The meaning of these steps is that knowing $\mathbf{r}$ tells a lot about $\mathbf{p}$ and one can't say anymore its distribution is uniform. Using B.5, one has

$$\mathbb{E}[p_x] = \mathcal{N}\left(\prod_i \int_{\max(0,\frac{r_i-\varepsilon}{1-\varepsilon})}^{\min(1,\frac{r_i}{1-\varepsilon})} dp_i\right) p_x \delta\left(\sum_j p_j - 1\right) \tag{B.6}$$

where

$$\frac{1}{\mathcal{N}} = \left(\prod_i \int_{\max(0,\frac{r_i-\varepsilon}{1-\varepsilon})}^{\min(1,\frac{r_i}{1-\varepsilon})} dp_i\right) \delta\left(\sum_j p_j - 1\right) \tag{B.7}$$

This are incomplete multivaried beta functions. It is however possible to compute this integrals when $\varepsilon$ is small enough to guarantee, for all $i$'s, that

$$v_i = \frac{r_i}{1-\varepsilon} < 1 \qquad\qquad u_i = \frac{r_i - \varepsilon}{1-\varepsilon} > 0 \tag{B.8}$$

In this case the integral becomes

$$\frac{1}{\mathcal{N}} = \left(\prod_i \int_{u_i}^{v_i} dp_i\right) \delta\left(\sum_j p_j - 1\right) \tag{B.9}$$

Now consider the following: if $M-1$ variables take the value of the upper extreme of integration, the last can take the value of the lower extreme and the delta constraint will be fulfilled; the opposite however is not true. The equation B.9 then becomes

$$\frac{1}{\mathcal{N}} = \int_{1-\sum_{i=2}^{M} v_i}^{v_1} dp_1 \int_{1-\sum_{i=3}^{M} v_i - p_1}^{v_2} dp_2 \ldots \int_{1-\sum_{i=1}^{M-2} p_i - v_M}^{v_{M-1}} dp_{M-1} \int_{1-\sum_{i=1}^{M-1} p_i}^{v_M} dp_M \delta\left(\sum_j p_j - 1\right) \tag{B.10}$$

It is then easy to obtain

$$\mathcal{N} = (M-1)!\left(\frac{1-\varepsilon}{\varepsilon}\right)^{M-1} \tag{B.11}$$

and

$$\mathbb{E}[p_x] = \frac{r_x - \frac{\varepsilon}{M}}{1-\varepsilon} \simeq r_x + \varepsilon\left(r_x - \frac{1}{M}\right) \tag{B.12}$$

# Appendix C

# Learning time of Modified Laplace

The starting point is

$$p_x^{MLAP,t+1} = \frac{n_x^t + r_x(\tau + M)}{t + \tau + M} \tag{C.1}$$

For large $t$ the CLT holds and one can write

$$n_x \underset{t\to\infty}{\approx} p_x t + \sqrt{t}\sigma_x z_t \tag{C.2}$$

Where $z_t$ has distribution $\mathcal{N}(0,1)$. We consider rates that are "close" to the p's, meaning that

$$r_x = p_x(1-\varepsilon) + \varepsilon\eta_x \tag{C.3}$$

with $\varepsilon$ small. We thus define $\varepsilon_x$ through

$$\frac{r_x}{p_x} = 1 + \varepsilon_x \tag{C.4}$$

$$1 + \varepsilon_x = \frac{r_x}{p_x} = \frac{p_x(1-\varepsilon) + \varepsilon\eta_x}{p_x} = 1 - \varepsilon + \varepsilon\frac{\eta_x}{p_x} \tag{C.5}$$

$$\varepsilon_x = \varepsilon\left(\frac{\eta_x}{p_x} - 1\right) \tag{C.6}$$

$$\langle\varepsilon_x\rangle = \varepsilon\left(\left\langle\frac{\eta_x}{p_x}\right\rangle - 1\right) = \varepsilon\left(\sum_x p_x\frac{\eta_x}{p_x} - 1\right) = 0 \tag{C.7}$$

$$\langle \varepsilon_x^2 \rangle = \varepsilon^2 \left( \left\langle \frac{\eta_x^2}{p_x^2} \right\rangle - 2 \left\langle \frac{\eta_x}{p_x} \right\rangle + 1 \right) = \varepsilon^2 \left( \sum_x \frac{\eta_x^2}{p_x} - 1 \right) \tag{C.8}$$

These averages are not taken on the distributions $\mathbf{p}$ and $\boldsymbol{\eta}$ but are just averages over the possible outcomes of a race.

$$\begin{aligned}
&\log b_x^{MLAP,t+1} \approx \log \left( p_x t + \sqrt{t}\sigma_x z_t + r_x(\tau + M) \right) - \log\left(t + \tau + M\right) = \\
&= \log p_x + \log\left(t+\tau\right) - \log\left(t+\tau\right) - \log\left(1 + \frac{M}{t+\tau}\right) + \log\left( \frac{t}{t+\tau} + \frac{\sqrt{t}\sigma_x z_t}{p_x(t+\tau)} + \right. \\
&\left. + \frac{\overbrace{r_x}^{1+\varepsilon_x}}{p_x} \left( \frac{\tau + M}{t+\tau} \right) \right) \approx \log p_x - \frac{M}{t+\tau} + \log\left( 1 + \frac{\sqrt{t}\sigma_x z_t}{p_x(t+\tau)} + \left( \frac{(1+\varepsilon_x)M}{t+\tau} \right) + \frac{\varepsilon_x \tau}{t+\tau} \right) \approx \\
&\approx \log p_x - \frac{M}{t+\tau} + \frac{\sqrt{t}\sigma_x z_t}{p_x(t+\tau)} - \frac{1}{2}\left( \frac{\sqrt{t}\sigma_x z_t}{p_x(t+\tau)} \right)^2 + \frac{M}{t+\tau} + \frac{\varepsilon_x \tau}{t+\tau} - \frac{1}{2}\left( \frac{\varepsilon_x \tau}{t+\tau} \right)^2 + \\
&- \frac{\tau\sqrt{t}\sigma_x z_t}{p_x(t+\tau)^2}\varepsilon_x \approx \log p_x + \frac{\sqrt{t}\sigma_x z_t}{p_x(t+\tau)} - \frac{1}{2}\frac{t(1-p_x)z_t^2}{p_x(t+\tau)^2} + \frac{\varepsilon_x \tau}{t+\tau} - \frac{1}{2}\frac{\varepsilon_x^2 \tau^2}{(t+\tau)^2} - \frac{\tau\sqrt{t}\sigma_x z_t}{p_x(t+\tau)^2}\varepsilon_x
\end{aligned} \tag{C.9}$$

We have expanded up to second order in $\varepsilon_x$ and up to first order in $\frac{1}{t+\tau}$. The averages considered in the following have to be intended as averages first over $z_t$ and then over $x$.

$$\begin{aligned}
\langle \log p_x - \log b_x^t \rangle \approx &- \underbrace{\left\langle \frac{\tau\sqrt{t}\sigma_x z_t}{p_x(t+\tau)^2}\varepsilon_x \right\rangle}_{=0} - \underbrace{\left\langle \frac{\sqrt{t}\sigma_x z_t}{p_x(t+\tau)} \right\rangle}_{=0} + \left\langle \frac{1}{2}\frac{t(1-p_x)z_t^2}{p_x(t+\tau)^2} \right\rangle + \\
&- \underbrace{\left\langle \frac{\varepsilon_x \tau}{t+\tau} \right\rangle}_{=0} + \left\langle \frac{1}{2}\frac{\varepsilon_x^2 \tau^2}{(t+\tau)^2} \right\rangle = \frac{M-1}{2}\frac{t}{(t+\tau)^2} + \frac{\tau^2}{2(t+\tau)^2}\varepsilon^2 \left( \sum_x \frac{\eta_x^2}{p_x} - 1 \right)
\end{aligned} \tag{C.10}$$

We use this result to calculate the log capital, since $\langle \log C_t \rangle = D_{KL}(p||r)t - \langle \Delta(t) \rangle$ where $\Delta(t)$ is the regret up to time $t$.

$$\langle \Delta(t) \rangle = \langle \Delta(t_0) \rangle + \left\langle \sum_{i=t_0+1}^{t} \log p_x - \log b_x^i \right\rangle \tag{C.11}$$

Now, having two series in the regret, one can approximate them in the following way

$$\sum_{i=t_0+1}^{t} \frac{i}{(i+\tau)^2} = \sum_{j=t_0+1+\tau}^{t+\tau} \frac{j-\tau}{j^2} = \sum_{j=t_0+1+\tau}^{t+\tau} \frac{1}{j} - \sum_{j=t_0+1+\tau}^{t+\tau} \frac{\tau}{j^2} \approx$$
$$\approx \log \frac{t+\tau}{t_0+1+\tau} - \tau \left( \frac{1}{t_0+1+\tau} - \frac{1}{t+\tau} \right) \quad \text{(C.12)}$$

$$\sum_{i=t_0+1}^{t} \frac{\tau^2}{(i+\tau)^2} = \sum_{j=t_0+1+\tau}^{t+\tau} \frac{\tau^2}{j^2} \approx \tau^2 \left( \frac{1}{t_0+1+\tau} - \frac{1}{t+\tau} \right) \quad \text{(C.13)}$$

$$\langle \Delta(t) \rangle = \langle \Delta(t_0) \rangle + \frac{M-1}{2} \left[ \log \frac{t+\tau}{t_0+1+\tau} - \tau \left( \frac{1}{t_0+1+\tau} - \frac{1}{t+\tau} \right) \right] +$$
$$+ \tau^2 \left( \frac{1}{t_0+1+\tau} - \frac{1}{t+\tau} \right) \frac{\varepsilon^2}{2} \left( \sum_x \frac{\eta_x^2}{p_x} - 1 \right) \quad \text{(C.14)}$$

This expression is still dependent on $\varepsilon$. Luckily this dependence can be removed using the following relation

$$D_{KL}(p||r) = \sum_x p_x \log \left( \frac{p_x}{r_x} \right) = -H(p) - \sum_x p_x \log \left( (1-\varepsilon)p_x + \varepsilon \eta_x \right) =$$
$$= -\sum_x p_x \log \left( 1 - \varepsilon + \varepsilon \frac{\eta_x}{p_x} \right) \approx -\sum_x p_x \varepsilon \left( \frac{\eta_x}{p_x} - 1 \right) + \frac{1}{2} \sum_x p_x \varepsilon^2 \left( \frac{\eta_x^2}{p_x^2} - 2\frac{\eta_x}{p_x} + 1 \right) =$$
$$= \frac{\varepsilon^2}{2} \left( \sum_x \frac{\eta_x^2}{p_x} - 1 \right) \quad \text{(C.15)}$$

This yields the following expression for the average of the log capital:

$$\langle \log C_t \rangle = D_{KL}(p||r)t - \langle \Delta(t) \rangle = D_{KL}(p||r)t - \langle \Delta(t_0) \rangle +$$
$$- \frac{M-1}{2} \left[ \log \frac{t+\tau}{t_0+1+\tau} - \tau \left( \frac{1}{t_0+1+\tau} - \frac{1}{t+\tau} \right) \right] - \tau^2 \left( \frac{1}{t_0+1+\tau} - \frac{1}{t+\tau} \right) D_{KL}(p||r) \quad \text{(C.16)}$$

To find the learning time $t^*$, as shown in the article by Despons et al., we consider

$$0 = \frac{\partial \langle \log C_t \rangle}{\partial t} = D_{KL}(p||r) - \frac{M-1}{2} \left( \frac{1}{t+\tau} - \frac{\tau}{(t+\tau)^2} \right) - \frac{\tau^2}{(t+\tau)^2} D_{KL}(p||r) \quad \text{(C.17)}$$

$$\Rightarrow D_{KL}(p||r) - \frac{M-1}{2(t+\tau)} + \frac{\tau}{(t+\tau)^2}\underbrace{\left(\frac{M-1}{2} - \tau D_{KL}(p||r)\right)}_{=const\times D_{KL}(p||r)} = 0 \tag{C.18}$$

$t + \tau = k$, we solve for $k$

$$k^2 - k\frac{M-1}{2D_{KL}(p||r)} + \tau \times const = 0 \tag{C.19}$$

$$k^* = \frac{M-1}{4D_{KL}(p||r)} \pm \sqrt{\left(\frac{M-1}{4D_{KL}(p||r)}\right)^2 - \tau \times const} \tag{C.20}$$

$$t^* = \frac{M-1}{2D_{KL}(p||r)} - 2\tau \tag{C.21}$$

This result goes along with the simulations.

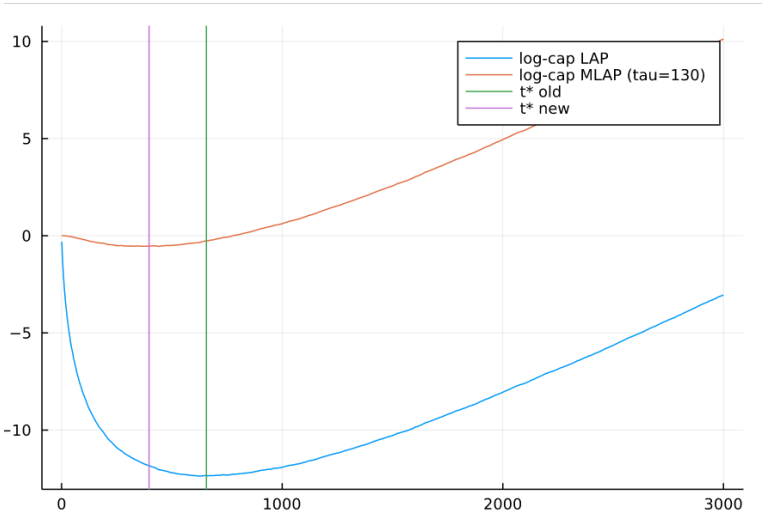


Figure C.1. Plot of log-capital as a function of time for L, ML. The vertical lines represent the two different $t^*$

# References


[1] Despons A, Peliti L, Lacoste D. Adaptive strategies in Kelly's horse races model. Journal of Statistical Mechanics: Theory and Experiment. 2022 sep;2022(9):093405. Available from: https://doi.org/10.1088%2F1742-5468%2Fac8e58.

[2] Haimovici A, Marsili M. Criticality of mostly informative samples: a Bayesian model selection approach. Journal of Statistical Mechanics: Theory and Experiment. 2015 oct;2015(10):P10013. Available from: https://doi.org/10.1088%2F1742-5468%2F2015%2F10%2Fp10013.

[3] Kelly Jr JL. A New Interpretation of Information Rate. Bell System Technical Journal. 1956;35(4):917-26. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1956.tb03809.x.

[4] Thorp E, Kassouf S. Beat the Market: A Scientific Stock Market System. Random House; 1967.

[5] MacLean L, Thorp E, Ziemba W. The Kelly Capital Growth Investment Criterion. World Scientific Publishing; 2011.

[6] Li Calzi M. Lecture notes on FINANCIAL MARKETS. Dipartimento di Matematica Applicata Universit'a "Ca' Foscari" di Venezia; 2002.

[7] Rivoire O, Leibler S. The Value of Information for Populations in Varying Environments. Journal of Statistical Physics. 2011 mar;142(6):1124-66. Available from: https://doi.org/10.1007%2Fs10955-011-0166-2.

[8] Kussell E, Leibler S. Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments. Science. 2005;309(5743):2075-8. Available from: https://www.science.org/doi/abs/10.1126/science.1114383.

[9] Cover T, Thomas J. Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing). USA: Wiley-Interscience; 2006.

[10] Cover T. Universal Portfolios. Mathematical Finance. 1991;1(1):1-29. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9965.1991.tb00002.x.

[11] Bouchaud JP, Potters M. Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management. 2nd ed. Cambridge University Press; 2003.