

Chapter 1

Some applications of first-passage ideas to finance

Rémy Chicheportiche, Jean-Philippe Bouchaud

Capital Fund Management,
23-25, rue de l'Université, 75 007 Paris

Many problems in finance are related to first passage times. Among all of them, we chose three on which we contributed personally. Our first example relates Kolmogorov-Smirnov like goodness-of-fit tests, modified in such a way that tail events and core events contribute equally to the test (in the standard Kolmogorov-Smirnov, the tails contribute very little to the measure of goodness-of-fit). We show that this problem can be mapped onto that of a random walk inside moving walls. The second example is the optimal time to sell an asset (modelled as a random walk with drift) such that the sell time is as close as possible to the time at which the asset reaches its maximum value. The last example concerns optimal trading in the presence of transaction costs. In this case, the optimal strategy is to wait until the predictor reaches (plus or minus) a threshold value before buying or selling. The value of this threshold is found by mapping the problem onto that of a random walk between two walls.

1. Introduction

Quantitative finance is the bounty land of statistics and probabilities. Bachelier proposed to model price paths as random walks in 1900, but his amazingly creative work was forgotten until the sixties, when Samuelson, Black & Scholes revived the Brownian motion framework, which is now the cornerstone of modern mathematical finance. This is probably unfortunate, because the continuous time, Gaussian random walk misses most of the important “stylized facts” of financial markets — fat (power-law) tails, intermittency and long memory, etc. Many of the results that hold true for a Gaussian process go awry in reality — for example, the well-known perfect hedge of Black-Scholes, that would enable one to sell option contracts without any risk, is a figment of the very specific assumptions of the model. It is all too easy to get carried away with the beauty of a mathematical model, and forget that it does not bear any relation with reality. This is of special concern in the case of financial markets, where inadequate models can contribute (and have contributed) to systemic risks [1].

This is however not to say that probabilistic methods are useless in this context. Quite on the contrary, empirically motivated, faithful models do help in controlling risks better and pricing derivative contracts more accurately. Many questions that are relevant in practice can be addressed. Some of them are directly related to first passage time problems, which is our topic here. For example, one might be interested to invest in financial markets with a profit objective, that would allow another project to be financed. What is the distribution of the time one should wait until this profit is reached? Conversely, one might be worried about the default of a company or a bank. This is often modeled as the first passage time when a random walk (the value of the asset) goes below a certain threshold (the equity) [2]. Another example is that of “barrier options”, which disappear when the price of the underlying hits a certain predefined value (the barrier); a related issue is the early exercise of so-called American options, when the option first reaches a value where it is optimal to exercise and cash the current pay-off rather than let the option run to maturity [3].

We review here three examples of the use of first passage ideas in finance: the design of a goodness-of-fit test whose law is the survival probability of a process before hitting a barrier, the optimal time to sell an asset, and the optimal value of the price threshold at which the expected benefit outweighs linear transaction costs. Extensions and open questions related to these three problems are briefly discussed in the conclusion.

2. Weighted Kolmogorov-Smirnov tests and first passages [4]

Our first example concerns goodness-of-fit (GoF) testing, which is ubiquitous in all fields of science and engineering. This is the problem of testing whether a null-hypothesis theoretical probability distribution is compatible with the empirical probability distribution of a sample of observations. GoF tests are designed to assess quantitatively whether a sample of N observations can statistically be seen as a collection of N realizations of a given probability law, or whether two such samples are drawn from the same hypothetical distribution.

The best known theoretical result is due to Kolmogorov and Smirnov (KS) [5, 6], and has led to the eponymous statistical test for an *univariate* sample of *independent* draws. The major strength of this test lies in the fact that the asymptotic distribution of its test statistics is completely independent of the null-hypothesis cdf.

Several specific extensions have been studied (and/or are still under scrutiny), including: different choices of distance measures, multivariate samples, investigation of different parts of the distribution domain, dependence in the successive draws (which is particularly important for financial applications), etc.

This class of problems has a particular appeal for physicists since the works of Doob [7] and Khmaladze [8], who showed how GoF testing is related to stochastic processes. Finding the law of a test amounts to computing a survival probability in a diffusion system. In a Markovian setting, this is often achieved by treating a Fokker-Planck problem, which in turn maps into a Schrödinger equation for a particle in a certain potential confined by walls.

2.1. Empirical cumulative distribution and its fluctuations

Let \mathbf{X} be a random vector of N independent and identically distributed variables, with marginal cumulative distribution function (cdf) F . One realization of \mathbf{X} consists of a time series $\{x_1, \dots, x_n, \dots, x_N\}$ that exhibits no persistence (see Ref. [9] when some non trivial dependence is present). The empirical cumulative distribution function

$$F_N(x) = \frac{1}{N} \sum_{n=1}^N \mathbb{1}_{\{X_n \leq x\}} \quad (1)$$

converges to the true CDF F as the sample size N tends to infinity. For finite N , the expected value and fluctuations of $F_N(x)$ are

$$\begin{aligned}\mathbb{E}[F_N(x)] &= F(x), \\ \text{Cov}(F_N(x), F_N(x')) &= \frac{1}{N} [F(\min(x, x')) - F(x)F(x')].\end{aligned}$$

The rescaled empirical CDF

$$Y_N(u) = \sqrt{N} [F_N(F^{-1}(u)) - u] \quad (2)$$

measures, for a given $u \in [0, 1]$, the difference between the empirically determined cdf of the X 's and the theoretical one, evaluated at the u -th quantile. It does not shrink to zero as $N \rightarrow \infty$, and is therefore the quantity on which any statistics for GoF testing is built.

Limit properties

One now defines the process $Y(u)$ as the limit of $Y_N(u)$ when $N \rightarrow \infty$. According to the Central Limit Theorem, it is Gaussian and its covariance function is given by:

$$I(u, v) = \min(u, v) - uv, \quad (3)$$

which characterizes the so-called Brownian bridge, i.e. a Brownian motion $Y(u)$ such that $Y(u=0) = Y(u=1) = 0$. Interestingly, the function F does not appear in Eq. (3) anymore, so the law of any functional of the limit process Y is independent of the law of the underlying finite size sample. This property is important for the design of *universal* GoF tests.

Norms over processes

In order to measure a limit distance between distributions, a norm $\|\cdot\|$ over the space of continuous bridges needs to be chosen. Typical such norms are the norm-2 (or 'Cramer-von Mises' distance)

$$\|Y\|_2 = \int_0^1 Y(u)^2 du,$$

as the bridge is always integrable, or the norm-sup (also called the Kolmogorov distance)

$$\|Y\|_\infty = \sup_{u \in [0,1]} |Y(u)|,$$

as the bridge always reaches an extremal value.

Unfortunately, both these norms mechanically overweight the core values $u \approx 1/2$ and disfavor the tails $u \approx 0, 1$: since the variance of $Y(u)$ is zero at both extremes and maximal in the central value, the major contribution to $\|Y\|$ indeed comes from the central region and not from the tails. To alleviate this effect, in particular when the GoF test is intended to investigate a specific region of the domain, it is preferable to introduce additional weights and study $\|Y\sqrt{\psi}\|$ rather than $\|Y\|$ itself. Anderson and Darling show in Ref. [10] that the solution to the problem with the Cramer-von Mises norm and arbitrary weights ψ is obtained by spectral decomposition of the covariance kernel. They design an eponymous test [11] with the specific choice of $\psi(u) = 1/I(u, u)$ equal to the inverse variance, which equi-weights all quantiles of the distribution to be tested. We analyze here the case of the same weights but with the Kolmogorov distance.

So again $Y(u)$ is a Brownian bridge, i.e. a centered Gaussian process on $u \in [0, 1]$ with covariance function $I(u, v)$ given in Eq. (3). In particular, $Y(0) = Y(1) = 0$ with probability equal to 1, no matter how distant F is from the sample cdf around the core values. In order to put more emphasis on specific regions of the domain, let us weight the Brownian bridge as follows: for given $a \in]0, 1[$ and $b \in [a, 1[$, we define

$$\tilde{y}(u) = y(u) \cdot \begin{cases} \sqrt{\psi(u)}, & a \leq u \leq b \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

We will characterize the law of the supremum $K(a, b) \equiv \sup_{u \in [a, b]} |\tilde{y}(u)|$:

$$\mathcal{P}_<(k|a, b) \equiv \mathbb{P}[K(a, b) \leq k] = \mathbb{P}[|\tilde{y}(u)| \leq k, \forall u \in [a, b]].$$

2.2. The equi-weighted Brownian bridge: Kolmogorov-Smirnov

In the case of a constant weight, corresponding to the classical KS test, the probability $\mathcal{P}_<(k; 0, 1)$ is well defined and has the well known KS form [5]:

$$\mathcal{P}_<(k; 0, 1) = 1 - 2 \sum_{n=1}^{\infty} (-1)^{n-1} e^{-2n^2 k^2}, \quad (5)$$

which, as expected, grows from 0 to 1 as k increases. The value k^* such that this probability is 95% is $k^* \approx 1.358$ [6]. This can be interpreted as follows: if, for a data set of size N , the maximum value of $|Y_N(u)|$ is larger than ≈ 1.358 , then the hypothesis that the proposed distribution is a “good fit” can be rejected with 95% confidence.

Diffusion in a cage with fixed walls

The Brownian bridge Y is nothing else than a Brownian motion with imposed terminal condition, and can be written as $Y(u) = X(u) - uX(1)$ where X is a Brownian motion. The survival probability of Y in a cage with absorbing walls can be found by counting the number of Brownian paths that go from $Y(0) = 0$ to $Y(1) = 0$ without ever hitting the barriers. More precisely, the survival probability of the Brownian bridge in the same stripe can be computed as $f_1(0; k)/f_1(0; \infty)$, where $f_u(y; k)$ is the transition kernel of the Brownian motion within the allowed region $[-k, k]$. It satisfies the simple Fokker-Planck equation

$$\begin{cases} \partial_u f_u(y; k) = \frac{1}{2} \partial_y^2 f_u(y; k), & \forall u \in [0, 1]. \\ f_u(\pm k; k) = 0 \end{cases}$$

By spectral decomposition of the Laplacian, the solution is found to be

$$f_u(y; k) = \frac{1}{k} \sum_{n \in \mathbb{Z}} e^{-E_n u} \cos(\sqrt{2E_n} y), \quad \text{where } E_n = \frac{1}{2} \left(\frac{(2n-1)\pi}{2k} \right)^2$$

and the free propagator in the limit $k \rightarrow \infty$ is the usual

$$f_u(y; \infty) = \frac{1}{\sqrt{2\pi u}} e^{-\frac{y^2}{2u}},$$

so that the survival probability of the constrained Brownian bridge is

$$\mathcal{P}_<(k; 0, 1) = \frac{\sqrt{2\pi}}{k} \sum_{n \in \mathbb{Z}} \exp\left(-\frac{(2n-1)^2 \pi^2}{8k^2}\right). \quad (6)$$

Although it looks different from Eq. (5), the two expressions can be shown to be exactly identical. But the above proof looks to us way easier than the canonical ones [10].

Diffusion in a cage with moving walls

The problem can be looked at differently. Under the following change of variable and time

$$W(t) = (1+t)Y\left(\frac{t}{1+t}\right), \quad t = \frac{u}{1-u} \in \left[\frac{a}{1-a}, \frac{b}{1-b}\right], \quad (7)$$

the problem can be transformed into that of a Brownian diffusion inside a box with walls *moving at constant velocity*. Indeed, one can check that

$$\text{Cov}(W(t), W(t')) = \min(t, t'),$$

and that $\mathcal{P}_<(k|0, 1)$ can be now written as

$$\mathcal{P}_<(k|0,) = \mathbb{P}[|W(t)| \leq k(1 + t), \forall t \in [0, \infty]].$$

Since the walls expand as $\sim t$ faster than the diffusive particle can move ($\sim \sqrt{t}$), the survival probability converges to a positive value, which is again given by the usual Kolmogorov distribution (5) [12–14].

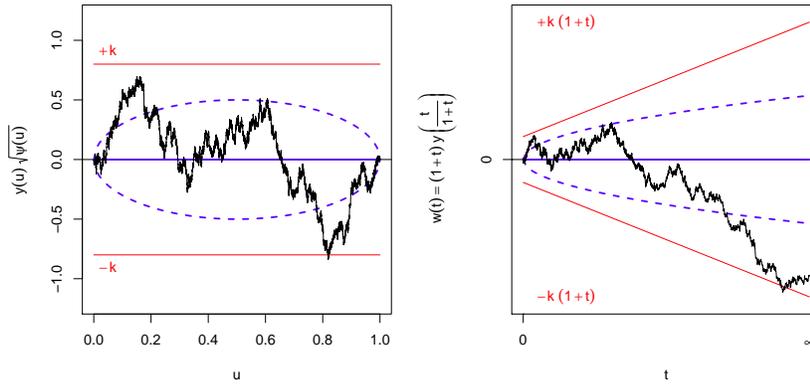


Fig. 1. The equi-weighted Brownian bridge, $\psi(u) = 1$. The time-changed rescaled process lives in a geometry with boundaries receding at constant speed.

**2.3. The variance-weighted Brownian bridge:
Accounting for the tails**

As mentioned above, the classical KS test is only weakly sensitive to the quality of the fit in the tails of the distribution, when it is often these tail events (corresponding to centennial floods, devastating earthquakes, financial crashes, etc.) that one is most concerned with (see, e.g., Ref. [15]).

A simple and elegant GoF test for the tails *only* can be designed starting with digital weights in the form $\psi(u; a) = \mathbb{1}_{\{u \geq a\}}$ or $\psi(u; b) = \mathbb{1}_{\{u \leq b\}}$ for upper and lower tail, respectively. The corresponding test laws can be read off Eq. (5.9) in Ref. [10].^a Investigation of both tails is attained with $\psi(u; q) = \mathbb{1}_{\{u \leq 1-q\}} + \mathbb{1}_{\{u \geq q\}}$ (where $q > \frac{1}{2}$).

^aThe quantity M appearing there is the volume under the normal bivariate surface between specific bounds, and it takes a very convenient form in the unilateral cases $\frac{1}{2} \leq a \leq u \leq 1$ and $0 \leq u \leq b \leq \frac{1}{2}$. Mind the missing j exponentiating the alternating (-1) factor.

Here we rather focus on a GoF test for a univariate sample of size $N \gg 1$, with the Kolmogorov distance but equi-weighted quantiles, which is equally sensitive to *all regions* of the distribution.^b We unify two earlier attempts at finding asymptotic solutions, one by Anderson and Darling in 1952 [10] and a more recent, seemingly unrelated one that deals with “life and death of a particle in an expanding cage” by Krapivsky and Redner [12, 16]. We present here the exact asymptotic solution of the corresponding stochastic problem, and deduce from it the precise formulation of the GoF test, which is of a fundamentally different nature than the KS test.

So in order to zoom on the tiny differences in the tails of the Brownian bridge, we weight it as explained earlier, with its variance

$$\psi(u) = \frac{1}{u(1-u)}.$$

Solutions for the distributions of such variance-weighted Kolmogorov-Smirnov statistics were studied by Noé, leading to the laws of the one-sided [17] and two-sided [18] finite sample tests. They were later generalized and tabulated numerically by Niederhausen [19, 20]. However, although exact and appropriate for small samples, these solutions rely on recursive relations and are not in closed form. We instead come up with an analytic closed-form solution for large samples that relies on an elegant analogy from statistical physics.

Diffusion in a cage with moving walls

After performing the above change of variable (7) that converts a Brownian bridge into a Brownian motion, $\mathcal{P}_<(k|a, b)$ can be written as

$$\mathcal{P}_<(k|a, b) = \mathbb{P}\left[|W(t)| \leq k\sqrt{t}, \forall t \in \left[\frac{a}{1-a}, \frac{b}{1-b}\right]\right].$$

The problem with initial time $\frac{a}{1-a} = 0$ and horizon time $\frac{b}{1-b} = T$ has been treated by Krapivsky and Redner in Ref. [12] as the survival probability $S(T; k = \sqrt{\frac{A}{2D}})$ of a Brownian particle diffusing with constant D in a cage with walls expanding as \sqrt{At} . Their result is that for large T ,

$$S(T; k) \equiv \mathcal{P}_<(k|0, \frac{T}{1+T}) \propto T^{-\theta(k)}.$$

They obtain analytical expressions for $\theta(k)$ in both limits $k \rightarrow 0$ and $k \rightarrow \infty$. The limit solutions of the very same differential problem were

^bOther choices of ψ generally result in much harder problems.

found earlier by Turban for the critical behavior of the directed self-avoiding walk in parabolic geometries [21].

We take here a slightly different route, suggested (but not finalized) by Anderson and Darling in Ref. [10]. Our specific contributions are: (i) we treat the general case $a > 0$ for *any* k ; (ii) we explicitly compute the k -dependence of both the exponent *and* the prefactor of the power-law decay; and (iii) we provide the link with the theory of GoF tests and compute the pre-asymptotic distribution when $]a, b[\rightarrow]0, 1[$ of the weighted Kolmogorov-Smirnov test statistics.

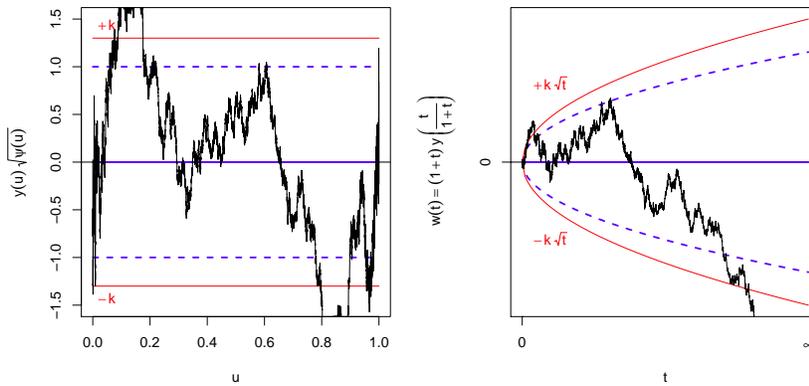


Fig. 2. The variance-weighted Brownian bridge, $\psi(u) = 1/[u(1-u)]$. The time-changed rescaled process lives in a geometry with boundaries receding as $\sim \sqrt{t}$.

Mean-reversion in a cage with fixed walls

Introducing now the new time change $\tau = \ln \sqrt{\frac{1-a}{a}} t$, the variable $Z(\tau) = W(t)/\sqrt{t}$ is a stationary Ornstein-Uhlenbeck process on $[0, T]$ where

$$T = \ln \sqrt{\frac{b(1-a)}{a(1-b)}}, \tag{8}$$

and

$$\text{Cov}(Z(\tau), Z(\tau')) = e^{-|\tau-\tau'|}.$$

Its dynamics is described by the stochastic differential equation

$$dZ(T) = -Z(T)dT + \sqrt{2} dB(T), \tag{9}$$

with $B(T)$ an independent Wiener process. The initial condition for $T = 0$ (corresponding to $b = a$) is $Z(0) = Y(a)/\sqrt{\mathbb{V}[Y(a)]}$, a random Gaussian variable of zero mean and unit variance. The distribution $\mathcal{P}_<(k|a, b)$ can now be understood as the unconditional survival probability of a mean-reverting particle in a cage with fixed absorbing walls:

$$\begin{aligned}\mathcal{P}_<(k|T) &= \mathbb{P}[-k \leq Z(\tau) \leq k, \forall \tau \in [0, T]] \\ &= \int_{-k}^k f_T(z; k) dz,\end{aligned}$$

where

$$f_T(z; k) dz = \mathbb{P}[Z(T) \in [z, z + dz] | \{Z(\tau)\}_{\tau < T}]$$

is the density probability of the particle being at z at time T , when walls are in $\pm k$. Its dependence on k , although not explicit on the right hand side, is due to the boundary condition associated with the absorbing walls (it will be dropped in the following for the sake of readability)^c.

The Fokker-Planck equation governing the evolution of the density $f_T(z)$ reads

$$\partial_\tau f_\tau(z) = \partial_z [z f_\tau(z)] + \partial_z^2 [f_\tau(z)], \quad 0 < \tau \leq T.$$

Calling \mathcal{H}_{FP} the second order differential operator $-\mathbb{1} + z\partial_z + \partial_z^2$, the full problem thus amounts to finding the general solution of

$$\begin{cases} -\partial_\tau f_\tau(z) = \mathcal{H}_{\text{FP}}(z) f_\tau(z) \\ f_\tau(\pm k) = 0, \forall \tau \in [0, T] \end{cases}.$$

We have explicitly introduced a minus sign since we expect that the density decays with time in an absorption problem. Because of the term $z\partial_z$, \mathcal{H}_{FP} is not hermitian and thus cannot be diagonalized. However, as is well known, one can define $f_\tau(z) = e^{-\frac{z^2}{4}} \phi_\tau(z)$ and the Fokker-Planck equation becomes

$$\begin{cases} -\partial_\tau \phi_\tau(z) = [-\partial_z^2 + \frac{1}{4}z^2 - \frac{1}{2}\mathbb{1}] \phi_\tau(z), \\ \phi_\tau(\pm k) = 0, \forall \tau \in [0, T] \end{cases},$$

Its Green's function, i.e. the (separable) solution *conditionally on the initial position* (z_i, T_i) , is the superposition of all modes

$$G_\phi(z, T | z_i, T_i) = \sum_\nu e^{-\theta_\nu(T-T_i)} \widehat{\varphi}_\nu(z) \widehat{\varphi}_\nu(z_i),$$

^cIn particular, $\mathcal{P}_<(k|0) = \text{erf}\left(\frac{k}{\sqrt{2}}\right)$.

where $\widehat{\varphi}_\nu$ are the normalized solutions of the stationary Schrödinger equation

$$\begin{cases} [-\partial_z^2 + \frac{1}{4}z^2] \varphi_\nu(z) = (\theta_\nu + \frac{1}{2}) \varphi_\nu(z), \\ \varphi_\nu(\pm k) = 0 \end{cases},$$

each decaying with its own energy θ_ν , where ν labels the different solutions with increasing eigenvalues, and the set of eigenfunctions $\{\widehat{\varphi}_\nu\}$ defines an orthonormal basis of the Hilbert space on which $\mathcal{H}_S(z) = [-\partial_z^2 + \frac{1}{4}z^2]$ acts. In particular,

$$\sum_\nu \widehat{\varphi}_\nu(z) \widehat{\varphi}_\nu(z') = \delta(z - z'), \quad (10)$$

so that indeed $G_\phi(z, T_i | z_i, T_i) = \delta(z - z_i)$, and the general solution writes

$$f_T(z_T; k) = \int_{-k}^k e^{\frac{z_i^2 - z_T^2}{4}} G_\phi(z_T, T | z_i, T_i) f_0(z_i) dz_i,$$

where $T_i = 0$, which corresponds to the case $b = a$ in Eq. (4), and f_0 is the distribution of the initial value z_i which is here, as noted above, Gaussian with unit variance.

\mathcal{H}_S figures out an harmonic oscillator of mass $\frac{1}{2}$ and frequency $\omega = \frac{1}{\sqrt{2}}$ within an infinitely deep well of width $2k$: its eigenfunctions are parabolic cylinder functions [22, 23]

$$\begin{aligned} y_+(\theta; z) &= e^{-\frac{z^2}{4}} {}_1F_1\left(-\frac{\theta}{2}, \frac{1}{2}, \frac{z^2}{2}\right) \\ y_-(\theta; z) &= z e^{-\frac{z^2}{4}} {}_1F_1\left(\frac{1-\theta}{2}, \frac{3}{2}, \frac{z^2}{2}\right) \end{aligned}$$

properly normalized. The only acceptable solutions for a given problem are the linear combinations of y_+ and y_- which satisfy orthonormality (10) and the boundary conditions: for periodic boundary conditions, only the integer values of θ would be allowed, whereas with our Dirichlet boundaries $|\widehat{\varphi}_\nu(k)| = -|\widehat{\varphi}_\nu(-k)| = 0$, real non-integer eigenvalues θ are allowed. For instance, the fundamental level $\nu = 0$ is expected to be the symmetric solution $\widehat{\varphi}_0(z) \propto y_+(\theta_0; z)$ with θ_0 the smallest possible value compatible with the boundary condition:

$$\theta_0(k) = \inf_{\theta > 0} \{\theta : y_+(\theta; k) = 0\}. \quad (11)$$

In what follows, it will be more convenient to make the k -dependence explicit, and a hat will denote the solution with the normalization relevant to

our problem, namely $\widehat{\varphi}_0(z; k) = y_+(\theta_0(k); z)/\|y_+\|_k$, with the norm

$$\|y_+\|_k^2 \equiv \int_{-k}^k y_+(\theta_0(k); z)^2 dz,$$

so that $\int_{-k}^k \widehat{\varphi}_\nu(z; k)^2 dz = 1$.

Asymptotic survival rate

Denoting by $\Delta_\nu(k) \equiv [\theta_\nu(k) - \theta_0(k)]$ the gap between the excited levels and the fundamental, the higher energy modes $\widehat{\varphi}_\nu$ cease to contribute to the Green's function when $\Delta_\nu T \gg 1$, and their contributions to the above sum die out exponentially as T grows. Eventually, only the lowest energy mode $\theta_0(k)$ remains, and the solution tends to

$$f_T(z; k) = A(k) e^{-\frac{z^2}{4}} \widehat{\varphi}_0(z; k) e^{-\theta_0(k)T},$$

when $T \gg (\Delta_1)^{-1}$, with

$$A(k) = \int_{-k}^k e^{\frac{z^2}{4}} \widehat{\varphi}_0(z; k) f_0(z) dz. \quad (12)$$

Let us come back to the initial problem of the weighted Brownian bridge reaching its extremal value in $[a, b]$. If we are interested in the limit case where a is arbitrarily close to 0 and b close to 1, then $T \rightarrow \infty$ and the solution is thus given by

$$\begin{aligned} \mathcal{P}_<(k|T) &= A(k) e^{-\theta_0(k)T} \int_{-k}^k e^{-\frac{z^2}{4}} \widehat{\varphi}_0(z; k) dz \\ &= \widetilde{A}(k) e^{-\theta_0(k)T}, \end{aligned}$$

with $\widetilde{A}(k) \equiv \sqrt{2\pi} A(k)^2$.

We now compute explicitly the limit behavior of both $\theta_0(k)$ and $\widetilde{A}(k)$.

$k \rightarrow \infty$ As k goes to infinity, the absorption rate $\theta_0(k)$ is expected to converge toward 0: intuitively, an infinitely far barrier will not absorb anything. At the same time, $\mathcal{P}_<(k|T)$ must tend to 1 in that limit. So $\widetilde{A}(k)$ necessarily tends to unity. Indeed,

$$\begin{aligned} \theta_0(k) &\xrightarrow{k \rightarrow \infty} \sqrt{\frac{2}{\pi}} k e^{-\frac{k^2}{2}} \rightarrow 0, \\ \widetilde{A}(k) &\xrightarrow{k \rightarrow \infty} \left(\int_{-\infty}^{\infty} \widehat{\varphi}_0(z; \infty)^2 dz \right)^2 = 1. \end{aligned} \quad (13)$$

In principle, we see from Eq. (12) that corrections to the latter arise both (and jointly) from the functional relative difference of the solution $\epsilon(z; k) = y_+(\theta_0(k); z)/y_+(0; z) - 1$, and from the finite integration limits ($\pm k$ instead of $\pm\infty$). However, it turns out that the correction of the first kind is of second order in ϵ , see [4]. The correction to $A(k)$ is thus dominated by the finite integration limits $\pm k$, so that

$$\tilde{A}(k \rightarrow \infty) \approx \operatorname{erf}\left(\frac{k}{\sqrt{2}}\right)^2. \quad (14)$$

$k \rightarrow 0$ For small k , the system behaves like a free particle in a sharp and infinitely deep well, since the quadratic potential is almost flat around 0. The fundamental mode becomes then

$$\widehat{\varphi}_0(z; k \rightarrow 0) = \frac{1}{\sqrt{k}} \cos\left(\frac{\pi z}{2k}\right),$$

and consequently

$$\theta_0(k) \xrightarrow{k \rightarrow 0} \frac{\pi^2}{4k^2} - \frac{1}{2}, \quad (15)$$

$$\tilde{A}(k) \xrightarrow{k \rightarrow 0} \frac{16}{\pi^2 \sqrt{2\pi}} k. \quad (16)$$

We show in Fig. 3 the functions $\theta_0(k)$ and $\tilde{A}(k)$ computed numerically from the exact solution, together with their asymptotic analytic expressions. In intermediate values of k (roughly between 0.5 and 3) these limit expressions fail to reproduce the exact solution.

Higher modes and validity of the asymptotic ($N \gg 1$) solution

Higher modes $\nu > 0$ with energy gaps $\Delta_\nu \lesssim 1/T$ must in principle be kept in the pre-asymptotic computation. This, however, is irrelevant in practice since the gap $\theta_1 - \theta_0$ is never small. Indeed, $\widehat{\varphi}_1(z; k)$ is proportional to the asymmetric solution $y_-(\theta_1(k); z)$ and its energy

$$\theta_1(k) = \inf_{\theta > \theta_0(k)} \{\theta : y_-(\theta; k) = 0\}$$

is found numerically to be very close to $1 + 4\theta_0(k)$. In particular, $\Delta_1 > 1$ (as we illustrate in Fig. 4) and thus $T\Delta_1 \gg 1$ will always be satisfied in cases of interest.

2.3.1. Back to GoF testing

Let us now come back to GoF testing. In order to convert the above calculations into a meaningful test, one must specify values of a and b . The natural choice would be $a = 1/N$, corresponding to the min of the sample series since $F(\min x_n) \approx F_N(\min x_n) = \frac{1}{N}$. Eq. (8) above motivates a slightly different value of $a = 1/(N+1)$ and $b = 1 - a$, such that the relevant value of T is given correspondingly by

$$T = \ln \sqrt{\frac{b(1-a)}{a(1-b)}} = \ln N.$$

This leads to our central result for the cdf of the weighted maximal Kolmogorov distance $K(\frac{1}{N+1}, \frac{N}{N+1})$ under the hypothesis that the tested and the true distributions coincide:

$$\boxed{S(N; k) = \mathcal{P}_<(k | \ln N) = \tilde{A}(k) N^{-\theta_0(k)}}, \quad (17)$$

which is valid whenever $N \gg 1$ since, as we discussed above, the energy gap Δ_1 is greater than unity.

The final cumulative distribution function (the test law) is depicted in Fig. 5 for different values of the sample size N . Contrarily to the standard KS case, this distribution *still depends on N* : as N grows toward infinity, the curve is shifted to the right, and eventually $S(\infty; k)$ is zero for any k . In particular, the threshold value k^* corresponding to a 95% confidence level (represented as a horizontal grey line) increases with N . Since for large N , $k^* \gg 1$ one can use the asymptotic expansion above, which soon becomes quite accurate, as shown in Fig. 5. This leads to:

$$\theta_0(k^*) \approx -\frac{\ln 0.95}{\ln N} \approx \sqrt{\frac{2}{\pi}} k^* e^{-\frac{k^{*2}}{2}},$$

which gives $k^* \approx 3.439, 3.529, 3.597, 3.651$ for, respectively, $N = 10^3, 10^4, 10^5, 10^6$. For exponentially large N and to logarithmic accuracy, one has: $k^* \sim \sqrt{2 \ln(\ln N)}$. This variation is very slow, but one sees that as a matter of principle, the “acceptable” maximal value of the weighted distance is much larger (for large N) than in the KS case.

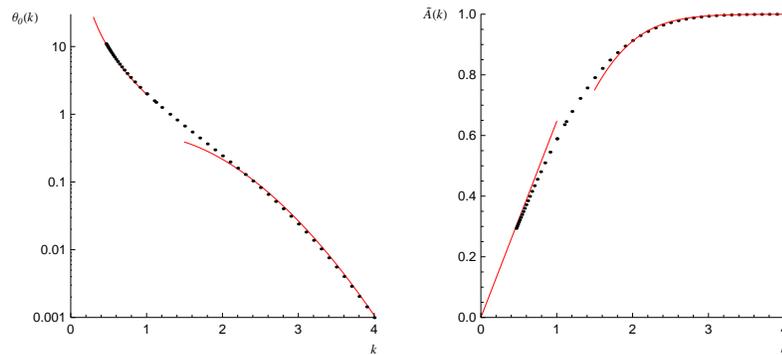


Fig. 3. **Left:** Dependence of the exponent θ_0 on k ; similar to Fig. 2 in Ref. [12], but in lin-log scale; see in particular Eqs. (9b) and (12) there. **Right:** Dependence of the prefactor \hat{A} on k . The red solid lines illustrate the analytical behavior in the limiting cases $k \rightarrow 0$ and $k \rightarrow \infty$.

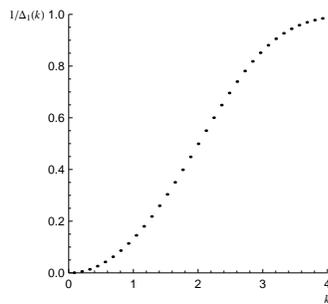


Fig. 4. $1/\Delta_1(k)$ saturates to 1, so that the condition $N \gg \exp[1/\Delta_1(k)]$ is virtually always satisfied.

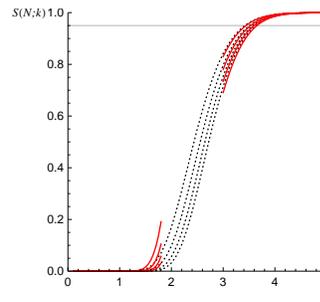


Fig. 5. Dependence of $S(N;k)$ on k for $N = 10^3, 10^4, 10^5, 10^6$ (from left to right). The red solid lines illustrate the analytical behavior in the limiting cases $k \rightarrow 0$ and $k \rightarrow \infty$.

3. Optimal time to sell a stock^d

Consider the problem of holding a stock at an initial date $t = 0$, and hoping to sell it back at a time $t = \tau$ before a deadline $t = T$. The goal is to find the *ex ante* optimal selling time τ , i.e. take a decision at $t = 0$ as of when to sell it in the future for optimal profit. If the (log-)price can be modelled as a stationary random walk, this is not a restriction of generality, since at any later time the problem is identical with however a reduced horizon.

3.1. Minimizing the expected distance to the maximum

In technical terms, we aim at minimizing the expected (relative) spread

$$S(\tau; T) = \frac{M_T - X_\tau}{M_T}$$

between the instantaneous price X_τ and the *ex post* realized maximum over the allowed horizon $M_T = \max\{X_t, t \in [0, T]\}$.

In order for this forward-looking problem to be handled analytically, we impose that the price process follows a (possibly drifted) geometric Brownian motion $X_t = e^{x_t}$ with

$$dx_t = \mu dt + \sigma dB_t \quad \text{or} \quad \dot{x} = \mu + \sigma \eta$$

where B_t in the Stochastic Differential Equation (left) is a Wiener process, and $\eta = \frac{dB_t}{dt}$ in the Langevin equation (right) is a standard White Gaussian Noise. We rewrite the maximum value as $M_T = e^{m_T}$ with obviously $m_T = \max\{x_t, t \in [0, T]\}$. The problem is clearly invariant under a shift of both x_t and m_T , so that we can arbitrarily set $x_0 = 0$.

The optimal time to sell is then defined as the solution of the minimization problem

$$\tau^* = \arg \min_{\tau \in [0, T]} \mathbb{E}[\ln S(\tau; T)] = \arg \min_{\tau \in [0, T]} \mathbb{E}[s(\tau; T)]$$

where $s(\tau; T) = m_T - x_\tau$. The expectation estimator is clearly intertemporal, and the probability distribution function \mathcal{P}_μ of $s(\tau; T)$ can be written in terms of the joint density f_μ of (x_τ, m_T) as

$$\begin{aligned} \mathcal{P}_\mu(s; \tau, T) &= \int_0^\infty \int_{-\infty}^\infty f_\mu(x, m; \tau, T) \delta(m - x - s) dx dm \\ &= \int_0^\infty f_\mu(m - s, m; \tau, T) dm. \end{aligned} \quad (18)$$

^dJoint work with S. Majumdar [24], motivated by a paper by Shiryaev, Xu and Zhou [25].

This is equivalent to writing $\mathcal{P}_\mu(s; \tau, T) = \int dF_\mu(m - s, m; \tau, T)$, with the (partial) cumulative distribution function

$$F_\mu(x, m; \tau, T) = \mathbb{P}[x_\tau = x, m_T \leq m]$$

counting the fraction of the paths arriving in $[x, x + dx]$ at time τ among all paths never crossing m from below over the whole horizon $[0, T]$. It is expressed in terms of the *causal* propagator G_μ as the probability of arriving in x at time τ without ever hitting m , and then arriving anywhere below m in the remaining time $T - \tau$:

$$F_\mu(x, m; \tau, T) = G_\mu(x, \tau; m) \int_{-\infty}^m G_\mu(x', T - \tau; m - x) dx'. \quad (19)$$

The propagator $G_\mu(x, \tau; m)$ describes a μ -drifted diffusion close to a fixed absorbing boundary, or equivalently a pure diffusion close to a boundary moving at constant velocity (“daredevil at the edge of a receding cliff”, [12]), see Fig. 6. It can be written in terms of the propagator G_0 of the zero-drift diffusion:

$$G_\mu(x, t; m) = \exp\left(-\frac{\mu^2 - 2\mu x}{2\sigma^2}\right) G_0(x, t; m),$$

where G_0 can be computed in several ways — method of images [16, 26], path-integral method [24, 27], solution of the Fokker-Planck equation. The solution writes (up to a normalizing constant) as the difference $G_0(x, t; m) \propto G_0(x, t; \infty) - G_0(x - 2m, t; \infty)$ between the free propagator

$$G_0(x, t; \infty) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{(x - x_0)^2}{2\sigma^2 t}\right)$$

with initial positions at $x_0 = 0$ and $x_0 = 2m$. As expected, $G_0(m, t; m) = 0$ at all times, saying that the probability of presence *at* the boundary is nil.

Once G_μ is known, the joint distribution f_μ of x_τ and m_T is obtained by differentiating Eq. (19) with respect to m , and the distribution of the spread is found from Eq. (18) to be

$$\boxed{\mathcal{P}_\mu(s; \tau, T) = a_\mu(s; \tau) b_{-\mu}(s, T - \tau) + a_{-\mu}(s; T - \tau) b_\mu(s, \tau)}, \quad (20)$$

where

$$a_\mu(s; \tau) = \frac{\mu}{2\sigma^2} \exp\left(-\frac{2s\mu}{\sigma^2}\right) \operatorname{erfc}\left(\frac{s - \mu\tau}{\sqrt{2\sigma^2\tau}}\right) + \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left(-\frac{(s + \mu\tau)^2}{2\sigma^2\tau}\right)$$

$$b_\mu(s; \tau) = - \exp\left(-\frac{2s\mu}{\sigma^2}\right) \operatorname{erfc}\left(\frac{s - \mu\tau}{\sqrt{2\sigma^2\tau}}\right) + \operatorname{erfc}\left(-\frac{s + \mu\tau}{\sqrt{2\sigma^2\tau}}\right).$$

Finally, the expected spread is

$$\mathbb{E}[s(\tau, T)] = \int_0^\infty s \mathcal{P}_\mu(s; \tau, T) ds = \iint_0^\infty s f_\mu(m - s, m; \tau, T) dm ds, \quad (21)$$

and the optimal τ^* is found by minimizing this function with respect to τ :

$$\tau^* = \begin{cases} T & , \mu \geq 0 \\ 0 & , \mu \leq 0 \end{cases}. \quad (22)$$

It is degenerate at $\mu = 0$ where both $\tau^* = 0$ and $\tau^* = T$ are optimal. This result states in technical terms a very intuitive truism: whenever the log-prices are expected to increase in average ($\mu > 0$) one should keep the stock as long as possible, and conversely if the log-prices are expected to fall ($\mu < 0$) one should sell immediately. It should not be surprising that τ^* is not affected by the value of the so-called “volatility” parameter σ , since the optimization program only focused on “maximizing the gain” without controlling for the encountered risk, and thus the solution applies to a risk-neutral agent only.

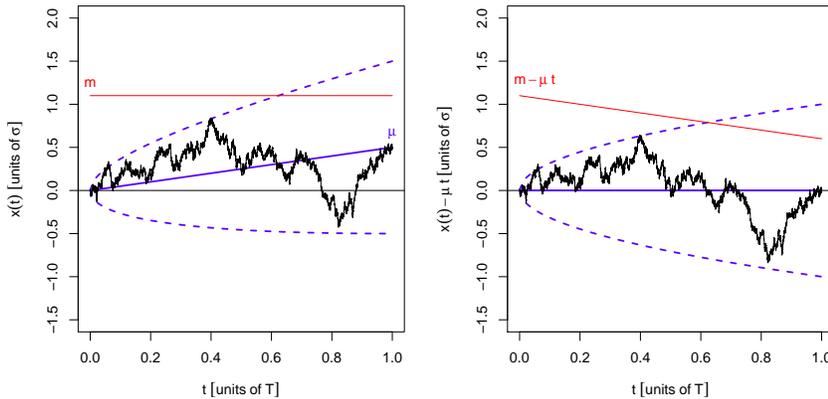


Fig. 6. Illustration for a positive drift ($\mu > 0$). In the case where $\mu < 0$, the upper barrier moves *away* linearly. See Sect. 2.2 page 5 for an application of the bilateral case.

3.2. Maximizing the occurrence time probability of maximum

Alternatively to minimizing the spread $s(\tau; T)$ between the expected maximum m_T and the log-price x_τ at the selling time, one can try to maximize

the probability $p_\mu(\tau; T)d\tau = d\mathbb{P}[x_\tau = m_T]$ that m_T will occur at time τ , whatever value it takes. But the joint probability that the global maximum over $[0, T]$ has a value $m_T = m$, and that this maximum is reached for the first (and in fact only) time at $t = \tau$ is nothing else than $f_\mu(m, m; \tau, T)$.

In order to avoid the issue of infinite crossings of the continuous Brownian motion [24, 28, 29], we allow for an infinitesimal spread $s = m - x$ that we eventually take to 0:

$$p_\mu(\tau; T) \propto \int_0^\infty \lim_{s \rightarrow 0} f_\mu(m - s, m; \tau, T) dm = \mathcal{P}_\mu(0^+; \tau, T),$$

with \mathcal{P}_μ given in Eq. (20). Noticing that a_μ converges to a finite value $a_\mu(0, \tau)$ when $s \rightarrow 0$, and that the expansion of b_μ to first order in s around 0 is $b_\mu(s \rightarrow 0, \tau) \approx 4s a_\mu(0, \tau)$, we have

$$\mathcal{P}_\mu(s \rightarrow 0; \tau, T) = 8s a_\mu(0, \tau) a_{-\mu}(0, T - \tau)$$

and finally, normalizing with $\int_0^T \mathcal{P}_\mu(s; \tau, T)d\tau$, we get

$$\boxed{p_\mu(\tau; T) = 2\sigma^2 a_\mu(0, \tau) a_{-\mu}(0, T - \tau)}. \quad (23)$$

Notice that when $\mu = 0$ the function $a_0(s, \tau)$ is the centered normal distribution with variance $\sigma^2\tau$. In particular $a_0(0, \tau) = 1/\sqrt{2\pi\sigma^2\tau}$ and one recovers Lévy's result [30]:

$$p_0(\tau; T) = \frac{1}{\pi} \frac{1}{\sqrt{\tau(T - \tau)}},$$

with two global maxima at $\tau^* = 0$ and $\tau^* = T$. For non zero μ 's, the distribution still have inverse square-root singularities both at $\tau = 0$ and $\tau = T$, but with unequal amplitudes. For $\mu < 0$, the amplitude of the $\tau = 0$ singularity is larger than that of the $\tau = T$ singularity, and vice-versa when $\mu > 0$. Therefore one concludes that:

$$\tau_m = \arg \max_{\tau \in [0, T]} p_\mu(\tau; T) = \begin{cases} T & , \mu \geq 0 \\ 0 & , \mu \leq 0 \end{cases} \quad (24)$$

and is thus equal to τ^* , see Eq. (22).

Whereas the minimization program of the previous section embedded the information of all the possible gaps s (as revealed by Eq. (21)), maximizing the occurrence time distribution only cares for the infinitesimal proximity of the maximum $s = m - x \rightarrow 0$. Nevertheless, although the objective function is not the same (minimize the spread or maximize the probability), the solution of the optimal time is not sensitive to the chosen criterion.

4. Optimal trading with linear costs^e

The problem addressed in this section is to determine the optimal strategy in the presence of “linear” trading costs (i.e. a fixed cost *per share*, neglecting any price impact) and a constraint on the maximum size of the position (both long and short). This problem is of very significant interest in practice, at least for small sizes. For large sizes, a quadratic cost can be added to mimic price impact; the problem is however not (yet ?) solved in full generality.

We consider an agent who wants to maximize his/her expected gains, by trading a single asset, of current price Price_t , over a long period $[0, T]$ (we will later consider the limit $T \rightarrow \infty$). The position (signed number of shares/contracts) of the trader at time t is π_t . We assume that the agent has some signal p_t that predicts the next price change $r_t = \text{Price}_{t+1} - \text{Price}_t$, and is faced with the following constraints:

- His/her risk control system is simply a cap on the absolute size of his/her position : $|\pi_t| \leq M$, with no other risk control.
- He/she has to pay linear costs $\Gamma|\Delta\pi_t|$ whenever he/she trades a quantity $\Delta\pi_t \equiv \pi_{t+1} - \pi_t$

We assume that the predictor has a number of “nice” (but natural) properties; in particular, the predictability $\mathfrak{L}_t(p) = \mathbb{E}[r_t | p_t = p]$ is an odd, continuous and strictly increasing function of p . We also assume that it is Markovian: $\forall \omega_{t+1}, \mathbb{P}[\omega_{t+1} | p_t, p_{t-1}, \dots] = \mathbb{P}[\omega_{t+1} | p_t]$ where ω_{t+1} is any event at $t + 1$. In what follows, we will use the notation $P_t(p' | p) dp' = \mathbb{P}[p_{t+1} = p' | p_t = p] dp'$. We also define an integrated predictability at $t = \infty$, depending on p_t :

$$p_\infty(p_t) = \mathbb{E}[\text{Price}_\infty - \text{Price}_t | p_t] = \sum_{n=0}^{\infty} \mathbb{E}[r_{t+n} | p_t].$$

This quantity indicates how much one will gain in the future if one keeps a fixed position $\pi_{t' \geq t} = \pi$: the expected gain is then $p_\infty(p_t) \pi$.

4.1. The Optimal Strategy

A naïve solution

At first sight, the solution to this problem seems straightforward: if the expected future gain (given by the integrated predictability) exceeds the

^eJoint work with J. de Lataillade, C. Deremble and M. Potters [31].

trading cost per contract Γ , then one trades in the direction of the signal (if not already at the maximum position), otherwise one does not. This solution obviously generates a positive average gain, but it has no reason to be the optimal solution. Indeed, because the predictor is auto-correlated in time, it might be worthy (and in general it will be) to wait for a larger value of the predictor, in order to grab the opportunities that have the most chances to get realized, and discard the others. As we shall see, the mistake in this naïve reasoning is not to compare the future gain with the cost, but rather comes from a wrong definition of the future gain, which does not include future trading decisions.

The Bellman method: general solution

The framework to attack this problem is Bellman's optimal control theory, or dynamic programming [32], which consists in solving the problem backwards: by assuming one follows the optimal strategy for all future times $t' > t$, one can find the optimal solution at time t . As is usual in dynamic programming, one has a *control variable* π_t , which needs to be optimized, and a *state variable* p_t , which parameterizes the solution. The optimization is done through a *value function* $V_t(\pi, p)$, which gives the maximal expected gains between time t and $+\infty$, considering that the position at $t - 1$ is π and the predictor's value at t is p . The optimal solution of the system will be denoted $(\pi_t^*)_{t \in [0, T]}$.

At the last time step $t = T$, the expected future return is really $p_\infty(p)\pi_T$ where $p = p_T$, since no trading is allowed beyond that time. Any trade $\Delta\pi_T$ induces a cost $\Gamma|\Delta\pi_T|$, so:

If $p_\infty(p) \geq +\Gamma$ then $\pi_T^* = +M$, and $V_T(\pi, p) = +p_\infty(p)M - \Gamma(M - \pi)$
 If $p_\infty(p) \leq -\Gamma$ then $\pi_T^* = -M$, and $V_T(\pi, p) = -p_\infty(p)M - \Gamma(M + \pi)$
 If $|p_\infty(p)| < \Gamma$ then $\pi_T^* = \pi$, and $V_T(\pi, p) = p_\infty(p)\pi$.

Hence, one recovers exactly the naïve solution in this case, but this is only because there is no trading beyond $t = T$. Now at $t < T$, the quantity to be maximized includes immediate gains, costs and future gains. This leads to the following recursion relation:

$$V_t(\pi, p) = \max_{|\pi'| \leq M} \left\{ \mathfrak{L}_t(p) \cdot \pi' - \Gamma|\pi' - \pi| + \int V_{t+1}(\pi', p') P_t(p'|p) dp' \right\}, \quad (25)$$

and π_t^* is the value of π' which realises this maximum when $\pi = \pi_{t-1}^*$ and $p = p_t$. The general solution is given by the following construction [31]:

- $\pi_t^* = \begin{cases} \pi_{t-1}^* & \text{if } |p_t| < q_t \\ \text{sign}(p_t) \cdot M & \text{if } |p_t| \geq q_t \end{cases}$ (with $\pi_{-1} = 0$)
- q_t is such that $q_t \geq 0$ and $g(t, q_t) = \Gamma$, where $g(t, p)$ is a continuous, strictly increasing function of p which satisfies, for $t < T$:

$$g(t, p) = \mathfrak{L}_t(p) + \Gamma \left[\int_{q_{t+1}}^{\infty} - \int_{-\infty}^{-q_{t+1}} \right] P_t(p'|p) dp' \quad (26)$$

$$+ \int_{-q_{t+1}}^{q_{t+1}} g(t+1, p') P_t(p'|p) dp'$$

The stationary solution

The solution provided by Bellman's method above exhibits in general a dependence in t . Let us now consider the case where $T \rightarrow \infty$, and suppose that the predictor is stationary, i.e. $P_t(p'|p) = P(p'|p)$ is independent of t . Then we obtain a *telescopic* (self-consistent) equation for the one-variable function g , and the solution for the threshold q^* :

$$g(p) = \mathfrak{L}(p) + \Gamma \left[\int_{q^*}^{\infty} - \int_{-\infty}^{-q^*} \right] P(p'|p) dp' + \int_{-q^*}^{q^*} g(p') P(p'|p) dp' \quad (27)$$

$$g(q^*) = \Gamma \quad (28)$$

The optimal solution π_t^* to the system is then similar to the general solution, but with a constant threshold q^* . Thus, we obtain a very simple trading system, always saturated at $\pm M$, with a threshold to decide at each step whether we should revert the position or not. This of course looks a lot like the naïve solution of page 20. The only difference lies in the value of the threshold $q^* = g^{-1}(\Gamma)$, defined by Eqs. (27,28), instead of $q_{\text{naïve}} = p_{\infty}^{-1}(\Gamma)$ for the naïve solution. Intuitively, these equations take our future trading into account, whereas the naïve solution does not.

If we look closely at Eq. (27), its interpretation becomes transparent: $2M g(p)$ is equal the expected difference in total future profit between the situation where $\pi = +M$ and the situation where $\pi = -M$. This difference is made up of:

- $\Delta\pi \mathfrak{L}(p)$ which represents the difference in immediate gain
- $\Delta\pi \Gamma \mathbb{P}[p_{t+1} > q^* | p_t = p]$ which represents the loss if the current position is $-M$ and in the next time step the predictor goes over the positive threshold q^* (hence π will go to $+M$)

- $\Delta\pi \Gamma \mathbb{P}[p_{t+1} < -q^* | p_t = p]$ which represents the loss if the current position is $+M$ and in the next time step the predictor goes below the negative threshold $-q^*$ (hence π will go to $-M$)
- $\int_{-q^*}^{q^*} P(p'|p) 2Mg(p') dp'$ which is the expected difference in total future profit if, in the next step, the predictor remains between the two thresholds (leaving π unchanged).

Since the change of position between $-M$ and $+M$ costs $2M\Gamma$, it makes sense to compare $2Mg(p)$ with it and only trade when $g(p)$ is greater than Γ . Hence, $g(p)$ can be seen as the “gain per traded lot”.

According to Eq. (27), $|g(p)| \geq |\mathfrak{L}(p)|$ and with Eq. (28) this implies in particular that $\mathfrak{L}(q^*) \leq \Gamma$. This property is actually rather intuitive: indeed, if the immediate expected gain was higher than the trading cost, then there would be no reason not to trade the maximal possible amount.

From here on, we only consider a linear predictor, $\mathfrak{L}_t(p) = p$.

Reformulation as a path integral

Although Eq. (27) is easy to interpret, it proves very difficult to solve in concrete cases. It can be rewritten by expanding the function g :

$$\begin{aligned}
 g(p) &= p + \int_{-q^*}^{q^*} p_1 P(p_1|p) dp_1 + \int_{-q^*}^{q^*} \int_{-q^*}^{q^*} p_2 P(p_2|p_1) P(p_1|p) dp_1 dp_2 + \dots \\
 &+ \Gamma \cdot \left[\int_{q^*}^{+\infty} P(p_1|p) dp_1 + \int_{q^*}^{+\infty} \int_{-q^*}^{q^*} P(p_2|p_1) P(p_1|p) dp_1 dp_2 + \dots \right] \\
 &- \Gamma \cdot \left[\int_{-\infty}^{-q^*} P(p_1|p) dp_1 + \int_{-\infty}^{-q^*} \int_{-q^*}^{q^*} P(p_2|p_1) P(p_1|p) dp_1 dp_2 + \dots \right] \\
 &= \mathcal{G}(p) + \Gamma [\mathcal{P}_+(p) - \mathcal{P}_-(p)], \tag{29}
 \end{aligned}$$

and can thus be understood as a path integral: $\mathcal{G}(p)$ can be interpreted as the average over all possible exit times n of the cumulated predictor $\sum_{i=0}^{n-1} p_i$, where the expectation is taken only over all paths that stay in $[-q^*, q^*]$ until n . Similarly $\mathcal{P}_+(p)$ and $\mathcal{P}_-(p)$ are the probabilities for a path starting at $p_0 = p$ to exit (at any possible later time) above q^* or below $-q^*$, respectively.

Using now the fact that $g(q^*) = \Gamma$, and $\mathcal{P}_+(q^*) + \mathcal{P}_-(q^*) = 1$, we get:

$$\mathcal{G}(q^*) - 2\Gamma \mathcal{P}_-(q^*) = 0. \tag{30}$$

In some cases, both sides of this equation will tend to be infinitesimal, so it is rather the ratio $\lim_{p \rightarrow q^*} \mathcal{G}(p)/\mathcal{P}_-(p)$ that we will ask to equal 2Γ . Figure 7

illustrates our reformulation of the problem in terms of first passage times properties.

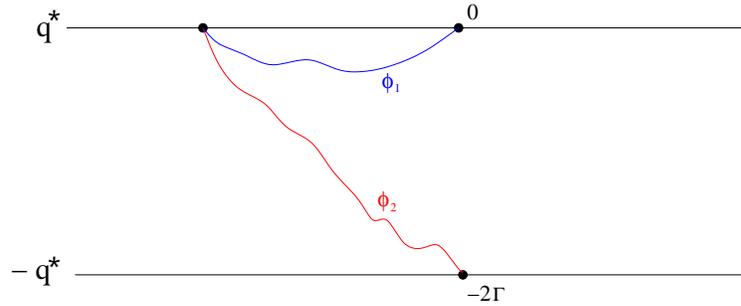


Fig. 7. Path integral representation of Eq. (30). The value of q^* is such that the “penalty” 2Γ over all paths exiting through $-q^*$ is equal to the average gain over all paths exiting either through q^* (e.g. ϕ_1 , blue) or through $-q^*$ (e.g. ϕ_2 , red).

Note that Eq. (30) is completely general provided the assumptions of page 20 are satisfied, it does not rely on any specific statistics of the predictor. In the next section, we will explicitly solve this equation when the predictor is Gaussian and follows an auto-regressive evolution.

4.2. Application to an auto-regressive linear predictor

Let us assume that the predictor follows a discrete auto-regressive dynamics:

$$p_{t+1} = \rho \cdot p_t + \beta \cdot \xi_t, \quad (31)$$

where $(\xi_t)_{t \in \mathbb{R}}$ is a set of independent $\mathcal{N}(0, 1)$ Gaussian random variables. One classical example of such a predictor is an exponential moving average of price returns:

$$p_t^{\text{EMA}} = K \sum_{t' < t} \rho^{t'-t-1} r_{t'}.$$

If we suppose, as is usual, that the returns r_t are $\mathcal{N}(0, \sigma_r)$ random variables, then p_t^{EMA} follows the dynamics in Eq. (31) with $\beta = K\sigma_r$. Note however that the r_t must have some small correlations in order to be predictable! Therefore, in this case, the discussion in terms of an auto-regressive process is only consistent in the limit $K \ll 1$.

When $\rho = 0$, the predictor is a white noise in time: $\mathbb{E}[p_t p_{t+1}] = 0$. Since we assume a perfect $p \leftrightarrow -p$ symmetry, the self-consistent equation becomes simply $g(p) = p$, which trivially implies that $q^* = \Gamma$ in that case. This threshold is what we expect from such a system: without any auto-correlation, the best strategy is to trade as soon as the instantaneous predictability is above the trading cost. Note that in this case $p_\infty = p_t$, so this threshold also coincides with the naïve solution.

At the other extreme, when $\epsilon \equiv 1 - \rho \ll 1$, we have $\mathbb{E}[p_{t+n}|p_t] \approx e^{-\epsilon n} p_t$, so $\tau = 1/\epsilon$ is the auto-correlation time of the predictor p_t . The standard deviation of the predictor, i.e. its average predictability, is $\sigma_p = \sqrt{\mathbb{E}[p_t^2]} = \beta/\sqrt{2\epsilon}$. The integrated predictability is given by

$$p_\infty(p) = \sum_{n=0}^{\infty} \mathbb{E}[p_{t+n}|p_t] \approx \sum_{n=0}^{\infty} e^{-\epsilon n} p_t \approx p/\epsilon,$$

what implies that the naïve threshold value is given by $q_{\text{naïve}} = \Gamma\epsilon$, while the integrated average predictability is $\sigma_\infty = \beta/\sqrt{2\epsilon^3}$. In what follows, we will study the problem by distinguishing between two cases:

$\beta \gg \Gamma$: the predictor can easily beat its transaction costs at every step.

This situation (which is not very realistic) requires to keep a *discrete time* approach of the problem.

$\beta \ll \Gamma$: the predictor needs in general a large number of steps to beat the costs. This will lead to a *continuous* formulation of the problem.

Discrete case: $\beta \gg \Gamma$

We already explained in page 23 that $q^* \leq \Gamma$ with a linear predictor. Consequently, whenever $\beta \gg \Gamma$, we also have $\beta \gg q^*$. This means that, starting at $p = q^*$, one will typically jump beyond q^* or $-q^*$ in just one step. Thus:

$$\mathcal{G}(q^*) = q^* \quad \text{and} \quad \mathcal{P}_-(q^*) = \int_{x^*}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx,$$

where $x^* = (2 - \epsilon)q^*/\beta$. Since $\beta \gg q^*$, one has $x^* \ll 1$ and thus $\mathcal{P}_-(q^*) \approx 1/2$. Equation (30) finally gives: $q^* = \Gamma$. Hence, if the volatility of each predictor change is very large compared to the trading costs, then one needs to be as selective as possible.

Continuous case: $\beta \ll \Gamma$

If the threshold was of the order of the predictor's surprise $q^* \approx \beta$, the predictor would have a significant probability of switching from above q^*

to below $-q^*$ in just one step. The optimal strategy would then require to resell everything at cost 2Γ , whereas the immediate gain would only be of the order of magnitude of β . So when $\beta \ll \Gamma$, we necessarily have $q^* \gg \beta$, and many steps are required for the predictor to get from $q^* \gg \beta$ to $-q^* \ll -\beta$. This is effectively the continuum limit, where the variation of the predictor at each time step is infinitesimal compared to q^* . We can then approximate the dynamics of the predictor by the Ornstein-Uhlenbeck drift-diffusion process:

$$dp_t = -\epsilon p_t dt + \beta dX_t, \quad (32)$$

where $(X_t)_t$ is a Wiener process.

In such a continuous setting, the quantities $\mathcal{G}(q^*)$ and $\mathcal{P}_-(q^*)$ are actually ill-defined because the diffusion process starts on an absorbing boundary. This is a classical problem, which is handled by starting infinitesimally close to q^* . Therefore we consider $\mathcal{G}(p)$ and $\mathcal{P}_-(p)$ for $p = q^* - \delta < q^*$. It can be shown that these two functions obey two Kolmogorov backward equations, that read:

$$\frac{1}{2}\beta^2 \frac{\partial^2 \mathcal{G}}{\partial p^2} - \epsilon p \frac{\partial \mathcal{G}}{\partial p} = -p \quad \text{and} \quad \frac{1}{2}\beta^2 \frac{\partial^2 \mathcal{P}_-}{\partial p^2} - \epsilon p \frac{\partial \mathcal{P}_-}{\partial p} = 0, \quad (33)$$

with boundary conditions: $\mathcal{G}(\pm q^*) = 0$ and $\mathcal{P}_-(q^*) = 0$, $\mathcal{P}_-(-q^*) = 1$. We therefore encounter again the problem of a Brownian harmonic oscillator confined between two walls, already discussed in Sect. 2.3. The solution of these equations are

$$\mathcal{G}(p) = \frac{1}{\epsilon} \left(p - q^* \frac{I(p\sqrt{a})}{I(q^*\sqrt{a})} \right) \quad \text{and} \quad \mathcal{P}_-(p) = \frac{1}{2} \left(1 - \frac{I(p\sqrt{a})}{I(q^*\sqrt{a})} \right),$$

with

$$I(x) = \int_0^x e^{v^2} dv \quad \text{and} \quad a = \frac{\epsilon}{\beta^2}.$$

To first order in $\delta \rightarrow 0$, Eq. (30) becomes

$$-\frac{\delta}{\epsilon} + \frac{\delta q^*}{\epsilon} \sqrt{a} \cdot \frac{I'(q^*\sqrt{a})}{I(q^*\sqrt{a})} \approx \Gamma \delta \sqrt{a} \frac{I'(q^*\sqrt{a})}{I(q^*\sqrt{a})}.$$

As expected, δ disappears from the equation, to give the following solution for the threshold q^* :

$$\boxed{q^* = \frac{\beta}{\sqrt{\epsilon}} F^{-1} \left(\frac{\Gamma \epsilon^{3/2}}{\beta} \right) \quad \text{where} \quad F(x) = x - I(x)/I'(x)}. \quad (34)$$

Note that when $\epsilon \ll 1$, this equation can be expressed entirely in terms of the integrated predictability:

$$p_\infty(q^*) = \sigma_\infty \sqrt{2} \cdot F^{-1} \left(\frac{\Gamma}{\sigma_\infty \sqrt{2}} \right).$$

This means that we can find the optimal threshold for a predictor by studying only its total predictive power (if we suppose of course that it satisfies all the required properties).

One can now study the limits of Eq. (34) for large and small values of the only remaining adimensional parameter $\eta = \Gamma \epsilon^{3/2} / \beta$. Interestingly, $\eta \approx 1$ is the regime of practical interest where predictability beats costs whenever the predictor's value is of the order of its rms. The asymptotes of $F(x)$ are as follows:

- If $x \gg 1$, then $\int_0^x e^{v^2} dv \ll e^{x^2}$, so $F(x) \approx x$.
- If $x \ll 1$, then $F(x) \simeq x - (1 - x^2) \int_0^x (1 - v^2) dv \approx 2x^3/3$.

Therefore when $\eta \gg 1$, the threshold is simply given by $q^* = \Gamma \epsilon$. This result is rather intuitive: if β is very small then the predictability of the predictor is weak, compared to the trading cost. Hence, it makes sense to try to catch any profitable opportunity, without taking future trading into account. That is why we recover the naïve solution of page 20. If on the other hand $\beta \gg \Gamma \epsilon^{3/2}$ then $\eta \ll 1$, and $F^{-1}(\eta) \approx \sqrt[3]{3\eta/2}$, which yields

$$q^* = \sqrt[3]{\frac{3}{2} \cdot \Gamma \beta^2}.$$

This says that if β is large enough, the optimal threshold is independent of the mean-reversion parameter ϵ . The surprise, however, is the rather unexpected dependence of the threshold q^* as the 1/3 power of the trading costs. This result was obtained in the literature before, in the limit considered here of a continuous time random walk, see Refs. [33–35]. Our formulation is however much more general, and would allow one to treat non Gaussian and non stationary situations as well.

5. Some open problems

We presented three very different examples of “first passage time” problems coming from quantitative finance. Let us discuss some extensions and open issues concerning these three problems.

As far as the Kolmogorov-Smirnov goodness-of-fit test is concerned, we believe that extensions of this test to higher-dimensional, multivariate settings, would be quite interesting. More precisely, the concept of “copulas” (that describe the correlation structure between dependent variables) has become an important one in theoretical finance in the recent years. For pairs of dependent variables, the copula $C(u, v)$ is an increasing function of both its arguments, from $[0, 1] \times [0, 1]$ to $[0, 1]$. The trivial copula, corresponding to independent variables, is such that $C(u, v) = uv$. It turns out that it is always possible to transform an arbitrary copula into the independent one by an appropriate change of variables, $(u, v) \rightarrow (s, t)$ [36, 37]. One can then, in the spirit of KS, test the GoF in a *copula independent* manner. The problem boils down to estimating the distribution of the maximum of a pinned “Brownian sheet” that generalizes the Brownian bridge described above. This is still an unsolved problem, but there is a hope that an exact solution can be found. Extensions to weights that emphasize the “tails” of the copula, similar to our one-dimensional problem above, would be quite interesting too.

The second problem, concerning the optimal selling time, is interesting from a mathematical/pedagogical point of view, but the final result turns out to be quite trivial from a financial point of view. A more interesting problem would be to add some correlations in the returns, accounting for trends or mean-reversion, for example with an exponentially decaying correlation function of the lag that would allow to make the problem Markovian.

Finally, the issue of optimal strategies in the presence of transaction costs would deserve much more attention. One particularly relevant endeavor would be to solve the problem in the presence of both linear *and* quadratic costs, i.e. when the cost of a change of position $\Delta\pi$ is of the form $\Gamma|\Delta\pi| + \Gamma'|\Delta\pi|^2$. The case treated in this review corresponds to $\Gamma' = 0$, but in practice *price impact* is very important: prices tend to go up when one buys, and down when one sells.

References

- [1] E. Derman and P. Wilmott. The financial modelers' manifesto. Online blog.
- [2] R. C. Merton, On the pricing of corporate debt: The risk structure of interest rates, *The Journal of Finance*. **29**(2), 449–470 (1974).
- [3] J. C. Hull, *Options, Futures and Other Derivatives*. Prentice Hall finance series, Prentice Hall Higher Education (2009).
- [4] R. Chicheportiche and J.-P. Bouchaud, Weighted Kolmogorov-Smirnov test: Accounting for the tails, *Physical Review E*. **86**, 041115 (Oct, 2012).
- [5] A. N. Kolmogorov, Sulla determinazione empirica di una legge di distribuzione, *Giornale dell'Istituto Italiano degli Attuari*. **4**(1), 83–91 (1933).
- [6] N. Smirnov, Table for estimating the goodness of fit of empirical distributions, *The Annals of Mathematical Statistics*. **19**(2), 279–281 (1948).
- [7] J. L. Doob, Heuristic approach to the Kolmogorov-Smirnov theorems, *The Annals of Mathematical Statistics*. **20**(3), 393–403 (1949).
- [8] E. V. Khmaladze, Martingale approach in the theory of goodness-of-fit tests, *Theory of Probability and its Applications*. **26**, 240 (1982).
- [9] R. Chicheportiche and J.-P. Bouchaud, Goodness-of-fit tests with dependent observations, *Journal of Statistical Mechanics: Theory and Experiment*. **2011**(09), P09003 (2011).
- [10] T. W. Anderson and D. A. Darling, Asymptotic Theory of Certain “Goodness of Fit” Criteria Based on Stochastic Processes, *The Annals of Mathematical Statistics*. **23**(2), 193–212 (1952).
- [11] D. A. Darling, The Kolmogorov-Smirnov, Cramer-von Mises Tests, *The Annals of Mathematical Statistics*. **28**(4), 823–838 (1957).
- [12] P. L. Krapivsky and S. Redner, Life and death in an expanding cage and at the edge of a receding cliff, *American Journal of Physics*. **64**(5), 546–551 (1996).
- [13] A. J. Bray and R. Smith, The survival probability of a diffusing particle constrained by two moving, absorbing boundaries, *Journal of Physics A: Mathematical and Theoretical*. **40**(10), F235 (2007).
- [14] A. J. Bray and R. Smith, Survival of a diffusing particle in an expanding cage, *Journal of Physics A: Mathematical and Theoretical*. **40**(36), 10965 (2007).
- [15] A. Clauset, C. R. Shalizi, and M. E. J. Newman, Power-law distributions in empirical data, *SIAM review*. **51**(4), 661–703 (2009).
- [16] S. Redner, *A guide to first-passage processes*. Cambridge University Press, Cambridge, UK (2001).
- [17] M. Noé and G. Vandewiele, The calculation of distributions of Kolmogorov-Smirnov type statistics including a table of significance points for a particular case, *The Annals of Mathematical Statistics*. **39**(1), 233–241 (1968).
- [18] M. Noé, The calculation of distributions of two-sided Kolmogorov-Smirnov type statistics, *The Annals of Mathematical Statistics*. **43**(1), 58–64 (1972).
- [19] H. Niederhausen. Tables of significance points for the variance-weighted Kolmogorov-Smirnov statistics. Technical report, Stanford University, Department of Statistics (February, 1981).

- [20] R. R. Wilcoxon, Percentage points of a weighted Kolmogorov-Smirnov statistic, *Communications in Statistics - Simulation and Computation*. **18**(1), 237–244 (1989).
- [21] L. Turban, Anisotropic critical phenomena in parabolic geometries: the directed self-avoiding walk, *Journal of Physics A: Mathematical and General*. **25**, L127 (1992).
- [22] W. N. Mei and Y. C. Lee, Harmonic oscillator with potential barriers — exact solutions and perturbative treatments, *Journal of Physics A: Mathematical and General*. **16**, 1623 (1983).
- [23] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products. Corrected and Enlarged Edition*. Academic Press, New York, London, Toronto and Tokyo (1980).
- [24] S. N. Majumdar and J.-P. Bouchaud, Optimal time to sell a stock in the Black-Scholes model: comment on Ref. [25], *Quantitative Finance*. **8**(8), 753–760 (2008).
- [25] A. Shiryaev, Z. Xu, and X. Y. Zhou, Thou shalt buy and hold, *Quantitative Finance*. **8**(8), 765–776 (2008).
- [26] W. Feller, *An introduction to probability theory and its applications*. vol. 1 and 2, *Wiley Series in Probability and Mathematical Statistics*, John Wiley & Sons, New York (1968).
- [27] S. N. Majumdar, Brownian functionals in physics and computer science, *Current Science*. **89**(12), 2076 (2005).
- [28] J. Randon-Furling and S. N. Majumdar, Distribution of the time at which the deviation of a Brownian motion is maximum before its first-passage time, *Journal of Statistical Mechanics: Theory and Experiment*. **2007**(10), P10008 (2007).
- [29] S. N. Majumdar, J. Randon-Furling, M. J. Kearney, and M. Yor, On the time to reach maximum for a variety of constrained Brownian motions, *Journal of Physics A: Mathematical and Theoretical*. **41**(36), 365005 (2008).
- [30] P. Lévy, Sur certains processus stochastiques homogènes, *Compositio Mathematica*. **7**, 283–339 (1939).
- [31] J. De Lataillade, C. Deremble, M. Potters, and J.-P. Bouchaud, Optimal trading with linear costs, *arXiv preprint q-fin.PM/1203.5957* (2012).
- [32] R. E. Bellman, *Dynamic Programming*. Dover Books on Mathematics, Dover Publications (2003).
- [33] L. C. G. Rogers. Why is the effect of proportional transaction costs $O(\delta^{2/3})$? In *Contemporary Mathematics*, vol. 351, pp. 303–308, American Mathematical Society, Providence, RI (2004).
- [34] R. J. Martin and T. Schöneborn, Mean reversion pays, but costs, *Risk Magazine*. pp. 96–101 (Feb, 2011).
- [35] R. J. Martin, Optimal multifactor trading under proportional transaction costs, *arXiv preprint q-fin.TR/1204.6488* (2012).
- [36] R. Chicheportiche. *Non linear Dependences in Finance*. PhD thesis, École Centrale Paris (Jun, 2013).
- [37] R. Chicheportiche and J.-P. Bouchaud. Ironing the copula surface. In preparation (2013).