



HAL
open science

Adaptive strategy in Kelly's horse races model

Armand Despons, Luca Peliti, David Lacoste

► **To cite this version:**

Armand Despons, Luca Peliti, David Lacoste. Adaptive strategy in Kelly's horse races model. *Journal of Statistical Mechanics: Theory and Experiment*, 2022, 2022, pp.093405. 10.1088/1742-5468/ac8e58 . hal-03783347

HAL Id: hal-03783347

<https://hal.science/hal-03783347>

Submitted on 5 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adaptive Strategies in Kelly's Horse Races Model

Armand Despons¹, Luca Peliti² and David Lacoste¹

¹ Gulliver Laboratory, UMR CNRS 7083, PSL Research University, ESPCI, 10 rue Vauquelin, F-75231 Paris Cedex 05, France

² SMRI, Santa Marinella (RM), 00058 Italy

E-mail: armand.despons@espci.fr

May 2022

Abstract. We formulate an adaptive version of Kelly's horse model in which the gambler learns from past race results using Bayesian inference. We characterize the cost of this gambling strategy and we analyze the asymptotic scaling of the difference between the growth rate of the gambler and the optimal growth rate, known as the gambler's regret. We also explain how this adaptive strategy relates to the universal portfolio strategy, and we build improved adaptive strategies in which the gambler exploits the information contained in the bookmaker odds distribution.

Contents

1	Introduction	2
2	Definition of Kelly's model	3
3	Adaptive strategy	4
3.1	Optimality of the strategy	6
3.2	The cost of learning	7
3.3	Extension to correlated races	8
4	Laplace's rule of succession with prior information	9
4.1	The modified Laplace's rule with prior information	9
4.2	Limitations of the modified Laplace's strategy	10
5	Conclusion	11

1. Introduction

In 1956, R. Kelly [1] extended the work of C. Shannon on communication to gambling models. This now classic piece of information theory [2] turned out to have fundamental implications for investment strategies in finance and beyond. In the context of biology, Kelly's paper lead to a clarification of the notion of value of information [3]. This value of information quantifies the quality of the strategies adopted by individuals in a population when faced with unpredictably varying environments [4]. In these conditions, it is sometimes advantageous for individuals to accept a reduction of their short-term reproductive success, in exchange for longer-term risk reduction, a strategy known as *bet-hedging* [5]. Such a strategy is employed for instance by cells to cope with antibiotics or by plants to cope with a fluctuating climate [6].

Kelly's strategy is known to be risky, and for this reason most gamblers use fractional Kelly's strategies, with reduced risk and growth rate [7]. This observation hints at a trade-off between the risk the gambler is ready to take and the average long-term growth rate of his capital. In a recent work, we have studied the properties of this trade-off inspired by ideas in Stochastic Thermodynamics [8]. We then studied another manifestation of this trade-off this time in the context of a biological population with phenotypic switching in a fluctuating environment [9].

In that work, the phenotypic switching rate of individuals in the population is not affected by the state of the environment, which means that no sensing mechanism is present. Naturally, real biological systems are able to sense their environment, extract relevant information from noisy measurements, and adjust their phenotypic response accordingly [10]. An open question in this field is thus the precise quantification of the cost that biological systems must pay for this sensing.

This question motivates the present paper. To evaluate this cost, we go back to Kelly’s model, and we modify it to add some adaptation mechanisms based on Bayesian inference [11]. The adaptive strategy we build in that way has similarities with the one used in universal portfolios [12] and is therefore optimal among all the strategies of this type. We quantify the cost of our adaptive strategy, which in the context of Kelly’s horse races takes the form of a capital loss of the gambler with respect to his/her optimal capital. We derive an asymptotic form for this cost both for uncorrelated and correlated races. We then investigate improved strategies in which the gambler leverages the information contained in the odd distribution chosen by the bookmaker.

2. Definition of Kelly’s model

Let us recall the basic elements of Kelly’s horse race model [1]. A race involves M horses, and is described by a normalized vector of winning probabilities \mathbf{p} , an inverse-odds vector \mathbf{r} and a gambler strategy \mathbf{b} . The latter corresponds to a specific allocation of the gambler’s capital on the M horses: if we denote by C_t the gambler’s capital at time t , the amount of capital invested on horse x reads $b_x C_t$. We further assume that, after each race, the gambler invests his whole capital, i.e., $\sum_{x=1}^M b_x = 1$, always betting a non-zero amount on all horses, i.e., $\forall x \in [1, M] : b_x \neq 0$. The inverse-odds vector \mathbf{r} is set by the bookmaker and is not necessarily normalized, in fact the value of $\sum_x r_x$ controls whether or not the bookmaker is extracting/injecting additional capital: if the sum of the components of the inverse-odds is strictly bigger (resp. strictly lower) than 1 the bookmaker extracts (resp. injects) additional capital. In the following, we focus on the neutral case where the bookmaker neither extracts nor injects capital, which means that the inverse-odds vector is normalized: $\sum_x r_x = 1$. This case is usually referred to as gambling without *track-take* in the horse-racing literature [1, 2].

Thus, the evolution of the gambler’s capital after one race reads:

$$C_{t+1} = \frac{b_{x_t}}{r_{x_t}} C_t, \quad \text{with a probability } p_{x_t}, \quad (1)$$

which implies that the log of the capital, $\log\text{-cap}(t) \equiv \log C_t$, evolves additively:

$$\log\text{-cap}(t) = \sum_{\tau=1}^t \log \left(\frac{b_{x_\tau}}{r_{x_\tau}} \right), \quad (2)$$

where x_τ denotes the index of the winner of the τ -th race and we assumed $\log\text{-cap}(0) = 0$ (i.e. $C_0 = 1$). Since the races are assumed to be independent, the terms $\log(b_{x_\tau}/r_{x_\tau})$ in (2) are independent and identically distributed, and we can use the weak law of the large numbers :

$$\frac{\log\text{-cap}(t)}{t} \xrightarrow[t \rightarrow \infty]{} \mathbb{E} \left[\log \left(\frac{b_x}{r_x} \right) \right] \quad (3)$$

in probability. Then

$$\mathbb{E} \left[\log \left(\frac{b_x}{r_x} \right) \right] \equiv \sum_x p_x \log \left(\frac{b_x}{r_x} \right) = D_{KL}(\mathbf{p} \parallel \mathbf{r}) - D_{KL}(\mathbf{p} \parallel \mathbf{b}), \quad (4)$$

where D_{KL} stands for the Kullback-Leibler divergence [2, sec. 2.3]. From an information theoretic point of view, (3) and (4) imply that the capital of the gambler increases in the long term only if the gambler has a better knowledge of \mathbf{p} than the bookmaker, otherwise it decreases.

It also follows from this analysis that the optimal strategy $\mathbf{b}^{\text{KELLY}} = \mathbf{p}$, called Kelly's strategy [1], overtakes any other strategies in the long-term. Its optimum growth rate is the positive quantity $D_{KL}(\mathbf{p} \parallel \mathbf{r})$.

Note that thanks to the assumption that the inverse-odds vector is normalized, another strategy is possible, namely $\mathbf{b}^{\text{NULL}} = \mathbf{r}$. We have called this strategy the *null strategy* [8], because it yields a constant capital as can be seen from Eqs. (3)-(4). Indeed, in this null strategy, the gambler has exactly the same knowledge as the bookmaker, hence he/her neither loses nor gains capital.

3. Adaptive strategy

Kelly's strategy assumes that the probability vector \mathbf{p} is known, but this, in practice, is never the case. Hence, it is of great interest to build strategies to learn \mathbf{p} , in order to play the optimal betting strategy. In information theory, these strategies are called universal precisely because they do not require any knowledge of \mathbf{p} . From a theoretical point of view, universal strategies are related to lossless source coding [15].

While general theories of history-dependent gambling are available [16], here we want to construct an adaptive strategy specifically for Kelly's problem. We use an estimator based on *Laplace's rule of succession* [17] :

$$b_x^{\text{LAPL}, t+1} \equiv \frac{n_x^t + 1}{t + M}, \quad (5)$$

where the vector \mathbf{n}^t contains the history of all race results up to time t : $\mathbf{n}^t \equiv [n_1^t, \dots, n_M^t]^T$. In this notation, n_x^t represents the number of times the horse x has won among the t previous races, and obviously $\sum_{x=1}^M n_x^t = t$.

This result (5) follows from a standard application of Bayesian inference : one starts from the likelihood of \mathbf{n}^t , which is the multinomial distribution:

$$\mathbb{P}[\mathbf{n}^t \mid \mathbf{b}] = \frac{t!}{n_1^t! \dots n_M^t!} b_1^{n_1^t} \dots b_M^{n_M^t}. \quad (6)$$

The posterior distribution is then given by *Bayes' rule*: $\mathbb{P}[\mathbf{b} \mid \mathbf{n}^t] = \frac{\mathbb{P}[\mathbf{n}^t \mid \mathbf{b}] \mathbb{P}[\mathbf{b}]}{\mathbb{P}[\mathbf{n}^t]}$.

Assuming a uniform prior $\mathbb{P}[\mathbf{b}] \propto \delta(\sum_x b_x - 1)$, the posterior distribution reads:

$$\mathbb{P}[\mathbf{b} \mid \mathbf{n}^t] = \frac{(t + M - 1)!}{n_1^t! \dots n_M^t!} b_1^{n_1^t} \dots b_M^{n_M^t} \delta\left(\sum_x b_x - 1\right). \quad (7)$$

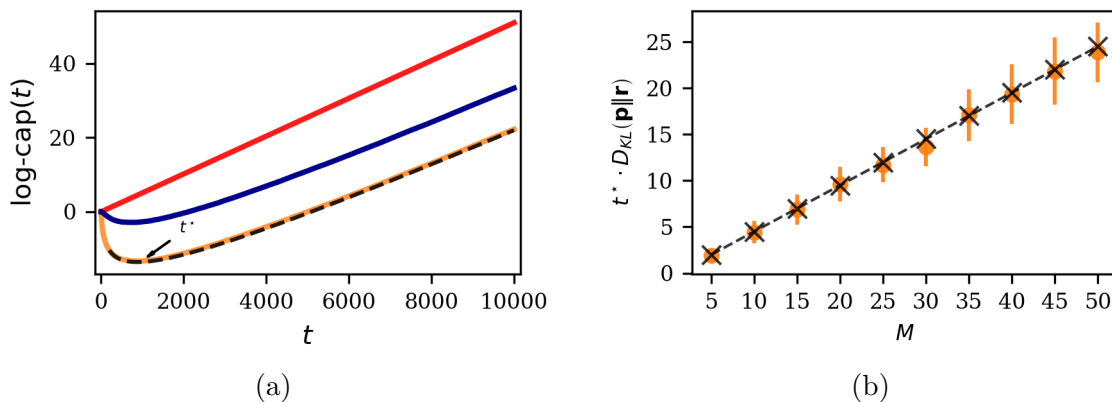


Figure 1: (a) Time evolution of the log-capital of Kelly's strategy (red solid curve), of Laplace's strategy (5) (orange solid curve) and of the modified Laplace's strategy using the odds distribution as prior (25) (blue solid curve). Both adaptive strategies undergo a *learning phase* in which the log-capital decreases during a time t^* , and in the long run they perform as well as Kelly's strategy. Note that the modified Laplace's strategy reduces the initial loss of capital associated to the learning phase. The average log-capital associated with Laplace's strategy (13) is the black dotted curve. In this figure, the log-capital is averaged over 3000 realizations with $M = 10$ and $\epsilon = 0.1$. (b) Linear dependence of the rescaled t^* as function of M , the number of horses. The color dots represent numerical estimates with error bars. The black crosses and the dashed curve represent the analytical prediction of (14). For the each colored dot, we have chosen \mathbf{p} and \mathbf{r} randomly for a fixed M and $\epsilon = 0.1$ and we have averaged over 50 realizations of t^* , itself computed from the average over 30 realizations of the log-capital.

Integrating (7) with respect to all the components except one yields :

$$\mathbb{P} [\mathbf{b}_x | \mathbf{n}^t] = \frac{(t + M - 1)!}{n_x^t! (t - n_x^t + M - 2)!} b_x^{n_x^t} (1 - b_x)^{t - n_x^t + M - 2}. \quad (8)$$

Laplace's rule of succession (5) is then obtained by taking the expectation value of the posterior for each component b_x . Alternate estimators are acceptable such as the one based on the maximum of the posterior, but typically the Laplace estimator performs better specially when one or several horses have a very low probability of winning. In fact, the form of Eq. (5) guarantees that all components of the vector of estimated bets are strictly positive even if the conditional probability given by Eq. (8) may become zero. In particular at $t = 0$, the Laplace estimator is $b_x = 1/M$ which means that the gambler divides his capital uniformly among the M horses, clearly the best thing to do in the absence of any information prior to the first race. The asymptotic regime of b_x^t at large times t can be analyzed using the central limit theorem. Only at large times does the estimator based on Laplace's rule of succession become unbiased. In Appendix A (38), we show that the KL divergence between \mathbf{p} and $\mathbf{b}^{\text{LAPL}, t}$ behaves as:

$$D_{\text{KL}} (\mathbf{p} \| \mathbf{b}^{\text{LAPL}, t}) \underset{t \rightarrow \infty}{\sim} \frac{M - 1}{2t} \underset{t \rightarrow \infty}{\longrightarrow} 0. \quad (9)$$

The same calculation also leads to an explicit expression of the standard deviation of the Laplace estimator. We obtain:

$$\text{STD} [b_x | \mathbf{n}^t] = \sqrt{\frac{b_x^{\text{LAPL}, t} (1 - b_x^{\text{LAPL}, t})}{t + M + 1}}. \quad (10)$$

Here we assumed that the distribution \mathbf{p} is constant in time. Laplace's strategy would still perform well if \mathbf{p} was a slowly varying function of times, i.e. if that distribution was varying on a time scale much longer than that of \mathbf{b}^{LAPL} , which by (9) is given by $(M - 1)/2$.

3.1. Optimality of the strategy

As mentioned in the introduction, Laplace's rule qualifies as universal because no foreknowledge of the distribution of the races \mathbf{p} is needed, instead the distribution is learned over time. In the literature on portfolios, a strategy of this type is known under the name of *universal portfolio*. This strategy is important in practice because it performs asymptotically as well as the best constant balanced portfolios which would have a foreknowledge of the stock prices [12]. Here, we explain precisely in what sense these two strategies are related. In the portfolio case, \mathbf{b} denotes the proportion of wealth invested on a certain day in an ensemble of stocks. The total capital is multiplied after each day by the product of \mathbf{b} with a matrix which describes the increase in the price of all the stocks from one day to the next. Thus, in the portfolio case, the gain of the capital is distributed on all the selected stocks, in contrast to Kelly's model in which a single horse is selected at each race.

To connect this strategy to Kelly's horse race, let us consider this portfolio in the case where only a single stock contributes each day. Let us then denote the index of that stock up to time a t in a vector $\mathbf{x}^t \equiv [x_1, \dots, x_t]^T$, and let $C[\mathbf{b}, t | \mathbf{x}^t]$ be the total capital a portfolio \mathbf{b} would have gained given the history \mathbf{x}^t up to time t . According to Ref. [12], the bet component b_x in the next race is given by

$$b_x^{\text{PORTF}, t+1} \equiv \frac{\int_{\mathbb{R}_+^M} d\mathbf{b} \delta(\sum_{\bar{x}} b_{\bar{x}} - 1) b_x C[\mathbf{b}, t | \mathbf{x}^t]}{\int_{\mathbb{R}_+^M} d\mathbf{b} \delta(\sum_{\bar{x}} b_{\bar{x}} - 1) C[\mathbf{b}, t | \mathbf{x}^t]}. \quad (11)$$

In the case of Kelly's model, where $C[\mathbf{b}, t | \mathbf{x}^t] = \prod_{i=1}^t b_{x_i} o_{x_i}$, the integrals above can be performed explicitly and one recovers exactly (5). Therefore, the adaptive strategy based on Laplace's rule of succession outlined above and the universal portfolio strategy are equivalent when considering a single stock. Since the universal portfolio is known to outperform any other strategy asymptotically, it follows that the adaptive strategy based on Laplace's rule is also optimal among all strategies given the assumptions of the model. Naturally, the universal portfolio has a wider applicability and is more suited to the stock market, because it accounts for distributed gains among many stocks. In this case of distributed gains, interesting and far reaching connections between Kelly's horse race and optimality in game theory have been studied [18].

3.2. The cost of learning

A typical evolution of the log-capital in the adaptive strategy is shown in Figure 1a (orange curve), together with the optimal strategy based on Kelly's criterion (red solid curve). The figure shows that the gambler loses capital at short times because of his/her poor knowledge of the distribution \mathbf{p} , but he/she catches up on long times once the distribution has been learned.

The duration of the burn-in phase is an interesting quantity, which we call t^* and which is shown in Figure 1a. This time is a proxy for the actual loss of capital during the burn-in phase, and for this reason we may regard it as the cost of learning. Clearly, this time depends on the *knowledge* of the bookmaker : if the bookmaker has a perfect knowledge of the races, he can set the odds according to $(\mathbf{r} = \mathbf{p})$, in which case the burn-in time is very large and the log-capital will only decrease. On the contrary, if the bookmaker has a poor knowledge of the system, the inverse odds are far from \mathbf{p} , and in that case, the burn-in time tends to be very short. In practice, we introduce a parameter ϵ and a random vector $\boldsymbol{\eta}$ uniformly distributed in the simplex with $\mathbf{r} = \mathbf{p} + \epsilon (\boldsymbol{\eta} - \mathbf{p})$, such that for $\epsilon = 0$ one has $\mathbf{r} = \mathbf{p}$ and for $\epsilon = 1$, \mathbf{r} is independent of \mathbf{p} .

To quantify the effect more precisely, it is convenient to define t^* as the minimum of the log-capital in the adaptive strategy, and to focus on the difference of log-capital between Kelly's strategy and Laplace's strategy: $\Delta(t) \equiv \log\text{-cap}^{\text{KELLY}}(t) - \log\text{-cap}(t)$. The quantity $\Delta(t)$ has been considered before in the multi-armed bandit problem [19] or in the portfolio theory [12, 20], it is usually called the *regret* associated with the adaptive strategy. The important point here is that the regret is odd-independent: $\Delta(t) = \sum_{i=1}^t [\log p_{x_i} - \log b_{x_i}^{\text{LAPL}, i}]$.

Averaging the regret associated with Laplace's strategy on the race realizations yields (see Appendix A for details on the derivation):

$$\langle \Delta(t) \rangle = \langle \Delta(t_0) \rangle + \frac{M-1}{2} \log \frac{t}{t_0+1}, \quad (12)$$

for $t_0 > \max_x p_x^{-1}$. From (12) one sees that $\langle \Delta(t) \rangle \sim \log t$, a feature already observed for the optimal adaptive strategy in the multi-armed bandit problem [19]. Since we know the long-term behavior of Kelly's strategy from (3), we deduce an expression for the average log-capital of our adaptive strategy:

$$\langle \log\text{-cap}(t) \rangle = D_{\text{KL}}(\mathbf{p} \parallel \mathbf{r}) t - \frac{M-1}{2} \log \frac{t}{t_0+1} - \langle \Delta(t_0) \rangle. \quad (13)$$

This theoretical prediction is shown in Figure 1a (black dotted curve) for a time t_0 respecting the condition for (12). By taking the derivative of (13) with respect to time, one obtains the following estimate for t^* :

$$t^* = \frac{M-1}{2} \frac{1}{D_{\text{KL}}(\mathbf{p} \parallel \mathbf{r})}. \quad (14)$$

As shown in Figure 1b, this time t^* indeed increases linearly with M and is inversely proportional to the KL-divergence between \mathbf{p} and \mathbf{r} , which quantifies the knowledge of

the races by the bookmaker. Since (14) is based on an asymptotic result, namely (13), this expression over-estimates the real t^* but becomes accurate in the limit when t^* is larger than the time t_0 .

At this point, it is important to appreciate that t^* in the above formula still depends on \mathbf{p} , which is the very quantity we are trying to learn in the adaptive strategy, therefore (14) could appear useless. In fact, this is not so, because the gambler can still use $\mathbf{b}^{\text{LAPL}, t}$ in (14) to obtain an estimate for the time t^* . In practice, if the gambler is allowed not to play initially, he/she should wait this time t^* before starting to bet.

3.3. Extension to correlated races

It is straightforward to extend the above strategy to the case of correlated races, provided they are Markovian, in other words, when the results of each race only depends on the result of the previous race. In such an extension, the quantity to be learned is no longer a vector of M components, but an $M \times M$ matrix $\{\mathbf{p}_{x,y}\}$ whose columns $\mathbf{p}_{\odot|y}$ are the conditional probability of winning given that horse y has won in the previous race. Similarly, the odds and the bets are also matrices of size $M \times M$.

Let us assume that the associated Markov chain is irreducible, and thus admits a unique stationary distribution \mathbf{p}^s . The equation governing the evolution of the log-capital is

$$\text{log-cap}(t) \underset{t \rightarrow \infty}{\sim} \left(\sum_{x,y=1}^M p_{x,y} \log \frac{b_{x|y}}{r_{x|y}} \right) t, \quad (15)$$

where $p_{x,y} \equiv p_{x|y} \cdot p_y^s$ and where $r_{x|y}$ are the inverse odds conditioned on the last race's result. Kelly's strategy is again the one that maximizes the long-term growth rate, given the normalization of the conditional bets $b_{x,y}^{\text{KELLY}} = p_{x,y}$ [2].

We now adapt Laplace rule's of succession to take into account correlations:

$$b_{x|y}^{\text{LAPL}, t+1} = \frac{n_{x|y}^t + 1}{\tau_y^t + M}, \quad (16)$$

where $n_{x|y}^t$ and τ_y^t are respectively the number of times horse x wins immediately following a victory of horse y , and the number of times horse y has won irrespective of the previous events.

Qualitatively the behavior for correlated or uncorrelated races are similar, except that the data to be learned in the correlated case lies in a higher dimensional space, and as a result, the learning phase lasts longer and the losses in terms of log-capital are heavier. Indeed, the regret now scales as M^2 while still being logarithmic in time (see Appendix C):

$$\langle \Delta(t) \rangle = \langle \Delta(t_0) \rangle + \frac{M(M-1)}{2} \log \frac{t}{t_0 + 1}. \quad (17)$$

In the end, one obtains the following estimate of the learning time :

$$t^* = \frac{M(M-1)}{2 \left\langle \log \frac{p_{x|y}}{r_{x|y}} \right\rangle}. \quad (18)$$

In fact one could have predicted the behavior in $M(M - 1)$ because in the learning of the $M \times M$ matrix, the last row is determined by the constraint on the normalization.

The extension to Markov races of higher order n is straightforward. Now, the relevant distribution is $p_{x|\mathbf{Y}}$ where \mathbf{Y} store the n previous winners. Thus, the matrix of the conditional probability becomes a $M \times M^n$ matrix. As before, the last row is determined by the normalization of each conditional probability, hence the effective number of degrees of freedom is $M^n(M - 1)$. Therefore, one obtains:

$$\langle \Delta(t) \rangle = \langle \Delta(t_0) \rangle + \frac{M^n(M - 1)}{2} \log \frac{t}{t_0 + 1}, \quad (19)$$

and:

$$t^* = \frac{M^n(M - 1)}{2 \left\langle \log \frac{p_{x|\mathbf{Y}}}{r_{x|\mathbf{Y}}} \right\rangle}. \quad (20)$$

4. Laplace's rule of succession with prior information

4.1. The modified Laplace's rule with prior information

In order to decrease the duration of the learning phase which is responsible for the initial loss in the log capital, a possibility is to use a non-uniform prior in the derivation of Laplace rule. Before starting to bet, the gambler could reasonably assume that the odds chosen by the bookmaker are probably not random, but contain some useful information which can be exploited in a non-uniform prior. To do so, we use a conjugate prior which, in the case of Kelly's horse races, is the multinomial distribution:

$$\mathbb{P}[\mathbf{b}] = \frac{\tau!}{a_1^\tau! \dots a_M^\tau!} b_1^{a_1^\tau} \dots b_M^{a_M^\tau} \delta \left(\sum_x b_x - 1 \right), \quad (21)$$

where the a_x^τ 's are free parameters and $\tau \equiv \sum_x a_x^\tau$. As both the prior and the likelihood are multinomial the two merge into a multinomial posterior:

$$\mathbb{P}[\mathbf{b} | \mathbf{n}^t] \propto b_1^{n_1^t + a_1^\tau} \dots b_M^{n_M^t + a_M^\tau} \delta \left(\sum_x b_x - 1 \right). \quad (22)$$

Taking the expectation of the x -th component b_x with the measure defined by (22) gives a modified version of Laplace's rule of succession for a non-uniform prior :

$$b_x^{\text{md-LAPL}, t+1} = \frac{\tilde{n}_x^{\tilde{t}} + 1}{\tilde{t} + M}, \quad (23)$$

where $\tilde{t} \equiv t + \tau$ and $\tilde{n}_x^{\tilde{t}} \equiv n_x^t + a_x^\tau$. The interpretation of (23) is that this prior is equivalent to taking into account $\tau \in \mathbb{R}$ fictional races before the first one, with a_x^τ being the equivalent of n_x^t during these τ fictive races.

Following the same route as with the uniform prior, one finds this reduced posterior (21):

$$\mathbb{P} [b_x | \mathbf{n}^t] = \frac{(\tilde{t} + M - 1)!}{\tilde{n}_x^{\tilde{t}}! (\tilde{t} - \tilde{n}_x^{\tilde{t}} + M - 2)!} b_x^{\tilde{n}_x^{\tilde{t}}} (1 - b_x)^{\tilde{t} - \tilde{n}_x^{\tilde{t}} + M - 2}. \quad (24)$$

At this point, the parameters of the prior (the a_i^τ 's) need to be specified. Assuming the bookmaker is well informed of the real \mathbf{p} , the gambler should favor an initial bet distribution closer to \mathbf{r} over a random distribution. This can be done by imposing:

$$b_x^{\text{md-LAPL}, t=1} = \frac{a_x^\tau + 1}{\tilde{t} + M} \equiv r_x, \quad (25)$$

which leads to:

$$a_x^\tau = r_x \cdot (\tau + M) - 1. \quad (26)$$

In order for the information in possession of the bookmaker to be exploitable, the prior should remain bounded. Thus, for all components x , the parameters of the reduced prior (24) should be strictly positive. This means :

$$\forall x, \begin{cases} a_x^\tau > 0 \\ \tau - a_x^\tau + M - 2 > 0 \end{cases}. \quad (27)$$

Replacing a_x^τ by its expression given by (26) one finds that the first condition reduces to $\tau > \frac{1}{\min_x r_x} - M$, while the second implies $\tau > \frac{1}{1 - \max_x r_x} - M$. Finally :

$$\tau > \max \left[\frac{1}{\min_x r_x}, \frac{1}{1 - \max_x r_x} \right] - M \geq 0. \quad (28)$$

Since (26) already enforces $\sum_x a_x^\tau = \tau$, τ becomes the only free parameter of our problem. Hence choosing any value of τ fulfilling (28) induces a set of a_x^τ fulfilling (25).

Note that a similar strategy could be built with any other form of prior information. For instance, when this information is contained in a distribution $\boldsymbol{\pi}$, the gambler should bet according to :

$$b_x^{\text{md-LAPL}, t=1} = \frac{a_x^\tau + 1}{\tau + M} = \pi_x. \quad (29)$$

4.2. Limitations of the modified Laplace's strategy

It is important to highlight that introducing a non-uniform prior only benefits the short term behavior of the log-capital, because one can show that the asymptotic regret of the modified Laplace's strategy is unchanged by the new choice of the prior: $\langle \Delta(t) \rangle \sim \frac{M-1}{2} \log t$ (see. Appendix B for the proof). This also implies that the learning time t^* is unchanged in the modified Laplace's strategy.

Further, it is expected that using the inverse-odds as prior information is advantageous only if the latter are more informative than using the uniform bet $1/M$

initially in the original Laplace’s rule. In other words, there is the prerequisite that this inverse-odds distribution is closer to \mathbf{p} than a uniform distribution would be. From Eq. 4, one can see that this means that the original strategy should start with a learning phase with a negative growth rate.

If the original strategy has already a positive growth rate initially, then the bets are already better than what the odds distribution could provide, and in that case, the modified Laplace’s strategy should not be used because it will not bring a benefit over the original Laplace’s strategy. Indeed, using the inverse-odds as prior information results in a strategy *between* Laplace’s original strategy and the null strategy $\mathbf{b}^{\text{NULL}} = \mathbf{r}$, and the null strategy, as we mentioned earlier, has a vanishing capital growth rate. Figure 1a summarizes what we discussed, showing alongside, Laplace’s strategy (orange solid curve) and the modified Laplace’s based on the inverse-odds (blue solid curve) when the bookmaker has an accurate knowledge of \mathbf{p} . We also tested that when the original Laplace strategy has a positive slope initially, the modified Laplace’s strategy can also start with a positive slope, albeit with a smaller value, and therefore the modified Laplace strategy in this case does not bring any benefit.

Finally, we note that the idea of using non-uniform priors has also been used in the litterature on portfolios. The latter is referred as *μ -weighted universal portfolios* [20]. The basic idea is to replace the uniform measure, which was only enforcing the normalization of the bets in (11) by a different arbitrary measure μ , which plays the role of a prior. The authors of this reference used a Dirichlet distribution for μ , which is a continuous measure, analogous to the multinomial prior we used.

5. Conclusion

In this work, we build an adaptive version of Kelly’s horse race model based on Bayesian inference, and we clarify its connection to the theory of universal portfolios. In practice, this strategy suffers a significant drop in the capital of the gambler up to a certain time. We recover the known logarithmic dependence on time for this gambler’s regret, for both correlated and uncorrelated races. We then build an improved adaptive strategy, in which the gambler exploits the information contained in the distribution of the odds of the bookmaker. When this distribution is informative, the cost of the learning phase can be significantly reduced, without affecting the long term growth rate. An interesting question in this context would be to study the limit of the performance of our adaptive strategy using the notion of directed information [21, 22].

As mentioned in the introduction, one motivation for this work was to understand the process of adaptation in biological evolution and in particular its cost. In this correspondance between gambling and biology, the environment plays a similar role as the race results, phenotypic decisions taken by biological systems are the equivalent of the bets, the multiplicative rate of the biological population is the equivalent of the odds, and the population growth rate is the equivalent of the long term capital growth rate. The adaptive strategy which we have described in this paper is a form of learning, by

which individuals leverage the information acquired from observing past environmental states. One benefit of our work is the characterization of this initial learning phase with its specific time scale.

Naturally, the biological problem is vastly more complex than the model we looked at: the information acquired by biological systems from the environment is often partial, corrupted and distorted [23]. The benefit of sensing is significant when the environment is time-dependent and individuals have memory [3]. Further, the growth of a biological population can also act back on the environment. Some of these effects could be taken into account in our model by making odds time-dependent or dependent on the bets [24], extensions which could also be interesting in the context of investment strategies. While adaptation in biology comes with complex costs associated with the making and the maintenance of the sensing/learning mechanisms, the approach of this paper provides a simple and direct way to model adaptation and its cost, which we hope will be useful for future endeavours on this topic.

Acknowledgements

We acknowledge fruitful discussions with R. Pugatch. DL acknowledges support from (ANR-11-LABX-0038, ANR-10-IDEX-0001-02).

Appendix A

Average log-capital behavior for uncorrelated races

Each component of \mathbf{n}^t is given by the sum of independent R.V.:

$$n_x^t = \sum_{i=1}^t (n_x^i - n_x^{i-1}) \equiv \sum_{i=1}^t w_x^i, \quad (30)$$

where $w_x^i = 1$ if x wins the race i (with probability p_x), 0 if it loses (with probability $(1 - p_x)$). One immediately sees that: $\mathbb{E}_W [w_x] = p_x$ and $\sigma_x^2 \equiv \mathbb{E}_W [(w_x - \mathbb{E}_W [w_x])^2] = p_x(1 - p_x)$. Hence the application of the central limit theorem (CLT) gives:

$$n_x^t \underset{t \rightarrow \infty}{\approx} p_x t + \sqrt{t} \sigma_x \cdot z_t, \quad (31)$$

where z_t is the realization at time t of Z , a standard Gaussian R.V. of mean 0 and variance 1. Note that the leading order of (31) ensures that Laplace's rule converges to p_x when the number of race increases. Plugging the expression of n_x^t obtained with the CLT in $\log b_{x_{t+1}}^{\text{LAPL}, t+1}$ reads:

$$\log b_{x_{t+1}}^{\text{LAPL}, t+1} \approx \log p_x + \frac{1}{\sqrt{t}} \frac{\sigma_x}{p_x} z_{t+1} - \frac{1}{2} \left(\frac{\sigma_x}{p_x} z_{t+1} \right)^2 + \frac{1}{p_x t} - \frac{M}{t}. \quad (32)$$

The expansion (32) of the log will be valid when $t > \max \left[\frac{1}{\min_x p_x}, M \right]$. Because the sum of the components of \mathbf{p} is equal to 1, $\min_x p_x \leq \frac{1}{M}$. The condition of the expansion is then $t > \min_x p_x^{-1}$. Hence:

$$\langle \log p_x - \log b_x^{\text{LAPL}, t+1} \rangle \approx \left\langle \frac{p_x M - 1}{p_x t} \right\rangle + \left\langle \frac{1}{2} \left(\frac{\sigma_x}{p_x} z_{t+1} \right)^2 \right\rangle - \left\langle \frac{1}{\sqrt{t}} \frac{\sigma_x}{p_x} z_{t+1} \right\rangle. \quad (33)$$

The ensemble average is defined here as: $\langle (\cdot) \rangle \equiv \mathbb{E}_X [\mathbb{E}_Z [(\cdot)] | \mathbf{p}] = \sum_x p_x \cdot \mathbb{E}_Z [(\cdot)]$. Following this definition (33) reduces to:

$$\langle \log p_x - \log b_x^{\text{LAPL}, t+1} \rangle \approx \frac{M - 1}{2t}. \quad (34)$$

As the regret is an additive quantity, it can be split:

$$\langle \Delta(t) \rangle = \langle \Delta(t_0) \rangle + \left\langle \sum_{i=t_0+1}^t \log p_x - \log b_x^{\text{LAPL}, i} \right\rangle. \quad (35)$$

Our previous result on the average discrepancy between Kelly's strategy (34) rests upon the expansion of the logarithm (32), so t_0 should be chosen such that the latter is valid. Then taking $t_0 > \min_x p_x^{-1}$ one could insert (34) in (35):

$$\langle \Delta(t) \rangle \approx \langle \Delta(t_0) \rangle + \frac{M - 1}{2} \sum_{i=t_0+1}^t \frac{1}{i}. \quad (36)$$

The terms of the harmonic series are known to diverge in a logarithmic fashion, hence we recover the expression given in the main text:

$$\langle \Delta(t) \rangle \approx \langle \Delta(t_0) \rangle + \frac{M - 1}{2} \log \frac{t}{t_0 + 1}. \quad (37)$$

Finally noticing that:

$$D_{\text{KL}}(\mathbf{p} \| \mathbf{b}^{\text{LAPL}, t}) = \langle \log p_x - \log b_x^{\text{LAPL}, t} \rangle, \quad (38)$$

then (34) provides (9) of the main text for the KL-divergence between Laplace's estimate and \mathbf{p} .

Appendix B

Long-run equivalence between $\mathbf{b}^{\text{md-LAPL}}$ and \mathbf{b}^{LAPL}

One can proceed in the same way as in the case of pseudo-Laplace's estimate to find the discrepancy from Kelly's estimate :

$$\log p_x - \log b_x^{\text{md-LAPL}, t+1} \approx \frac{\tau + M}{t} + \frac{1}{2t} \frac{1 - p_x}{p_x} z_t^2 - \sqrt{\frac{1 - p_x}{p_x t}} z_t - \frac{a_x^\tau + 1}{p_x t}. \quad (39)$$

As the average of the last term in (39) is $\left\langle \frac{a_x^r + 1}{p_x \cdot t} \right\rangle = \sum_x p_x \cdot \frac{a_x^r + 1}{p_x t} = \frac{\tau + M}{t}$, the average discrepancy from Kelly's estimate is asymptotically the same for a pseudo and a classical Laplace's estimate:

$$\langle \log p_x - \log b_x^{\text{md-LAPL}, t+1} \rangle \approx \frac{M-1}{2t}. \quad (40)$$

This is reasonable because in the long run the information contained in the prior becomes negligible with respect to the information contained in the likelihood.

Appendix C

Proof of the equivalence between \mathbf{b}^{LAPL} and $\mathbf{b}^{\text{PORTF}}$

Keeping in mind that:

$$\int_{\mathbb{R}_+^M} d\mathbf{b} \delta \left(\sum_{\tilde{x}} b_{\tilde{x}} - r \right) \prod_{\tilde{x}=1}^M b_{\tilde{x}}^{n_{\tilde{x}}} = \frac{\prod_{\tilde{x}} n_{\tilde{x}}!}{(\sum_{\tilde{x}} n_{\tilde{x}} + M - 1)!} r^{\sum_{\tilde{x}} n_{\tilde{x}} + M - 1}, \quad (41)$$

where $r \in \mathbb{R}_+$ and $\forall x, n_x \in \mathbb{R}_+$ (for any real number a , we will write $a! \equiv \Gamma(a + 1)$), we can perform an explicit computation of (11) with the assumption of Kelly's model: (11):

$$\begin{aligned} \int_{\mathbb{R}_+^M} d\mathbf{b} \delta \left(\sum_{\tilde{x}} b_{\tilde{x}} - 1 \right) b_x C[\mathbf{b}, t | \mathbf{x}^t] &= O \int_{\mathbb{R}_+^M} d\mathbf{b} \delta \left(\sum_{\tilde{x}} b_{\tilde{x}} - 1 \right) \prod_{\tilde{x}} b_{\tilde{x}}^{n_{\tilde{x}}^t + \delta_{x, \tilde{x}}} \\ &= O \frac{(n_x^t + 1)! \prod_{\tilde{x} \neq x} n_{\tilde{x}}^t!}{(t + M)!}, \end{aligned} \quad (42)$$

where $O \equiv \prod_{x=1}^M o_x^{n_x^t}$. The denominator is computed similarly:

$$\int_{\mathbb{R}_+^M} d\mathbf{b} \delta \left(\sum_{\tilde{x}} b_{\tilde{x}} - 1 \right) C[\mathbf{b}, t | \mathbf{x}^t] = O \frac{\prod_{\tilde{x}} n_{\tilde{x}}^t!}{(t + M - 1)!}. \quad (43)$$

Taking the quotient leads to the claimed equivalence between universal portfolio and Laplace's estimate:

$$b_x^{\text{PORTF}, t+1} = \frac{n_x^t + 1}{t + M} = b_x^{\text{LAPL}, t+1}. \quad (44)$$

Appendix D

Average log-capital behavior for markovian races

The CLT on the components of \mathbf{n}^t should be modified to take into account correlations. Indeed correlations add a new random variable τ_y^t , defined as the total number of time horse y won during the t past races. τ_y^t could also be expressed via the CLT and the expansion of $n_{x|y}^t$ now involves τ_y^t instead of t :

$$\begin{aligned} n_{x|y}^t &\approx p_{x|y} \tau_y^t + \sqrt{\tau_y^t} \sigma_{x|y} \cdot z_t \\ \tau_y^t &\approx p_y^s t + \sqrt{t} \sigma_y \cdot w_t \end{aligned}, \quad (45)$$

where $\sigma_{x|y} \equiv p_{x|y}(1 - p_{x|y})$ (and similarly for σ_y) and w_t, z_t are two independent realizations of a standard normal random variable. Note in addition that, as our underlying Markov chain is irreducible, the steady-state probability of winning \mathbf{p}^s is well defined. Expanding the square root yields:

$$\sqrt{\tau_y^t} \approx \sqrt{p_y^s t} + \frac{1}{2} \frac{\sigma_y w_t}{\sqrt{p_y^s}}. \quad (46)$$

The logarithm of our estimate is given by:

$$\log b_{x|y}^{\text{LAPL}, t+1} = \log(n_{x|y}^t + 1) - \log(\tau_y^t + M). \quad (47)$$

In what comes next we will denote (\diamond) terms that are of null average. The first terms of (47) reads:

$$\log(n_{x|y}^t + 1) \approx \log p_{x|y} + \log p_y^s t + \frac{1}{p_y^s p_{x|y} t} - \frac{1}{2t} \left(\frac{\sigma_y^2 w_{t+1}^2}{p_y^s} + \frac{\sigma_{x|y}^2 z_{t+1}^2}{p_{x|y}^2 p_y^s} \right) + (\diamond). \quad (48)$$

While the expansion of the second term of (47) yields:

$$\log(\tau_y^t + M) \approx \log p_y^s t + \frac{M}{p_y^s t} - \frac{1}{2} \frac{\sigma_y^2 w_{t+1}^2}{\sqrt{t} p_y^s} + (\diamond). \quad (49)$$

The expansions (48) (49) are valid provided t follows $t > \max[t_1, t_2]$ with:

$$\begin{aligned} t_1 &\equiv \max_{x,y} \left[\frac{1}{p_y^s p_{x|y}} \right] \\ t_2 &\equiv \max_{x,y} \left[\frac{1}{2} \frac{1}{p_y^s} \sqrt{\frac{1 - p_y^s}{p_y^s}} \sqrt{\frac{1 - p_{x|y}}{p_{x|y}}} \right]. \end{aligned} \quad (50)$$

Note that the condition to expand $\sqrt{\tau_y^t}$ in the CLT of $n_{x|y}^t$ are automatically satisfied for any t fulfilling (50).

The ensemble average in the case of correlated races is defined by: $\langle (\cdot) \rangle \equiv \sum_{x,y} p_{x,y} \cdot \mathbb{E}_{Z,W}[(\cdot)] = \sum_{x,y} p_y^s p_{x|y} \cdot \mathbb{E}_{Z,W}[(\cdot)]$. Applying this definition reads:

$$\left\langle \log p_{x|y} - \log b_{x|y}^{\text{LAPL}, t+1} \right\rangle \approx \frac{M(M-1)}{2t}. \quad (51)$$

Summing (51) from a t_0 that satisfies the condition (50) gives:

$$\langle \Delta(t) \rangle \approx \langle \Delta(t_0) \rangle + \frac{M(M-1)}{2} \log \frac{t}{t_0 + 1}. \quad (52)$$

Appendix E Arbitrary long correlations

The case of longer correlations is similar to classical Markov races. Indeed let's call n the length of the correlation, the realization at time t of the R.V. X depends now

on $\mathbf{Y}_t \equiv [x_t, x_{t-1}, \dots, x_{t-n+1}]$. As for $t \geq n$, $\mathbb{P}[x_{t+1} | \mathbf{Y}_t \mathbf{Y}_{t-1} \dots \mathbf{Y}_n] = \mathbb{P}[x_{t+1} | \mathbf{Y}_t] \equiv p_{x_{t+1} | \mathbf{Y}_t}$, hence we recover a Markovian process. Then the CLT is expressed as before:

$$\begin{aligned} n_{x|\mathbf{Y}}^t &\approx p_{x|\mathbf{Y}} \tau_{\mathbf{Y}}^t + \sqrt{\tau_{\mathbf{Y}}^t} \sigma_{x|\mathbf{Y}} \cdot z_t \\ \tau_{\mathbf{Y}}^t &\approx p_{\mathbf{Y}}^s t + \sqrt{t} \sigma_{\mathbf{Y}} \cdot w_t \end{aligned} \quad (53)$$

Doing the exact same steps as above one finds that :

$$\log p_{x|\mathbf{Y}} - \log b_{x|\mathbf{Y}}^{\text{LAPL}, t+1} \approx \frac{1}{p_{x|\mathbf{Y}} p_{\mathbf{Y}}^s t} - \frac{M}{p_{\mathbf{Y}}^s t} - \frac{1}{2t} \frac{\sigma_{x|\mathbf{Y}}^2 z_{t+1}^2}{p_{x|\mathbf{Y}}^2 p_{\mathbf{Y}}^s} + (\diamond), \quad (54)$$

again (\diamond) denotes terms that are of null average. The ensemble average is defined as : $\langle (\cdot) \rangle \equiv \sum_{x, \mathbf{Y}} p_{\mathbf{Y}}^s p_{x|\mathbf{Y}} \cdot \mathbb{E}_{Z, W} [(\cdot)]$. Applying this definition, remembering that: $\sum_{x, \mathbf{Y}} 1 = M^{n+1}$ and $\sum_{x, \mathbf{Y}} p_{x|\mathbf{Y}} = M^n$, reads:

$$\left\langle \log p_{x|\mathbf{Y}} - \log b_{x|\mathbf{Y}}^{\text{LAPL}, t+1} \right\rangle \approx \frac{M^n(M-1)}{2t}. \quad (55)$$

Summing (55) yields the expression for the regret for arbitrary correlated races:

$$\langle \Delta(t) \rangle = \langle \Delta(t_0) \rangle + \frac{M^n(M-1)}{2} \log \frac{t}{t_0 + 1}, \quad (56)$$

which is valid for a t_0 following the same constraint (50) replacing $p_{x|y}$ and p_y^s by $p_{x|\mathbf{Y}}$ and $p_{\mathbf{Y}}^s$ respectively.

References

- [1] Kelly J L 1956 *IRE Trans. Inf. Theory* **2** 185–189
- [2] Cover T M and Thomas J A 2005 *Elements of Information Theory* (Wiley Interscience)
- [3] Rivoire O and Leibler S 2011 *J. Stat. Phys.* **142** 1124–1166
- [4] Bergstrom C T and Lachmann M 2004 *IEEE Inf. Theory Work. - Proceedings, ITW* 50–54
- [5] Levien E, Min J, Kondev J and Amir A 2020 *Rep. Prog. Phys.* **84** 116601
- [6] Venable D L 2007 *Ecology* **88** 1086–1090
- [7] MacLean L C, Thorp E O and Ziemba W T 2011 *The Kelly capital growth investment criterion: Theory and practice* vol 3 (Word Scientific)
- [8] Dinis L, Unterberger J and Lacoste D 2020 *Europhys. Lett.* **131** 1–23
- [9] Dinis L, Unterberger J and Lacoste D 2022 *J. Stat. Mech. Theory Exp.* **2022** 053503
- [10] Kussell E and Leibler S 2005 *Science (80-.)*. **309** 2075–2078
- [11] Perkins T J and Swain P S 2009 *Mol. Syst. Biol.* **5** 326
- [12] Cover T M 1991 *Math. Financ.* **1** 1–29
- [13] Vinkler D A, Permuter H H and Merhav N 2016 *J. Stat. Mech. Theory Exp.* **2016** 43403
- [14] Still S, Sivak D A, Bell A J and Crooks G E 2012 *Phys. Rev. Lett.* **109** 120604
- [15] Merhav N and Feder M 1998 *IEEE Trans. Inf. Theory* **44** 2124–2147
- [16] Hirono Y and Hidaka Y 2015 *J. Stat. Phys.* **161** 721–742
- [17] Jaynes E T 2003 *Probability theory: The logic of science* (Cambridge university press)
- [18] Pugatch R, Barkai N and Tlusty T 2013 Asymptotic Cellular Growth Rate as the Effective Information Utilization Rate (*Preprint*)
- [19] Lai T L and Robbins H 1985 *Adv. Appl. Math.* **6** 4–22

- [20] Cover T M and Ordentlich E 1996 *IEEE Trans. Inf. Theory* **42** 348–363
- [21] Permuter H H, Kim Y H and Weissman T 2011 *IEEE Trans. Inf. Theory* **57** 3248–3259
- [22] Kobayashi T Y and Sughiyama Y 2019 *Entropy* **21** 1002
- [23] Mitsokapas E and Harris R J 2022 *Phys. A Stat. Mech. its Appl.* **593** 126762
- [24] Giunta G 2020 *Spatio-Temporal Organization of Enzymatic Reactions* Ph.D. thesis Technische Universität München