Optimal stopping problems

Part II. Applications

A. N. Shiryaev   **M. V. Zhitlukhin**

Steklov Mathematical Institute, Moscow
The University of Manchester, UK

Bonn – May 2013

Outline

The second part of the course contains some particular optimal stopping problems related to **mathematical finance**.

The questions we consider concern problems of portfolio re-balancing and choosing optimal moments of time to sell (or buy) stock.

Problems of this type play the central role in

the **technical analysis** of financial markets,

the field, which is much less theoretically developed in comparison with the two other analyses: the **fundamental** and the **quantitative** ones.

Most of the methods in the technical analysis are based on empirical evidence of "rules of thumb", while we will try to present a mathematical foundation for them.

The contents of this part of the course

1. Overview of general facts from the optimal stopping theory
2. Stopping a Brownian motion at its maximum
   Peskir, Shiryaev. *Optimal stopping and free-boundary problems*, 2006; sec. 30
3. Trading rule "Buy and hold"
   Shiryaev, Xu, Zhou. *Thou shalt buy and hold*, 2008
4. Sequential hypothesis testing
   Shiryaev. *Optimal stopping rules*, 2007; ch. 4
5. Sequential parameters estimation
   Çetin, Novikov, Shiryaev. LSE preprint, 2012
6. Quickest disorder detection
   Shiryaev. *Optimal stopping rules*, 2000; ch. IV
   Shiryaev. *Quickest detection problems: 50 years later*, 2010

1. General optimal stopping theory

Formulation of an optimal stopping problem

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathrm{P})$ be a filtered probability space and a $G = (G_t)_{t\geqslant 0}$ be a stochastic process on it, where G_t is interpreted as the **gain** if the observation is stopped at time t.

For a given **time horizon** $T \in [0, \infty]$, denote by $\mathfrak{M}_T$ the class of all **stopping times** τ of the filtration $(\mathscr{F}_t)_t$ (i. e. random variables $0 \leqslant \tau \leqslant T$, which are finite a.s. and $\{\omega : \tau(\omega) \leqslant t\} \in \mathscr{F}_t$ for all $t \geqslant 0$).

The **optimal stopping problem**

$$V = \sup_{\tau \in \mathfrak{M}_T} \mathrm{E} G_\tau$$

consists in finding the quantity V and a stopping time $\tau^* \in \mathfrak{M}_T$ at which the supremum is attained (if such τ^* exists).

If G_t is $\mathscr{F}_t$-measurable for each $t \geqslant 0$, we say that the optimal stopping problem V is a **standard** problem.

If G_t is not $\mathscr{F}_t$-measurable, we say that the optimal stopping problem V is a **non-standard** problem.

The general optimal stopping theory is well-developed for **standard problems**. So, non-standard problems are typically solved by a reduction to standard ones.

There are two main approaches to solve standard OS problems:

- **Martingale approach** operates with $\mathscr{F}_t$-measurable functions G_t and is based on
  a) The method of **backward induction** (for the case of discrete time and a finite horizon)
  b) The method of **essential supremum**
- **Markovian approach** assumes that functions G_t have the Markovian representation, i. e. there exists a strong Markov process X_t such that

$$G_t(\omega) = G(t, X_t(\omega))$$

with some measurable functions $G(t, x)$, where $x \in E$ and E is a phase space of X.

The method of backward induction

Suppose we have a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\leqslant N}, \mathrm{P})$, where $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \ldots \subseteq \mathscr{F}_N \subseteq \mathscr{F}$, and a random sequence G_0, $G_1, \ldots, G_N$, such that (for simplicity)

$$\mathrm{E}|G_n| < \infty \text{ for each } n = 0, 1, \ldots, N.$$

Consider the problem

$$V = \sup_{\tau\in\mathfrak{M}_N} \mathrm{E}G_\tau,$$

where here $\mathfrak{M}_N$ is the class of **integer-valued** stopping times $\tau \leqslant N$.

To solve this problem we introduce (by **backward induction**) a special stochastic sequence $S_N, S_{N-1}, \ldots, S_0$:

$$S_N = G_N, \qquad S_n = \max\{G_n, \mathrm{E}(S_{n+1} \mid \mathscr{F}_n)\},$$

which represents the maximum gain that is possible to obtain starting at time n:

- If $n = N$, we have to stop and our gain is $S_N = G_N$.
- If $n < N$, we can either stop or continue. If we stop, our gain is G_n and if we continue our gain is $\mathrm{E}(S_{n+1} \mid \mathscr{F}_n)$.

This reasoning suggests to consider the following candidate for the optimal stopping time:

$$\widetilde{\tau} = \inf\{k \leqslant N : S_k = G_k\},$$

i. e. to stop as soon as the maximum gain we can receive does not exceed the immediate gain.

The main properties of the method of backward induction are given by the following theorem.

Theorem. The stopping time $\widetilde{\tau}$ and the sequence $(S_n)_{n\leqslant N}$ satisfy the properties

(a) $\widetilde{\tau}$ is an **optimal stopping time**;

(b) if τ^* is also optimal, then $\tau^N \leqslant \tau^*$ a.s.;

(c) the sequence $(S_n)_{n\leqslant N}$ is the smallest supermartingale which dominates $(G_n)_{n\leqslant N}$;

(d) the stopped sequence $(S_{n\wedge\widetilde{\tau}})_{n\leqslant N}$ is a martingale.

The method of essential supremum

This method extends the method of backward induction. Here, we consider only a continuous-time problem

$$V_t^T = \sup_{t \leqslant \tau \leqslant T} \mathrm{E} G_\tau.$$

We make the following assumptions:

1. $(G_t)_{t \geqslant 0}$ is right-continuous and left-continuous over stopping time (if $\tau_n \uparrow \tau$ then $G_{\tau_n} \to G_\tau$ a.s.);
2. $\mathrm{E} \sup\limits_{0 \leqslant t \leqslant T} |G_t| < \infty$ (where $G_\infty = 0$ if $T = \infty$).

Consider the **Snell's envelope** $S = (S_t)_{t \geqslant 0}$ of the process G_t,

$$S_t = \operatorname*{ess\,sup}_{\tau \geqslant t} \mathrm{E}(G_\tau \mid \mathscr{F}_t),$$

and define the Markov time

$$\widetilde{\tau}_t = \inf\{u \geqslant t : S_u = G_u\}, \qquad \text{where } \inf \emptyset := \infty.$$

Theorem If assumptions 1-2 above hold, and $\mathrm{P}(\widetilde{\tau}_t < \infty)$ for any $t \geqslant 0$, then

(a) $\widetilde{\tau}_t$ is an **optimal stopping time** for V_t;

(b) if τ_t^* is also optimal for V_t, then $\widetilde{\tau}_t \leqslant \tau_t^*$ a.s.;

(c) the process $(S_u)_{u \geqslant t}$ is the smallest right-continuous supermartingale which dominates $(G_u)_{u \geqslant t}$;

(d) the stopped process $(S_{u \wedge \widetilde{\tau}_t})_{u \geqslant t}$ is a right-continuous martingale;

(e) if $\mathrm{P}(\widetilde{\tau}_t = \infty) > 0$ then there is **no** optimal stopping time for V_t.

Markovian approach

Let $(X_t)_{t \geqslant 0}$ be a homogeneous strong Markov process on a probability space

$$(\Omega, \mathscr{F}, \mathrm{P}_x),$$

where $x \in E \, (= \mathbb{R}^d)$, $\mathrm{P}_x(X_0 = x) = 1$, and $x \mapsto \mathrm{P}_x(A)$ is measurable for each $A \in \mathscr{F}$.

Consider an optimal stopping problem in the **Markovian setting**:

$$V(x) = \sup_{\tau \in \mathfrak{M}} \mathrm{E}_x G(X_\tau), \qquad x \in E,$$

where $\mathfrak{M}$ is the class of stopping times (finite P_x-a.s for each $x \in E$) of the filtration $(\mathscr{F}_t^X)_{t \geqslant 0}$, $\mathscr{F}_t^X = \sigma(X_s; s \leqslant t)$.

$G(x)$ is called the **gain function**, $V(x)$ is called the **value function**.

For simplicity, we assume that $\mathrm{E}_x \sup\limits_{t \geqslant 0} |G(X_t)| < \infty$ for each $x \in E$.

Introduce the two sets:

$$\textbf{continuation set } C = \{x \in E : V(x) > G(x)\},$$
$$\textbf{stopping set } D = \{x \in E : V(x) = G(x)\}.$$

It turns out that under rather general conditions the optimal stopping time in problem $V(x)$ is the **first entry time** to the stopping set:

$$\tau_D = \inf\{t \geqslant 0 : X_t \in D\}$$

(τ_D is a Markov time if X is right-continuous and D is closed).

Sufficient and necessary conditions for this fact can be found in the first part of the course and in the book *Optimal Stopping and Free-boundary Problems* by Peskir, Shiryaev (2006).

Remark. The case of a finite time horizon or a non-homogeneous process X can be reduced to the above case by increasing the dimension of the problem and considering the process (t, X_t).

In fact, for a homogeneous Markovian optimal stopping problems with the infinite time horizon it is often easier

to "guess" a candidate solution

and then to **verify** that it is indeed a solution.

The basic idea is that $V(x)$ and C (or D) should solve the **free-boundary problem**

$$\begin{cases} \mathscr{L}_X V \leqslant 0, \\ V \geqslant G \quad (V > G \text{ in } C \text{ and } V = G \text{ in } D), \end{cases}$$

where $\mathscr{L}_X$ is the characteristic (infinitesimal) operator of X.

If X is a diffusion process, G is sufficiently smooth in a neighborhood of ∂C, and ∂C is "nice", then the above system of inequalities splits into the conditions

$$\begin{cases} \mathscr{L}_X V = 0 \text{ in } C \\ V = G \text{ in } D \\ \left.\dfrac{\partial V}{\partial x}\right|_{\partial C} = \left.\dfrac{\partial G}{\partial x}\right|_{\partial C} \quad \text{(smooth fit)} \end{cases}$$

If X is a homogeneous 1-dimensional process, usually it is possible to find the solution (V and ∂C) of this system **explicitly**, and then to **verify** that it is also a solution of the optimal stopping problem.

The verification is typically based on some manipulations with the Itô formula – the general idea will become clear when we consider specific problems below.

Extension: integral functionals and discounting

Suppose, as above, that X_t is a homogeneous strong Markov process and consider a more general optimal stopping problem:

$$V(x) = \sup_{\tau \in \mathfrak{M}} \mathrm{E}_x \left[e^{-\lambda_\tau} G(X_\tau) + \int_0^\tau e^{-\lambda_s} L(X_s) ds \right],$$

where $L(x)$ is a function, and the **discounting process** $\lambda = (\lambda_t)_{t \geqslant 0}$ is given by

$$\lambda_t = \int_0^t \lambda(X_s) ds,$$

with some function $\lambda(x) \colon E \to \mathbb{R}_+$.

If the functions G, L, λ are "nice", we obtain the following modification of the condition on V in C (in addition to the condition $V = G$ in D and the smooth fit condition):

$$\mathscr{L}_X V(x) - \lambda(x) V(x) = -L(x) \text{ in } C.$$

2. Stopping a Brownian motion at its maximum

Let $B = (B_t)_{t \geqslant 0}$ be a standard Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathrm{P})$. Define the Brownian motion $B^\mu = (B_t^\mu)_{t \geqslant 0}$ with drift μ and its running maximum $S^\mu = (S_t^\mu)_{t \geqslant 0}$:

$$B_t^\mu = \mu t + B_t, \qquad S_t^\mu = \sup_{0 \leqslant s \leqslant t} B_s^\mu.$$

Let θ be the (P-a.s. unique) time at which the maximum of B^μ on $[0,1]$ is attained (i. e. $B_\theta^\mu = S_1^\mu$).

We consider the following **optimal stopping problem**:

$$V^\mu = \inf_{\tau \in \mathfrak{M}_1} E|\tau - \theta|,$$

where $\mathfrak{M}_1$ is the class of all stopping times $\tau \leqslant 1$ of the process B^μ.

Financial interpretation

Suppose the price of stock is described by a geometric Brownian motion:

$$dX_t = \mu X_t dt + \sigma X_t dB_t \iff X_t = X_0 \exp\big((\mu - \tfrac{\sigma^2}{2})t + \sigma B_t\big)$$

If one holds the stock at time $t = 0$ and wants to **sell** it until time $t = 1$ then it would be the best to sell it at the time θ, when X attains its maximum on $[0, 1]$.

However

θ is not a stopping time,

so it is impossible to sell the stock at the time θ.

Thus, it is natural to consider the problem V^μ of finding a **stopping time**, which is as close as possible to θ.

(X attains its maximum when $B^{\widetilde{\mu}} = \log X$ attains its maximum).

Another approach could be to consider **expected utility maximization**:

$$\mathrm{E}U(S_\tau) \to \max \text{ over } \tau \in \mathfrak{M}_1,$$

where $U(x)$ is some utility function.

For example, for $U_\alpha(x) = x^\alpha$, $\alpha \in (0,1]$, or $U_0(x) = \log(x)$ we have

- $\tau^* = 0$ if $\mu \leqslant (1-\alpha)\sigma^2/2$ (in this case S_t^α is a supermartingale)
- $\tau^* = T$ if $\mu \geqslant (1-\alpha)\sigma^2/2$ (S_t^α is a submartingale)

However, the choice of a utility function is **subjective**; often it is not clear why we should prefer one utility function over another.

In contrast, the criterion $\mathrm{E}|\tau - \theta| \to \min$ **has a clear interpretation**.

Solution of the problem V^0

First we consider the case of a Brownian motion *without drift* ($\mu = 0$), as the case $\mu \neq 0$ is considerably more difficult.

$$\boxed{V^0 = \inf_{\tau \in \mathfrak{M}_1} \mathrm{E}|\tau - \theta|, \qquad \theta = \arg\max_{0 \leqslant s \leqslant 1} B_s.}$$

Observe that the optimal stopping problem V^0 is **non-standard** according to our terminology (because θ is not measurable w.r.t $\mathscr{F}_\tau$). Thus, our first step will be to reduce it to a **standard** problem.

The main result about the solution of the problem V^0 is as follows.

Theorem. *The value V^0 is given by the formula*

$$V^0 = 2\Phi(z_*) - 1 = 0.73\ldots$$

where $z_* = 1.12\ldots$ is the unique root of the equation

$$4\Phi(z) - 2z\varphi(x) - 3 = 0.$$

The following stopping time is optimal:

$$\tau_* = \inf\{0 \leqslant t \leqslant 1 : S_t - B_t \geqslant z_*\sqrt{1-t}\}$$

Proof of the theorem

Step 1. First we reduce the problem V^0 to a **standard** optimal stopping problem by showing that for any stopping time $\tau \in \mathfrak{M}_1$ we have

$$\mathrm{E}|\tau - \theta| = \mathrm{E}\left[\int_0^\tau F\left(\frac{S_t - B_t}{\sqrt{1-t}}\right) dt\right] + \frac{1}{2}, \qquad (*)$$

where $F(x) = 4\Phi(x) - 3$.

Indeed, observe that

$$\begin{aligned} |\tau - \theta| &= (\tau - \theta)^+ + (\tau - \theta)^- = (\tau - \theta)^+ + \theta - \tau \wedge \theta \\ &= \int_0^\tau \mathbf{I}(\theta \leqslant t) dt + \theta - \int_0^\tau \mathbf{I}(\theta > t) dt \\ &= \theta + \int_0^\tau (2\mathbf{I}(\theta \leqslant t) - 1) dt. \end{aligned}$$

Setting $\pi_t = \mathrm{P}(\theta \leqslant t \mid \mathscr{F}_t^B)$ and taking E of the both sides of the previous equation we get

$$\begin{aligned}
\mathrm{E}|\tau - \theta| &= \mathrm{E}\theta + \mathrm{E}\int_0^\tau (2\mathbf{I}(\theta \leqslant t) - 1)dt \\
&= \frac{1}{2} + \mathrm{E}\int_0^\infty (2\mathrm{P}(\theta \leqslant t \mid \mathscr{F}_t^B) - 1)\mathbf{I}(t \leqslant \tau)dt \\
&= \frac{1}{2} + \mathrm{E}\int_0^\tau (2\pi_t - 1)dt.
\end{aligned}$$

By stationary and independent increments of B we get

$$\begin{aligned}
\pi_t &= \mathrm{P}\big(S_t \geqslant \max_{t\leqslant s\leqslant 1} B_s \mid \mathscr{F}_t^B\big) = \mathrm{P}\big(S_t - B_t \geqslant \max_{t\leqslant s\leqslant 1} B_s - B_t \mid \mathscr{F}_t^B\big) \\
&= \mathrm{P}(z - x \geqslant S_{1-t})\big|_{z=S_t, x=B_t} = 2\Phi\left(\frac{S_t - B_t}{\sqrt{1-t}}\right) - 1.
\end{aligned}$$

Inserting this into the above formula, we get $(*)$.

Step 2. To solve the optimal stopping problem

$$V = \inf_{\tau \in \mathfrak{M}_1} \mathrm{E} \left[\int_0^\tau F\left(\frac{S_t - B_t}{\sqrt{1-t}} \right) dt \right] + \frac{1}{2}$$

we first note that the filtrations of the processes $S - B$ and B coincide, so we need to consider only stopping times τ of $S - B$.

According to the **Lévy's distributional theorem**,

$$\mathrm{Law}(S - B) = \mathrm{Law}(|B|),$$

so we get an equivalent problem

$$V = \inf_{\tau \leqslant 1} \mathrm{E} \left[\int_0^\tau F\left(\frac{|B_t|}{\sqrt{1-t}} \right) dt \right] + \frac{1}{2},$$

where the supremum is taken over stopping times $\tau \leqslant 1$ of B.

Step 3. To solve the above problem, we make use of a deterministic **change of time**.

Introduce the process $Z = (Z_t)_{t\geqslant 0}$ by

$$Z_t = e^t B_{1-e^{-2t}}.$$

By Itô's formula we find that Z solves the SDE

$$dZ_t = z_t dt + \sqrt{2} d\beta_t,$$

where the process $\beta = (\beta_t)_{0\leqslant t\leqslant 1}$ is given by

$$\beta_t = \frac{1}{\sqrt{2}} \int_0^t e^s dB_{1-e^{-2s}} = \frac{1}{\sqrt{2}} \int_0^{1-e^{-2t}} \frac{1}{\sqrt{1-s}} dB_s.$$

Observe that β is a continuous Gaussian martingale with zero mean and variance equal to t, so according to **Lévy's characterization theorem**,

$$\boldsymbol{\beta} \textbf{ is a standard Brownian motion.}$$

Next we pass from the **old time t** to the **new time s** by the formula

$$t = 1 - e^{-2s} \iff s = \log(1/\sqrt{1-t})$$

and we get

$$V = 2 \inf_{\tau \in \mathfrak{M}} \mathrm{E}\left[\int_0^{\sigma_\tau} e^{-2s} F(|Z_s|)ds\right] + \frac{1}{2},$$

where $\sigma_\tau = \log(1/\sqrt{1-\tau})$ and σ_τ is a stopping time w.r.t $(\mathscr{F}_s^Z)_{s\geqslant 0}$.

Thus, we need to solve the **Markovian** optimal stopping problem

$$W = \inf_{\sigma \geqslant 0} \mathrm{E}\left[\int_0^{\sigma} e^{-2s} F(|Z_s|)ds\right]$$

for the diffusion process Z, which has the infinitesimal generator

$$\mathscr{L}_Z = z\frac{d}{dz} + \frac{d^2}{dz^2}.$$

Step 4. Following the general theory, introduce the value function

$$W(z) = \inf_{\sigma \geqslant 0} \mathrm{E}_z \left[\int_0^{\sigma} e^{-2s} F(|Z_s|) ds \right].$$

The function $F(z) = 4\Phi(z) - 3$ is increasing for $z \geqslant 0$, which allows us **to guess** that the optimal stopping time σ^* should be of the form

$$\sigma^* = \inf\{t > 0 : |Z_t| \geqslant z_*\},$$

where $z_* > 0$ is a constant to be found.

In order to find z_* we formulate the associated **free-boundary problem**:

$$\begin{cases} (\mathscr{L}_Z - 2)W_*(z) = -F(|z|) \text{ for } z \in (-z_*, z_*), \\ W_*(\pm z_*) = 0 \quad \textit{(instantaneous stopping)} \\ W_*'(\pm z_*) = 0 \quad \textit{(smooth fit).} \end{cases}$$

Inserting $\mathscr{L}_Z$, the first equation transforms to

$$W_*''(z) + zW_*'(z) - 2W_*(z) = -F(z) \text{ for } z \in (-z_*, z_*).$$

The general solution of this equation is

$$W_*(z) = C_1(1+z^2) + C_2(z\varphi(x) + (1+z^2)\Phi(z)) + 2\Phi(z) - \tfrac{3}{2}.$$

From the formulation of the optimal stopping problem, it is clear that W_* should be an even function, and hence $W_*'(0) = 0$.

Using the conditions $W_*(z_*) = W_*'(z_*) = W_*'(0) = 0$ we find $C_1 = \Phi(z_*)$, $C_2 = -1$ and z_* is the unique root of the equation

$$4\Phi(z) - 2z\varphi(x) - 3 = 0.$$

Consequently,

$$W_*(z) = \Phi(z_*)(1+z^2) + -z\varphi(x) + (1-z^2)\Phi(z)) - \tfrac{3}{2}, \quad z \in [0, z_*].$$

Step 5. Now we need **to verify** that the solution $W_*(z)$ of the free-boundary problem is equal to the value function $W(z)$.

Observe that $W_*(z)$ is C^2 everywhere except at $\pm z_*$ where it is C^1. By the Itô–Tanaka–Meyer formula we find

$$\begin{aligned} e^{-2t}W_*(Z_t) = W_*(Z_0) &+ \int_0^t e^{-2s}\big(\mathscr{L}_Z W_*(Z_s) - 2W_*(Z_s)\big)ds \\ &+ \sqrt{2}\int_0^t e^{-2s}W_*'(Z_s)d\beta_s. \end{aligned} \tag{**}$$

Using that $\mathscr{L}_Z W_*(z) - 2W_*(z) = -F(|z|)$ for $z \in (-z_*, z_*)$, and $\mathscr{L}_Z W_*(z) - 2W_*(z) = 0 > -F(|z|)$ for $z \notin (-z_*, z_*)$, we get

$$e^{-2t}W_*(Z_t) \geqslant W_*(Z_0) - \int_0^t e^{-2s}F(|Z_s|)ds + \mathrm{Mart}_t.$$

Since $W_*(z) \leqslant 0$ for all z, applying the optional sampling theorem under P_z, we get that $W_*(z) \leqslant W(z)$ for all z.

Analyzing $(**)$ we also find that

$$0 = W_*(Z_0) - \int_0^{\sigma_*} e^{-2s} F(|Z_s|) ds + \mathrm{Mart}_{\sigma_*}.$$

Taking the expectation E_z, we get $W_*(z) \geqslant W(z)$, which implies $W_*(z) = W(z)$ and completes the proof of the claim.

Transforming σ_* back to the initial problem, we see that τ_* is the optimal stopping time for V.

Distributional properties of τ_*

Using that $\mathrm{Law}(S - B) = \mathrm{Law}(|B|)$, we see that τ^* is distributed as the stopping time

$$\widetilde{\tau} = \inf\{t \geqslant 0 : |B_t| = z_* \sqrt{1-t}\}.$$

This implies

$$\mathrm{E}\tau_* = \mathrm{E}\widetilde{\tau} = \mathrm{E}B_{\widetilde{\tau}}^2 = z_*^2(1 - \mathrm{E}\widetilde{\tau}) = z_*^2(1 - \mathrm{E}\tau_*).$$

Solving this equation for $\mathrm{E}\tau_*$ we find

$$\mathrm{E}\tau_* = z_*^2/(1 + z_*^2) = 0.55\ldots$$

Similarly, using that $(B_t^4 - 6tB_t^2 + 4t^2)$ is a martingale, we find

$$\mathrm{E}\tau_*^2 = \frac{2z_*^4}{(1+z_*^2)^2(3 + 6z_*^2 + z_*^4)} = 0.36\ldots \quad \text{and} \quad \mathrm{Var}\,\tau_* = 0.05\ldots$$

A related problem: minimizing the distance “in space”

It is remarkable that the optimal stopping time τ_* which minimizes the **distance in time** $\mathrm{E}|\tau-\theta|$ also minimizes the **distance in space**:

$$\mathrm{E}(B_{\tau_*}-S_1)^2=\inf_{\tau\leqslant 1}\mathrm{E}(B_\tau-S_1)^2, \qquad (***)$$

where $S_t=\max_{s\leqslant t}B_s$.

To prove (***), observe that S_1 is a square-integrable functional of the Brownian path on $[0,1]$. By the **Itô-Clark representation theorem**, there exists a unique $\mathscr{F}_t^B$-adapted process $H=(H_t)_{t\leqslant 1}$ such that

$$S_1=\mathrm{E}S_1+\int_0^1 H_t\,dB_t, \qquad \text{and} \qquad \mathrm{E}\int_0^1 H_t^2\,dt<\infty.$$

Moreover, the following explicit formula is valid

$$H_t=2\left(1-\Phi\left(\frac{S_t-B_t}{\sqrt{1-t}}\right)\right).$$

Define the square-integrable martingale $M = (M_t)_{t \geqslant 1}$:

$$M_t = \int_0^t H_s dB_s.$$

By the martingale property and the optional sampling theorem, obtain

$$\begin{aligned}\mathrm{E}(B_\tau - S_1)^2 &= \mathrm{E}B_\tau^2 - 2\mathrm{E}(B_\tau M_1) + \mathrm{E}S_1^2 \\ &= \mathrm{E}\tau - 2\mathrm{E}(B_\tau M_\tau) + 1 = \mathrm{E}\left(\int_0^\tau (1 - 2H_t)dt\right) + 1\end{aligned}$$

for each $\tau \in \mathfrak{M}$.

Using the explicit formula for H_t we find

$$\inf_{\tau \in \mathfrak{M}_1} \mathrm{E}(B_\tau - S_1)^2 = \inf_{\tau \in \mathfrak{M}_1} \mathrm{E}\left[\int_0^\tau F\left(\frac{S_t - B_t}{\sqrt{1-t}}\right) dt\right] + 1 = V + \frac{1}{2},$$

hence the optimal stopping times in the both problems coincide.

Remark 1: on Levý's theorem and its generalization

The **Levý's theorem** for a Brownian motion $B = (B_t)_{t \geqslant 0}$ states that

$$(\sup B - B,\ \sup B) \overset{\text{Law}}{=} (|B|,\ L(B)),$$

where $L(B) = (L_t(B))_{t \geqslant 0}$ is the **local time** of B at zero:

$$L_t(B) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{I}(|B_s| \leqslant \varepsilon) ds, \quad t \geqslant 0, \qquad \text{(lim exists a.s.)}$$

Graversen & Shiryaev (2000) showed that the following generalisation holds for a Brownian motion with drift B^μ

$$(\sup B^\mu - B^\mu,\ \sup B^\mu) \overset{\text{Law}}{=} (|X^\mu|,\ L(X^\mu)),$$

where $X^\mu = (X_t^\mu)_{t \geqslant 0}$ is the **bang-bang process**

$$dX_t^\mu = -\mu \operatorname{sgn} X_t^\mu dt + dB_t, \qquad X_0^\mu = 0,$$

and $L(X^\mu)$ is defined in the same way as $L(B)$.

Remark 2: on the Itô-Clark theorem for B^μ

Let B be a standard Brownian motion, and $\mathscr{F}_1^B = \sigma(B_t; t \leqslant 1)$. Suppose ξ is an $\mathscr{F}_1^B$-measurable square-integrable random variable.

The **Itô-Clark theorem** states that there exists a unique predictable process $H = (H_t)_{t \leqslant 1}$ such that $\mathrm{E} \int_0^1 H^2 ds < \infty$ and

$$\xi = \mathrm{E}\xi + \int_0^1 H_s dB_s.$$

In general, it is very difficult to find the process H explicitly.

However, let us show that for $\xi = S_1^\mu \equiv \max\limits_{t \leqslant 1} B_t^\mu$ we have

$$H_t^\mu = 1 - \Phi\left(\frac{(S_t^\mu - B_t^\mu) - \mu(1-t)}{\sqrt{1-t}}\right) + e^{2\mu(S_t^\mu - B_t^\mu)} \Phi\left(\frac{-(S_t^\mu - B_t^\mu) - \mu(1-t)}{\sqrt{1-t}}\right)$$

1°. Using that B^μ has stationary independent increments, we get

$$\begin{aligned}
\mathrm{E}(S_1^\mu \mid \mathscr{F}_t) &= S_t^\mu + \mathrm{E}\Big[\Big(\sup_{t\leqslant s\leqslant 1} B_s^\mu - S_t^\mu\Big)^+ \mid \mathscr{F}_t\Big] \\
&= S_t^\mu + \mathrm{E}\Big[\big(\sup_{t\leqslant s\leqslant 1}(B_s^\mu - B_t^\mu) - (S_t^\mu - B_t^\mu)\big)^+ \mid \mathscr{F}_t\Big] \\
&= S_t^\mu + \mathrm{E}(S_{1-t}^\mu - (z-x))^+\big|_{z=S_t^\mu, x=B_t^\mu}.
\end{aligned}$$

Using the formula $\mathrm{E}(X-c)^+ = \int_c^\infty \mathrm{P}(x>z)dz$, we get

$$\mathrm{E}(S_1^\mu \mid \mathscr{F}_t) = S_t^\mu + \int_{S_t^\mu - B_t^\mu}^\infty (1 - F_{1-t}^\mu(z))dx := f(t, B_t^\mu, S_t^\mu),$$

where $F_{1-t}^\mu(z) = \mathrm{P}(S_{1-t}^\mu \leqslant z)$.

2°. Applying the Itô formula to the right-hand side of the previous equation and using that the left-hand side defines a continuous martingale, we get

$$\begin{aligned}\mathrm{E}(S_1^\mu \mid \mathscr{F}_t) &= \mathrm{E}S_1^\mu + \int_0^t \frac{\partial f}{\partial x}(s, B_s^\mu, S_s^\mu)dB_s \\ &= \mathrm{E}S_1^\mu + \int_0^t (1 - F_{1-t}^\mu(S_s^\mu - B_s^\mu))dB_s\end{aligned}$$

as a nontrivial continuous martingale cannot have paths of bounded variation.

3°. Finally, we use the well-known formula

$$\begin{aligned}F_{1-t}^\mu(z) &:= \mathrm{P}(S_{1-t}^\mu \leqslant z) \\ &\qquad = \Phi\left(\frac{z-\mu(1-t)}{\sqrt{1-t}}\right) - e^{2\mu z}\varphi\left(\frac{-z-\mu(1-t)}{\sqrt{1-t}}\right).\end{aligned}$$

which gives the sought-for representation for H^μ.

The case $\mu \neq 0$

We consider only the problem of minimizing the distance in space:

$$\widetilde{V}^{\mu} = \inf_{\tau \in \mathfrak{M}_1} \mathrm{E}(B^{\mu}_{\tau} - S^{\mu}_1)^2,$$

where

$$B^{\mu}_t = \mu t + B_t, \qquad S_t = \max_{s \leqslant t} B^{\mu}_s.$$

The problem is **not standard**, so first we reduce it to a standard one.

Lemma. *For any $\tau \in \mathfrak{M}_1$ the following identity holds:*

$$\mathrm{E}\big[(S^{\mu}_1 - B^{\mu}_t)^2 \mid \mathscr{F}^B_{\tau}\big) = (S^{\mu}_{\tau} - B_{\tau})^2 + 2\int_{S^{\mu}_{\tau} - B^{\mu}_{\tau}}^{\infty} z(1 - F^{\mu}_{1-\tau}(z))dz,$$

where

$$F^{\mu}_{1-t}(z) = \mathrm{P}(S^{\mu}_{1-t} \leqslant z) = \Phi\left(\frac{z - \mu(1-t)}{\sqrt{1-t}}\right) - e^{2\mu z}\Phi\left(\frac{-z - \mu(1-t)}{\sqrt{1-t}}\right).$$

The process $S_t^\mu - B_t^\mu$ is Markov, and in order to apply the general theory we let it start from an arbitrary point (t,x) by introducing the process

$$X_{t+s}^x = x \vee S_s^\mu - B_s^\mu$$

Then we get the Markovian optimal stopping problem

$$V(t,x) = \inf_{0 \leqslant \tau \leqslant 1-t} \mathrm{E}_{t,x} G(t+\tau, X_{t+\tau}^x),$$

where G is given by

$$G(t,x) = x^2 + 2\int_x^\infty zR(t,z)dz, \qquad R(t,z) = 1 - F_{1-t}^\mu(z).$$

There is **no** closed-form analytical solution of this problem. However, the optimal stopping boundaries can be found numerically from a **system of integral equation**. For details, see Peskir, Shiryaev, *Optimal Stopping and Free-Boundary problems*, sec. 30.

In fact, for $\mu > 0$ the optimal stopping time is given by

$$\tau^* = \inf\{0 \leqslant t \leqslant T : S_t^\mu - B_t^\mu \in [b_1(t), b_2(t)]\}$$

and for $\mu < 0$ the optimal stopping time is given by

$$\tau^* = \inf\{0 \leqslant t \leqslant T : S_t^\mu - B_t^\mu \geqslant b_1(t)\},$$

where $b_1(t)$ and $b_2(t)$ are some functions (dependent on μ) that can be found by solving a system of non-linear **integral equations**.

Next we present the qualitative structure of the stopping and the continuation sets.

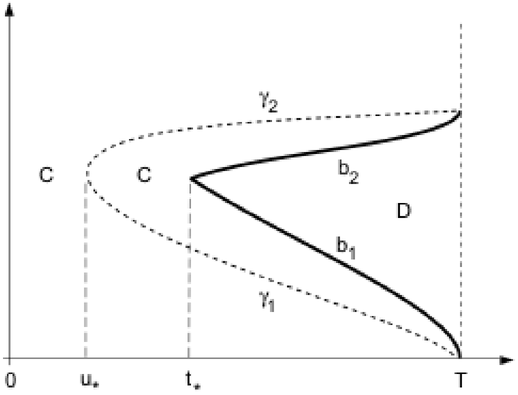


Figure VIII.6: (*The "black-hole" effect.*) A computer drawing of the optimal stopping boundaries b_1 and b_2 when $\mu > 0$ is away from 0.

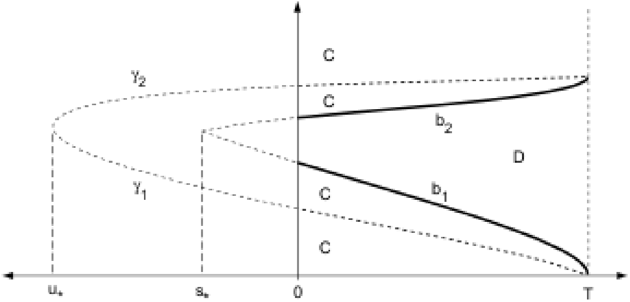


Figure VIII.7: A computer drawing of the optimal stopping boundaries b_1 and b_2 when $\mu \geq 0$ is close to 0.

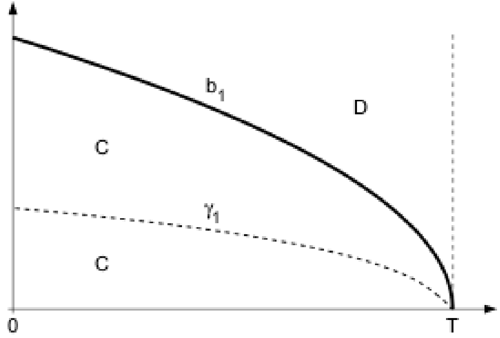


Figure VIII.8: A computer drawing of the optimal stopping boundary b_1 when $\mu \leq 0$ is close to 0.

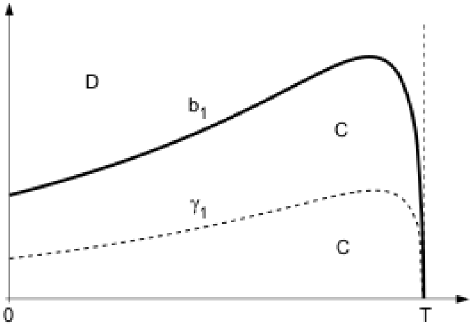


Figure VIII.9: (*The "hump" effect.*) A computer drawing of the optimal stopping boundary b_1 when $\mu < 0$ is away from 0.

3. Trading rule "Buy and Hold"

In this section we consider another optimality criterion when one wants to **sell stock** until a time $T > 0$.

Suppose the (discounted) stock price is described by a geometric Brownian motion

$$dX_t = aX_t dt + \sigma X_t dB_t \iff X_t = X_0 \exp((a - \tfrac{\sigma^2}{2})t + \sigma B_t).$$

and put

$$M_t = \sup_{s \leqslant t} X_s.$$

We consider the following **optimal stopping problem**:

$$W = \sup_{\tau \in \mathfrak{M}_T} \mathrm{E} \frac{X_\tau}{M_T},$$

which means that a trader wants to maximize the average percentage of the maximum possible gain.

This problem is particularly interesting, because it is equivalent to the maximization of the relative error between the selling price and the maximum price:

$$\mathrm{E}\left[\frac{M_T - X_\tau}{M_T}\right] \to \max\,.$$

As it was noted in the previous section, criteria of this type have clear meaning, unlike the standard approach of maximizing expected utility.

A simpler case: maximizing the logarithmic rate

Before we proceed to the solution of the above problem, let us consider a simpler problem

$$R = \sup_{\tau \in \mathfrak{M}_T} \mathrm{E}\left[\log\left(\frac{X_\tau}{M_T}\right)\right].$$

Clearly,

$$R = \sup_{\tau \in \mathfrak{M}_T} \mathrm{E}[(a - \sigma^2/2)\tau + \sigma B_\tau - \log M_T]$$

and the optimal stopping time is given by

$$\tau^* = \begin{cases} T, & a > \sigma^2/2, \\ \text{any time in } [0, T], & a = \sigma^2/2, \\ 0, & a < \sigma^2/2, \end{cases}$$

where we use that $\mathrm{E}B_\tau = 0$, and $\mathrm{E}M_T$ does not depend on τ.

It will be convenient to introduce the **goodness index** of the stock

$$\alpha = \frac{a}{\sigma^2}.$$

Then the optimal stopping time is

$$\tau^* = \begin{cases} T, & \alpha > 1/2, \\ \text{any time in } [0,T], & \alpha = 1/2, \\ 0, & \alpha < 1/2. \end{cases}$$

Remark. In the undiscounted case, the goodness index is given by

$$\widetilde{\alpha} = \frac{\widetilde{a} - r}{\sigma^2}.$$

where $\widetilde{a}$ is the drift coefficient of the undiscounted price, and r is the discounting rate.

Solution of the main problem

Remarkably, the same **deterministic** stopping time is optimal in the original problem

$$W = \sup_{\tau \in \mathfrak{M}_T} \mathrm{E}\left[\frac{X_\tau}{M_T}\right],$$

Theorem. *1) If $\alpha < 0$, then $\tau^* = 0$ is the unique optimal stopping time for W and the optimal relative error is given by*

$$\begin{aligned} W(\alpha,\sigma) = 1 - \frac{2\alpha-1}{2(\alpha-1)}\Phi\big[-(\alpha-1/2)\sigma\sqrt{T}\big] \\ - \frac{2\alpha-3}{2(\alpha-1)}e^{(1-\alpha)\sigma^2 T}\Phi\big[(\alpha-3/2)\sigma\sqrt{T}\big]. \end{aligned}$$

Theorem (cont.). *2) If $\alpha \geqslant 1/2$ then $\tau^* = T$ is the unique optimal stopping time for W when $\alpha > 1/2$, and both $\tau^* = 0$ and $\tau^* = T$ are optimal when $\alpha = 1/2$. The optimal relative error is given by*

$$\begin{aligned} W(\alpha, \sigma) = 1 &- \Big(1 - \frac{1}{2\alpha}\Big)\Phi\big[(\alpha - 1/2)\sigma\sqrt{T}\big] \\ &- \Big(1 + \frac{1}{2\alpha}\Big)e^{\alpha\sigma^2 T}\Phi\big[-(\alpha + 1/2)\sigma\sqrt{T}\big]. \end{aligned}$$

Moreover, $W(\alpha, \sigma)$ decreases in α and increases in σ and

$$0 \leqslant W(\alpha, \sigma) < \frac{1}{2\alpha} \text{ for any } \alpha \geqslant 1/2,\ \sigma \geqslant 0.$$

Remark. As in the original paper by Shiryaev, Xu, Zhou, we omit the case $\alpha \in (0, 1/2)$, which can be solved by the PDE approach, in favor of the probabilistic approach that will be used.

The answer for the case $\alpha < 1/2$ is $\tau^* = 0$.

Discussion of the result

1. When the stock goodness index $\alpha \geqslant 1/2$, one should hold on to the stock, i. e. the **stock is good** one. The better the stock (as measured by α) the smaller the relative error.

In particular, the error diminishes to zero when $\alpha \to \infty$, so the buy-and-hold rule almost realizes selling at the maximum price if the stock is sufficiently good.

2. If $\alpha < 1/2$, one should sell the stock immediately. This is **bad stock** the investor ought to get rid of as soon as possible.

3. The buy-and-hold rule is **insensitive** to the stock parameters as the definition of good and bad stocks involves a **range** of the parameters, instead of specific values for them.

Proof of the theorem

Using the self-similarity property of B, we can assume $\sigma = 1$.

Observe that

$$W = \sup_{\tau \leqslant T} \mathrm{E}\left[\frac{e^{B^\mu_\tau}}{e^{S^\mu_T}}\right],$$

where

$$B^\mu_t = \mu t + B_t, \qquad S^\mu_t = \sup_{s \leqslant t} B^\mu_s, \qquad \text{where } \mu = a - \sigma^2/2.$$

This is a **non-standard** optimal stopping problem, so we reduce it to a standard one:

$$\begin{aligned}
\mathrm{E}\big[e^{B_\tau^\mu}/e^{S_T^\mu}\big] &= \mathrm{E}\Big[\min\Big\{e^{-(S_\tau^\mu - B_\tau^\mu)}, e^{-\max\limits_{\tau\leqslant t\leqslant T}(B_t^\mu - B_\tau^\mu)}\Big\}\Big] \\
&= \mathrm{E}\Big[\mathrm{E}\Big[\min\Big\{e^{-(S_\tau^\mu - B_\tau^\mu)}, e^{-\max\limits_{\tau\leqslant t\leqslant T}(B_t^\mu - B_\tau^\mu)}\Big\}\Big|\,\mathscr{F}_t\Big]\Big] \\
&= \mathrm{E}\Big[\mathrm{E}\Big[\min\Big\{e^{-x}, e^{-S_{T-t}^\mu}\Big\}\Big]\,\Big|_{x=S_\tau^\mu - B_\tau^\mu}\Big] \\
&= \mathrm{E}[G(\tau, S_\tau^\mu - B_\tau^\mu)],
\end{aligned}$$

where

$$G(t,x) := \mathrm{E}\big[\min\big\{e^{-x}, e^{-S^{\mu}_{T-t}}\big\}\big] > 0.$$

Direct computations show that for $\mu \neq 1/2$

$$G(t,x) = \tfrac{2(\mu-1)}{2\mu-1} e^{-(\mu-1/2)(T-t)} \Phi\left(\tfrac{-x+(\mu-1)(T-t)}{\sqrt{T-t}}\right) + \tfrac{1}{2\mu-1} e^{-(1-2\mu)x} \Phi\left(\tfrac{-x-\mu(T-t)}{\sqrt{T-t}}\right) + e^{-x} \Phi\left(\tfrac{x-\mu(T-t)}{\sqrt{T-t}}\right)$$

and for $\mu = 1/2$

$$G(t,x) = [1 + x + (T-t)/2] \Phi\left(\tfrac{-x-(T-t)/2}{\sqrt{T-t}}\right) - \sqrt{\tfrac{T-t}{2\pi}} e^{-(x+(T-t)/2)^2/(2(T-t))} + e^{-x} \Phi\left(\tfrac{x-\mu(T-t)/2}{\sqrt{T-t}}\right).$$

Thus we need to solve the **standard** problem

$$W = \sup_{\tau \in \mathfrak{M}_T} \mathrm{E} G(\tau, X_\tau), \qquad \text{where } X_t = S^{\mu}_t - B^{\mu}_t.$$

Introduce the **value function**, letting the process X start from an arbitrary point (t, x):

$$W(t,x) = \sup_{\tau \leqslant T-t} \mathrm{E} G(t+\tau, X_\tau^x), \qquad 0 \leqslant t \leqslant T,\ x \geqslant 0.$$

where $X_t^x = x \vee S_t^\mu - B_t^\mu$.

We know that the optimal stopping time is the **first entry time** to the stopping set D:

$$D = \{(t,x) : W(t,x) = G(t,x)\}, \qquad \tau^* = \inf\{t \geqslant 0 : (t, X_t) \in D\}.$$

Thus, in order to solve the problem, we need to analyze the structure of the functions W and G.

Next we consider two cases, which differ in the methods used:

1. $\alpha \geqslant 1$ or $\alpha \leqslant 0$,
2. $1/2 \leqslant \alpha < 1$.

Case 1: $\alpha \geqslant 1$ or $\alpha \leqslant 0$

Observe that $\mathrm{Law}(X^x) = \mathrm{Law}(|Y|)$, where $Y = (Y_t)_{t\geqslant 0}$ is the **bang-bang process**

$$dY_t = -\mu\,\mathrm{sgn}(Y_t)dt + d\widetilde{B}_t, \qquad Y_0 = x,$$

with a Brownian motion $\widetilde{B}$ (may be different from B).

By the Itô-Tanaka formula we obtain

$$\begin{aligned} G(t+s, |Y_s|) = {} & G(t,x) + \int_0^s \mathscr{L}_Y G(t+u, |Y_u|)du \\ & + \int_0^s G'_x(t+u, |Y_u|)\,\mathrm{sgn}(Y_u)d\widetilde{B}_u \\ & + \int_0^s G'_x(t+u, |Y_u|)dL_u(Y) \\ = {} & G(t,x) + \int_0^s H(t+u, |Y_u|)du + m_s, \end{aligned}$$

where we used that $G'_x(t, 0+) = 0$ and $H(t,x)$ and m_s are defined by

...

$$H(t,x) := \mathscr{L}_Y G(t,x) \equiv G'_t(t,x) - \mu G'_x(t,x) + \frac{1}{2} G''_{xx}(t,x)$$

$$m_s = \int_0^s G'_x(t+u, |Y_u|)\,\mathrm{sgn}(Y_u) d\widetilde{B}_u.$$

One can show that $-1 \leqslant G'_x \leqslant 0$, so m_s is a martingale, which implies

$$W(t,x) = G(t,x) + \sup_{\tau \leqslant T-t} \mathrm{E}\left[\int_0^\tau H(t+u, X_u^x) du\right]. \qquad (*)$$

Then, a lengthy calculation shows that

$$H(t,x) = (\mu - 1/2) G(t,x) - G'_x(t,x).$$

If $\mu \geqslant 1/2 \Leftrightarrow \alpha \geqslant 1$, then $H(t,x) \geqslant 0$ since $G'_x(t,x) \leqslant 0$ by the monotonicity of G in x, and the inequality is strict if $\alpha > 1$.

This proves that $\tau^* = T - t$ is optimal in $(*)$, and hence, $\tau^* = T$ is optimal in the original problem.

In the case $\mu \leqslant -1/2 \Leftrightarrow \alpha \leqslant 0$ observe that

$$e^x G(t,x) = e^x \mathrm{E}\big[\min\big\{e^{-x}, e^{-S^{\mu}_{T-t}}\big\}\big] = \mathrm{E}\big[\min\big\{1, e^{-S^{\mu}_{T-t}+x}\big\}\big]$$

is strictly increasing w.r.t x, so

$$\frac{\partial(e^x G(t,x))}{\partial x} > 0, \text{ or } G'_x(t,x) + G(t,x) > 0.$$

Thus

$$\begin{aligned} H(t,x) &= (\mu - 1/2)G(t,x) - G'_x(t,x) \\ &= (\mu + 1/2)G(t,x) - (G(t,x) + G'_x(t,x)) < 0. \end{aligned}$$

The inequality $H(t,x) < 0$ implies $\tau^* = 0$ is the optimal stopping time.

Case 2: $1/2 \leqslant \alpha < 1$

The direct approach of case 1 does not work in case 2, and we provide another solution (which, in fact, applies also for $\alpha \geqslant 1$).

Lemma. *If* $\alpha > 1/2$ *then*

$$W(t,x) > G(t,x), \qquad t \in [0,T),\ x \geqslant 0.$$

If $\alpha = 1/2$ *then*

$$W(t,x) > G(t,x), \qquad t \in [0,T),\ x > 0.$$

Observe that directly from this lemma it follows that $\tau^* = T$ is optimal for $\alpha > 1/2$ by the definition of the stopping set $D = \{(t,x) : W(t,x) = G(t,x)\}$.

The case $\alpha = 1/2$ requires some additional reasoning, and will not be covered in the lectures – for details see the paper of Shiryaev et al.

Overview of proof. Note that $a \geqslant 1/2$ is equivalent to $\mu \geqslant 0$, and we assume $\sigma = 1$.

First, using the explicit formula for $G(t,x)$ provided above, one can show that in the case $\mu = 0$ we have

$$\mathrm{E}G(T, X_T^x) > G(0,x) \text{ for } x > 0, \qquad \mathrm{E}G(T, X_T^0) = G(0,0).$$

This implies that $W(t,x) > G(t,x)$ whenever $\alpha = 1/2$, $x > 0$ proving the second statement of the lemma.

Next, for $\mu > 0$ applying Girsanov's theorem we have

$$\mathrm{E}G(T, X_T^\alpha) - G(0,x) = \mathrm{E}\big[e^{-x \vee S_T}(e^{B_T} - 1)e^{-\mu^2 T/2 + \mu B_T}\big]$$

Then

$$\frac{\partial}{\partial \mu}\Big(e^{\mu^2 T/2}\big\{\mathrm{E}G(T, X_T^x) - G(0,x)\big\}\Big)$$
$$= \mathrm{E}\big[e^{-x \vee S_T}(e^{B_T} - 1)B_T e^{\mu B_T}\big] > 0, \quad \mu > 0,$$

which, together with the second statement, implies the first one.

Remark. Let $B = (B_t)_{t \geqslant 0}$ be a standard Brownian motion and $B_t^\mu = B_t + \mu t$ be a Brownian motion with drift μ.

Recall that the **Girsanov theorem** implies that for any measurable "good" functional $G_t(x)$ (e. g. non-negative or bounded) it holds that

$$\mathrm{E} G_t(B^\mu) = \mathrm{E} e^{\mu B_t - \mu^2 t/2} G_t(B).$$

Thus, we have proved that $\tau^* = T$ is optimal if $\alpha \geqslant 1/2$ and $\tau^* = 0$ is optimal if $\alpha < 0$.

We also need to find the relative errors

$$W^*(\alpha, \sigma) = \mathrm{E}\left[X_{\tau^*}/M_T\right].$$

Since the optimal stopping time is **deterministic**, W^* can be found using the explicit formula for the joint density of (X_t, M_t):

$$\begin{aligned}\mathrm{P}(X_t \in dx, M_t \in dm) = {} & \frac{2}{\sigma^3\sqrt{2\pi t^3}} \frac{\log(m^2/s)}{xm} \\ & \times \exp\Bigl(-\frac{\log^2(m^2/x)}{2\sigma^2 t} + \frac{\beta}{\sigma}\log(x) - \frac{1}{2}\beta^2 t\Bigr),\end{aligned}$$

where $\beta = a/\sigma - \sigma/2$ (see Karatzas & Shreve (1991), p. 368).

Details of the derivation of the formulas for W^* are in the paper by Shiryaev et al.

Some extensions

1. In the case $\alpha \in (0, 1/2)$ the optimal stopping time is is $\tau^* = 0$.
2. J. du Toit & G. Peskir (2009) considered also the problem

$$\mathrm{E}(M_T / X_\tau) \to \min$$

   and it turned out that the solution is **not deterministic** if $\alpha \in (0, 1)$, but deterministic in other cases.
   For $\alpha \in (0, 1)$ it can be represented as the first hitting time of M_t / Z_t to a boundary characterized by some integral equation.
3. K. Ano & R. V. Ivanov (2012) generalized the result to α-stable Levý processes, and found that the optimal stopping time is $\tau^* = 0$ or $\tau^* = T$.

Applications to real data

1. Shiryaev, Xu, Zhou. *Thou shalt buy and hold*, 2008.

The original paper provides an example of applying the buy-and-hold rule to the S&P500 index fund based on the data for 1889–1978.

The parameters (estimated by annual data) are $a = 6.18\%$, $\sigma = 16.67\%$, $\alpha = 2.2239 > 0.5$.

If one takes $T = 1$ (year), then $W^*(\alpha, \sigma) = 10.15\%$, i. e. if you buy and hold the S&P500 index fund for 1 year, you can expect to achieve almost 90% of the maximum possible return.

2. Hui, Yam, Chen. *Shiryaev-Zhou index – a noble approach to benchmarking and analysis of real estate stocks*, 2012.

This paper applies the buy-and-hold rule to real estate stock in Hong Kong.

4. Sequential hypothesis testing

Development of **sequential methods** of mathematical statistics started in the 1940s and was largely influenced by the book of A. Wald "Sequential analysis" (1947).

In contrast to classical statistical methods, where sample size is fixed, in sequential methods one can choose the sample size depending on observed data.

Typically, this opportunity leads to a smaller average sample size, while maintaining the same probabilities of errors.

We consider two models of observable processes $(X_t)_{t \geqslant 0}$, which are generated by a **Brownian motion** $(B_t)_{t \geqslant 0}$.

Model A: (related to the problem of hypotheses testing and estimation)

$$\boxed{X_t = \mu t + B_t} \qquad \text{or, equivalently,} \quad dX_t = \mu dt + dB_t,$$

where μ is an **unknown** parameter.

Model B: (related to the problem of detecting a disorder)

$$\boxed{X_t = \mu(t-\theta)^+ + B_t} \qquad \text{or} \quad dX_t = \begin{cases} dB_t, & t < \theta, \\ \mu dt + dB_t, & t \geqslant \theta, \end{cases}$$

where μ is a known parameter, but θ is **unknown**.

Hypotheses testing for a Brownian motion

Let $B = (B_t)_{t \geqslant 0}$ be a Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathrm{P})$. Suppose we sequentially observe the process $X = (X_t)_{t \geqslant 0}$

$$X_t = \mu t + B_t,$$

where $\mu = \mu_0$ or $\mu = \mu_1$ is the unknown drift coefficient.

A **decision rule** for testing the hypotheses $H_0 \colon \mu = \mu_0$ and $H_1 \colon \mu = \mu_1$ is a pair (τ, d), where

- τ is a stopping time of the filtration $(\mathscr{F}_t^X)_{t \geqslant 0}$, $\mathscr{F}_t^X = \sigma(X_s; s \leqslant t)$;
- d is an $\mathscr{F}_\tau$-measurable function taking values $\{\mu_0, \mu_1\}$.

The time τ is interpreted as the moment of stopping the observation, and d corresponds to the hypothesis accepted at time τ.

Generally, we want to find decision rules with a small observation time and a small rate of wrong decisions.

The Bayesian sequential testing problem

Suppose that μ is a **random variable** independent of B and taking values μ_1, μ_0 with probabilities π, $1-\pi$.

The **Bayesian sequential testing problem** of H_1 and H_2 consists in finding the decision rule (τ^*, d^*) which minimizes (over all decision rules) the Bayesian risk

$$R(\tau, d) = a\mathrm{P}(d = \mu_0, \mu = \mu_1) + b\mathrm{P}(d = \mu_1, \mu = \mu_0) + c\mathrm{E}\tau,$$

where $a, b, c > 0$ are given numbers: a, b are interpreted as penalties for wrong decisions, and c as observation cost.

The conditionally-extremal problem (Wald's problem)

Suppose μ is an unknown **real number**.

Let $\Delta(\alpha, \beta)$ denote the class of all decision rules with the probabilities of errors not exceeding α and β respectively:

$$(\tau, d) \in \Delta(\alpha, \beta) \;\Leftrightarrow\; \mathrm{P}(d = \mu_0 \mid \mu = \mu_1) \leqslant \alpha,\; \mathrm{P}(d = \mu_1 \mid \mu = \mu_0) \leqslant \beta.$$

The **conditionally-extremal sequential testing problem** of H_1, H_2 consists in finding $(\tau^*, d^*) \in \Delta(\alpha, \beta)$ such that

$$\mathrm{E}^0\tau^* \leqslant \mathrm{E}^0\tau, \quad \mathrm{E}^1\tau^* \leqslant \mathrm{E}^1\tau \quad \text{for any } (\tau, d) \in \Delta(\alpha, \beta),$$

where $\mathrm{E}^0 = \mathrm{E}[\,\cdot \mid \mu = \mu_0]$, $\mathrm{E}^1 = \mathrm{E}[\,\cdot \mid \mu = \mu_0]$.

In other words, we look for a decision rule with the minimal average observation time among all decision rules with given error probabilities.

Solution of the Bayesian problem

Without loss of generality, we will assume $c = 1$, $\mu_0 = 0$, $\mu_1 = m > 0$.

We have the problem

$$V(\pi) = \inf_{(\tau,d)} \mathrm{E}_\pi[a\mathbf{I}(d = 0, \mu = m) + b\mathbf{I}(d = m, \mu = 0) + \tau],$$

where E_π emphasises the prior distribution of μ (i. e. $\mathrm{P}(\mu = m) = \pi$).

Introduce the **a posteriori probability process** $\pi = (\pi_t)_{t \geqslant 0}$:

$$\pi_t = \mathrm{P}(\mu = m \mid \mathscr{F}_t^X).$$

Then we for any stopping time τ

$$\begin{aligned} \mathrm{E}_\pi \mathbf{I}(d = 0, \mu = m) &= \mathrm{E}\big[\mathrm{E}(\mathbf{I}(d = 0)\mathbf{I}(\mu = m) \mid \mathscr{F}_\tau^X)\big] \\ &= \mathrm{E}[\pi_\tau \mathbf{I}(d = 0)], \end{aligned}$$

since $\mathbf{I}(d = 0)$ is an $\mathscr{F}_\tau$-measurable function.

In the same way $\mathrm{E}_\pi \mathbf{I}(d=m, \mu=0) = \mathrm{E}[(1-\pi_\tau)\mathbf{I}(d=m)]$, so

$$V(\pi) = \inf_{(\tau,d)} \mathrm{E}_\pi[\tau + a\pi_\tau \mathbf{I}(d=0) + b(1-\pi_\tau)\mathbf{I}(d=m)].$$

Then for any decision rule (τ, d) the rule (τ, d') with

$$d' = \begin{cases} 0, & \text{if } a\pi_\tau \leqslant b(1-\pi_\tau) \;\Leftrightarrow\; \pi_\tau \leqslant b/(a+b), \\ m, & \text{if } a\pi_\tau > b(1-\pi_\tau) \;\Leftrightarrow\; \pi_\tau > b/(a+b), \end{cases}$$

will be not worse (in terms of the Bayesian risk) than (τ, d).

Thus the optimal decision rule (τ^*, d^*) should be such that τ^* solves the **optimal stopping problem**

$$V(\pi) = \inf_\tau \mathrm{E}_\pi[\tau + a\pi_\tau \wedge b(1-\pi_\tau)],$$

and d^* is given by the above formula.

Introduce the **likelihood process** $\varphi = (\varphi_t)_{t \geqslant 0}$

$$\varphi_t = \frac{d\mathrm{P}_t^1}{d\mathrm{P}_t^0}(\omega), \quad \text{where } \mathrm{P}_t^i = \mathrm{P}^i \mid \mathscr{F}_t^X.$$

It is well-known that $\varphi_t = \exp\bigl(mX_t - \frac{m^2}{2}t\bigr)$.

As follows from the **general Bayes formula** (see Liptser, Shiryaev, *Statistics of Random Processes*, ch. 7, § 9),

$$\pi_t = \pi \frac{d\mathrm{P}_t^1}{d[\pi \mathrm{P}_t^1 + (1-\pi)\mathrm{P}_t^0]},$$

and therefore

$$\pi_t = \frac{\frac{\pi}{1-\pi}\varphi_t}{1 + \frac{\pi}{1-\pi}\varphi_t}.$$

Applying the Itô formula, we obtain that π_t satisfies the SDE

$$d\pi_t = -m^2(1-\pi_t)dt + m\pi_t(1-\pi_t)dX_t, \qquad \pi_0 = \pi.$$

According to the **innovation representation**, the process

$$\widetilde{B}_t = X_t - \int_0^t \mathrm{E}_\pi[\mu \mid \mathscr{F}_s^X]ds$$

is a Brownian motion (this can be established by checking that $\widetilde{B}_t$ is a continuous square-integrable martingale such that $\mathrm{E}\widetilde{B}_t = 0$, and $\mathrm{E}(\widetilde{B}_t - \widetilde{B}_s)^2 \mid \mathscr{F}_s) = t - s$ for $t \geqslant s$).

Since $\mathrm{E}_\pi[\mu \mid \mathscr{F}_s^X] = m\pi_s$, we get that X is a **diffusion process** with the stochastic differential

$$dX_t = m\pi_t dt + d\widetilde{B}_t.$$

This representation implies that

$$d\pi_t = m\pi_t(1 - \pi_t)d\widetilde{B}_t, \qquad \pi_0 = \pi.$$

Thus, we have to solve the optimal stopping problem

$$V(\pi) = \mathrm{E}_\pi[\tau + G(\pi_\tau)], \qquad G(\pi) = a\pi \wedge b(1-\pi).$$

Due to the nature of the problem it is reasonable to assume that the continuation set is an interval

$$C = \{\pi : \pi \in (A, B)\}$$

for some $0 \leqslant A \leqslant b/(a+b) \leqslant B \leqslant 1$.

This assumptions suggests that we should look for $V(\pi), A, B$ as a solution of the **free-boundary problem**

$$\begin{cases} \mathscr{L}_\pi V = -1 \text{ for } \pi \in (A, B), \\ V = G \text{ for } \pi \notin (A, B), \\ V'(A) = a, \ V'(B) = B \end{cases}$$

where $L_\pi = \frac{m^2}{2}\pi^2(1-\pi)^2\frac{\partial^2}{\partial\pi^2}$.

Introduce the function

$$\psi(\pi) = (1 - 2\pi)\log\Big(\frac{\pi}{1-\pi}\Big).$$

We can find that the solution of the differential equation with the boundary conditions $V(A) = aA$, $V'(A) = a$ for a fixed $0 < A < b/(a+b)$ is given by

$$V(\pi; A) = \frac{2}{m^2}(\psi(\pi) - \psi(A)) + \Big(a - \frac{2}{m^2}\psi'(A)\Big)(\pi - A) + aA$$

for $\pi \geqslant A$. Choosing A and B in a such way that the conditions $V(B) = b(1-\pi)$, $V'(b) = b$ are satisfied, we obtain

$$V_*(\pi) = \begin{cases} \frac{2}{m^2}(\psi(\pi) - \psi(A)) + \Big(a - \frac{2}{m^2}\psi'(A)\Big)(\pi - A) + aA \\ \quad \text{if } \pi \in (A_*, B_*) \\ a\pi \wedge b(1-\pi) \text{ if } \pi \in [0, A_*] \cup [B_*, 1]. \end{cases}$$

where A_* and B_* form the **unique** solution of the equations

$$V(B_*; A_*) = b(1 - B_*), \qquad V'(B_*; A_*) = -b.$$

Now we need to **verify** that $V_*(\pi)$ coincides with the value function $V(\pi)$ in the optimal stopping problem.

It is clear that in the optimal stopping problem we only need to consider stopping times τ with $\mathrm{E}\tau < \infty$. Then we have

$$V(\pi) = \inf_\tau \mathrm{E}_\pi[\tau + G(\pi_\tau)] \geqslant \inf_\tau \mathrm{E}_\pi[\tau + V_*(\pi_\tau)] + \inf_\tau \mathrm{E}_\pi[G(\pi_\tau) - V_*(\pi_\tau)].$$

If $\pi \in (A_*, B_*)$, using that $\mathscr{L}_\pi V_* = -1$, for any τ, $\mathrm{E}\tau < \infty$, we get

$$E_\pi V_*(\pi_\tau) - V_*(\pi) = -\mathrm{E}_\pi \tau \iff \mathrm{E}_\pi[\tau + V_*(\pi_\tau)] = V_*(\pi).$$

Using that $G(\pi) \geqslant V_*(\pi)$ for any $\pi \in [0,1]$ we obtain

$$\inf_\tau \mathrm{E}_\pi[G(\pi_\tau) - V_*(\pi_\tau)] \geqslant 0,$$

and finally

$$V(\pi) \geqslant \inf_\tau \mathrm{E}_\pi[\tau + V_*(\pi_\tau)] \quad \Longrightarrow \quad V(\pi) \geqslant V_*(\pi).$$

By direct computations, we find that for any $\pi \in [0,1]$ the stopping time

$$\tau^* = \inf\{t \geqslant 0 : \pi_t \notin (A_*, B_*)\}$$

has the finite expectation $\mathrm{E}_\pi \tau_*$. Moreover,

$$\mathrm{E}_\pi[\tau^* + G(\pi_{\tau^*})] = \mathrm{E}_\pi[\tau^* + V_*(\pi_{\tau^*})] = V_*(\pi).$$

This implies that

$$V(\pi) \geqslant V_*(\pi) = \mathrm{E}_\pi[\tau^* + G(\pi_{\tau^*})].$$

However, since $V(\pi) \leqslant \mathrm{E}_\pi[\tau^* + G(\pi_{\tau^*})]$ by the definition of the value function, we get

$$V = V_* \text{ and } \tau^* \text{ is the optimal stopping time.}$$

Thus we have proved the following theorem.

Theorem. *The optimal decision rule in the sequential testing problem of H_0 and H_1 is (τ^*, d^*) with*

$$\tau^* = \inf\{t \geqslant 0 : \pi_t \notin (A_*, B_*)\}, \qquad d = \begin{cases} 0, & \pi_\tau \leqslant b/(a+b), \\ m, & \pi_\tau > b/(a+b), \end{cases}$$

where the constants A_*, B_* are the unique solution of the equations

$$V(B_*; A_*) = b(1 - B_*), \qquad V'(B_*; A_*) = -b,$$

for the function

$$V(\pi; A) = \frac{2}{m^2}(\psi(\pi) - \psi(A)) + \Big(a - \frac{2}{m^2}\psi'(A)\Big)(\pi - A) + aA.$$

Remark: the symmetric case

Suppose we test the hypotheses $H_+ : \mu = m$, $H_- : \mu = -m$ for $m > 0$ and $a = b > 0$, $c = 1$.

In this case, using the explicit formula for π_t through X_t we obtain the optimal decision rule (τ^*, d^*) with

$$\tau^* = \inf\{t \geqslant 0 : X_t \notin (-\widetilde{A}_*, \widetilde{A}_*)\}, \qquad d = m \operatorname{sgn}(X_{\tau^*}),$$

where

$$\widetilde{A}_* = 2m \left(\log \frac{A_*}{1 - A_*} - \log \frac{\pi}{1 - \pi} \right),$$

for the constant A_* being the unique solution of the equation

$$2am^2 = \frac{1 - A_*}{A_*} - \frac{A_*}{1 - A_*} + 2 \log \frac{1 - A_*}{A_*}.$$

(If $\widetilde{A}_* \geqslant 0$, we set $\tau^* = 0$.)

Solution of Wald's problem

We look for a decision rule $(\tau^*, d^*) \in \Delta(\alpha, \beta)$ such that

$$\mathrm{E}^1\tau^* \leqslant \mathrm{E}^0\tau, \quad \mathrm{E}^0\tau^* \leqslant \mathrm{E}^0\tau \quad \text{for any } (\tau, d) \in \Delta(\alpha, \beta).$$

assuming $\mu_1 = m > 0$, $\mu_0 = -m$, $\alpha + \beta < 1$. Recall that $(\tau, d) \in \Delta(\alpha, \beta)$ if $\mathrm{P}^1(d = -m) \leqslant \alpha$, $\mathrm{P}^0(d = m) \leqslant \beta$.

Theorem. *The optimal decision rule (τ^*, d^*) is given by*

$$\tau^* = \inf\{t \geqslant 0 : \varphi_t \notin (A_*, B_*)\}, \qquad d^* = \begin{cases} m, & \varphi_{\tau_*} \geqslant B_*, \\ -m, & \varphi_{\tau_*} \leqslant A_*, \end{cases}$$

where

$$\varphi_t = \exp\left(2mX_t - \tfrac{m^2}{2}t\right), \qquad A_* = \alpha/(1-\beta), \quad B_* = (1-\alpha)/\beta.$$

The average observation times

$$\mathrm{E}^0\tau^* = w(\beta, \alpha)/(2m^2), \qquad E^1\tau^* = w(\alpha, \beta)/(2m^2),$$

where $w(x, y) = (1-x)\log\frac{1-x}{y} + x\log\frac{x}{1-y}$.

Proof

We solve Wald's problem by reducing it to the **Bayesian problem**.

For simplicity we consider only the case $\alpha = \beta (< 1/2)$.

Consider the Bayesian problem with a parameter a:

$$V(\pi; a) = \inf_{(\tau,d)} \mathrm{E}_\pi[\tau + a\mathbf{I}(d \neq \mu)].$$

The optimal decision rule here is of the form

$$\tau^* = \inf\{t \geqslant 0 : X_t \notin (-A_*, A_*)\}, \qquad d^* = m \operatorname{sgn}(X_{\tau^*}),$$

where $A_* = \widetilde{A}_*(\pi; a)$. Moreover, for any $A_* > 0$, $\pi \in (0,1)$ it is possible to find $a > 0$ such that $\widetilde{A}_*(\pi; a) = A_*$.

Take $A_* > 0$ such that for $\tau^* = \tau^*(A_*)$, $d^* = d^*(A_*)$ we have

$$\mathrm{P}^1(d^* = -m) = \mathrm{P}^0(d^* = m) = \alpha.$$

Then for any $\pi \in (0,1)$ there exist $a = a(\pi)$ such that the optimal stopping rule $\tau^*(\pi; a) = \tau^*(A_*)$.

Using that for any decision rule (τ, d), any $\pi \in (0,1)$, $a > 0$ we have

$$\begin{aligned}
&\mathrm{E}_\pi[\tau + a\mathbf{I}(d \neq \mu)] \\
&\quad = \pi \mathrm{E}^1\tau + (1-\pi)\mathrm{E}^0\tau + a\big[\pi \mathrm{P}^1(d = -m) + (1-\pi)\mathrm{P}^0(d = m)\big],
\end{aligned}$$

we obtain that for any $(\tau, d) \in \Delta(\alpha, \beta)$ and $\pi \in (0,1)$

$$\pi \mathrm{E}^1\tau^* + (1-\pi)\mathrm{E}^0\tau^* \leqslant \pi \mathrm{E}^1\tau + (1-\pi)\mathrm{E}^0\tau,$$

where $\tau^* = \tau^*(\pi, a(\pi)) = \tau^*(A_*)$.

Since $\pi \in (0,1)$ is arbitrary, (τ^*, d^*) solves the Wald problem.

5. Sequential parameters estimation

In this part we consider the sequential estimation problem for an unknown drift coefficient of a Brownian motion.

We observe a random process $X = (X_t)_{t \geqslant 0}$

$$X_t = \mu t + B_t,$$

where μ is a **random** parameter independent of a Brownian motion B.

A **decision rule** for estimating μ is a pair (τ, d), where

- τ is a stopping time of the filtration $(\mathscr{F}_t^X)_{t \geqslant 0}$, $\mathscr{F}_t^X = \sigma(X_s; s \leqslant t)$;
- d is an $\mathscr{F}_\tau$-measurable function with values in $\mathbb{R}$.

Generally, we want τ to be small, and d to be close to μ.

Bayesian risk and reduction to an optimal stopping problem

We consider the **Bayesian risk** given by

$$\mathcal{R} = \inf_{(\tau,d)} \mathrm{E}[c\tau + W(\mu, d)],$$

where E is the expectation w.r.t the measure generated by the independent μ and B, and W is a **penalty function**; $\mathrm{E}\tau < \infty$.

Due to the representation

$$\mathrm{E}[c\tau + W(\mu, d)] = \mathrm{E}\big\{\mathrm{E}\big[c\tau + W(\mu, d) \mid \mathscr{F}_\tau^X\big]\big\}$$

and the measurability of τ and d w.r.t $\mathscr{F}_\tau^X$, we need to find

$$\mathrm{E}\big[W(\mu, d) \mid \mathscr{F}_\tau^X\big].$$

The **conditional distribution** of μ is given by

$$\mathrm{P}\big(\mu \leqslant y \mid \mathscr{F}_t^X\big) = \frac{\int\limits_{-\infty}^{y} \frac{d\mathrm{P}(X_0^t \,|\, \mu = z)}{d\mathrm{P}(X_0^t \,|\, \mu = 0)} dP_\mu(z)}{\int\limits_{-\infty}^{\infty} \frac{d\mathrm{P}(X_0^t \,|\, \mu = z)}{d\mathrm{P}(X_0^t \,|\, \mu = 0)} dP_\mu(z)},$$

with the **Radon–Nikodym** derivative

$$\frac{d\mathrm{P}(X_0^t \,|\, \mu = z)}{d\mathrm{P}(X_0^t \,|\, \mu = 0)}$$

of the measure of the process $X_0^t = (X_s, s \leqslant t)$ with the parameter $\mu = z$ w.r.t the measure of the process $X_0^t = (X_s, s \leqslant t)$ with $\mu = 0$.

Evaluating the Radon–Nikodym derivative, we obtain

$$\mathrm{P}\big(\mu \leqslant y \mid \mathscr{F}_t^X\big) = \frac{\int\limits_{-\infty}^{y} e^{zX_t - z^2 t/2} dP_\mu(z)}{\int\limits_{-\infty}^{\infty} e^{zX_t - z^2 t/2} dP_\mu(z)}.$$

If $P_\mu(z)$ has density, $dP_\mu(z) = p(z)dz$, then the **conditional density** of μ can be represented in the form

$$p(y, X_t; t) := \frac{dP(\mu \leqslant y \,|\, \mathscr{F}_t^X)}{dy} = \frac{e^{yX_t - y^2 t/2} p(y)}{\int\limits_{-\infty}^{\infty} e^{zX_t - z^2 t/2} p(z) dz}.$$

Thus, for $d = d(\tau)$ we find

$$\mathrm{E}[W(\mu, d) \mid \mathscr{F}_\tau^X] = \int_{\mathbb{R}} W(y, d(\tau)) \cdot p(y, X_\tau, \tau) dy.$$

If for any τ there exists $\mathscr{F}_\tau^X$-measurable function $d^*(\tau)$ such that

$$\inf_{d \in \mathscr{F}_\tau^X} \int_{\mathbb{R}} W(y, d) \cdot p(y, X_\tau; \tau) dy = \\ = \int_{\mathbb{R}} W(y, d^*(\tau)) \cdot p(y, X_\tau; \tau) dy \qquad (\equiv G(\tau, X_\tau)),$$

then the following equation holds (with the notation $p = \mathrm{Law}\ \mu$)

$$\inf_{(\tau,d)} \mathrm{E}[c\tau + W(\mu, d)] = \inf_\tau \mathrm{E}[c\tau + G(\tau, X_\tau)] \qquad (\equiv V(p)).$$

and if τ^* is the **optimal stopping time** in the right-hand side, then $(\tau^*, d^*(\tau^*))$ is the **optimal decision rule**.

Example 1: the mean-square criterion

$$\boxed{W(\mu, d) = (\mu - d)^2 \quad \text{and} \quad \mu \sim N(m, \sigma^2)}$$

In this case

$$V(p) = \inf_{\tau} \mathrm{E}[c\tau + v(\tau)], \qquad \text{where} \quad v(t) = 1/(t + \sigma^2).$$

The optimal time τ^* is **deterministic**:

(a) if $\sqrt{c} < \sigma^2$, then τ^* is the unique root of the equation

$$V(\tau^*) = \sqrt{c} \quad \Longleftrightarrow \quad \tau^* = c^{-1/2} - \sigma^{-2};$$

(b) if $\sqrt{c} \geqslant \sigma^2$, then $\tau^* = 0$;

The optimal d^* is the **a posteriori mean** $\mathrm{E}(\mu \,|\, \mathscr{F}^X_{\tau^*})$:

(c) $$d^* = \begin{cases} \sqrt{c} X_{\tau^*} + m\sqrt{c}/\sigma^2, & \text{if } \sqrt{c} < \sigma^2, \\ m, & \text{if } \sqrt{c} \geqslant \sigma^2. \end{cases}$$

Let us show how to obtain the representation

$$V(p) = \inf_{\tau} \mathrm{E}[c\tau + v(\tau)] \quad \text{for} \quad v(t) = 1/(t + \sigma^2).$$

Consider

$$\inf_{(\tau,d)} \mathrm{E}[c\tau + (\mu - d)^2].$$

For a given τ, the optimal $d^* = d^*(\tau)$ is the **conditional mean** of μ

$$d^*(\tau) = \mathrm{E}(\mu \,|\, \mathcal{F}_\tau^X) = \int_{\mathbb{R}} y \cdot p(y, X_\tau; \tau) dy,$$

and $\mathrm{E}[(\mu - d^*)^2 \mid \mathscr{F}_\tau^X]$ is the **conditional variance** of μ.

If $\mu \sim \mathcal{N}(m, \sigma^2)$, the **conditional variance**

$$\mathrm{E}\big(\mu - d_t^* \mid \mathscr{F}_t^X\big) = v(t),$$

where $v(t)$ solves the Ricatti equation (the **Kalman-Bucy filter**)

$$v'(t) = -v^2(t), \quad v(0) = \sigma^2.$$

Its solution is given by

$$v(t) = \frac{1}{t + \sigma^{-2}}.$$

As a result,

$$V(p) = \inf_{\tau} \mathrm{E}\left[c\tau + \frac{1}{t + \sigma^{-2}}\right],$$

which proves statements **(a)** and **(b)** for τ^*.

Representation **(c)** for $d^* = \mathrm{E}(\mu \,|\, \mathscr{F}^X_{\tau^*})$ follows from the formula

$$
\begin{aligned}
d^*_{\tau^*} &= \int_{\mathbb{R}} y p(y, X_{\tau^*}; \tau^*) dy \\
&= X_{\tau^*} v(\tau^*) + m \exp\left(- \int_0^{\tau^*} v(s) ds \right) = \\
&= X_{\tau^*} \frac{\sigma^2}{1+\sigma^2\tau^*} + \frac{m}{1+\sigma^2\tau^*},
\end{aligned}
$$

which implies

$$
d^* = \begin{cases} \sqrt{c} X_{\tau^*} + m\sqrt{c}/\sigma^2, & \text{if } \sqrt{c} < \sigma^2 \;\; (\tau^* = c^{-1/2} - \sigma^{-2}), \\ m, & \text{if } \sqrt{c} \geqslant \sigma^2 \;\; (\tau^* = 0). \end{cases}
$$

Example 2: exact estimation

Let δ_μ be the **Dirac function** and consider the penalty function

$$\boxed{W(\mu,\cdot) = -\delta_\mu(\cdot).}$$

In this case

$$\int_{\mathbb{R}} W(y,d)p(\tau,X_\tau,y)dy = -p(\tau,X_\tau,d),$$

which means that $d^*(\tau)$ is the **mode** of the conditional density $p(\tau,X_\tau,\cdot)$ (i.e. a maximum of $p(\tau,X_\tau,\cdot)$).

In the **normal case** $\mu \sim \mathcal{N}(m, \sigma^2)$ the mode coincides with the **conditional mean** (see Example 1):

$$d^*(\tau^*) = \begin{cases} \sqrt{c}X_{\tau^*} + m\sqrt{c}/\sigma^2, & \text{if } \sqrt{c} < \sigma^2 \ \ (\tau^* = c^{-1/2} - \sigma^{-2}), \\ m, & \text{if } \sqrt{c} \geqslant \sigma^2 \ \ (\tau^* = 0). \end{cases}$$

Therefore

$$G(\tau, X_\tau) = -p(\tau, X_\tau; d^*(\tau)) = -\frac{1}{v(\tau)\sqrt{2\pi}}.$$

It remains to find $\tau^* = t^*$ which minimizes $c\tau - \left(v(\tau)\sqrt{2\pi}\right)^{-1}$:

$$t^* = \begin{cases} 1/(8\pi c^2) - 1/(\sigma^2), & \text{if } 8\pi c^2 < \sigma^2, \\ 0, & \text{if } 8\pi c^2 \geqslant \sigma^2. \end{cases}$$

The corresponding function d^* is given by

$$d^* = v(\tau^*)X_{\tau^*} + m\frac{v(\tau^*)}{\sigma^2} = 8\pi c^2 X_{\tau^*} + m\frac{8\pi c^2}{\sigma^2}.$$

Remark

It would be interesting to consider problems, where μ belongs to a **finite segment** $[\mu_1, \mu_2]$, e.g. with a uniform distribution.

In this case the optimal stopping time τ^* will not be deterministic.

6. Sequential disorder detection

Generally speaking, a **moment of disorder** (another name – a **change-point**) of a stochastic process is a moment of time when its probabilistic structure changes.

We consider problems of **detecting** the disorder, when the moment of disorder is not observed directly, but shows up though changes in the behaviour of an observable process.

We will study **sequential** methods, when data arrives continuously and the aim is to stop the observation as soon as a disorder occurs, but not earlier.

The general disorder detection theory for discrete time was presented in the first part of the course. In this part we consider one particular discrete-time problem, and then study the case of continuous time.

One possible application of disorder detection methods can be found in question of **portfolio re-balancing**.

For example, if stock price is described by a geometric Brownian motion $dX_t = \mu X_t dt + \sigma X_t dB_t$, then we found earlier that the decision **when to sell the stock** depends on the ratio μ/σ^2, which, we assumed, stays constant.

But suppose that

μ/σ^2 **may change during the time segment** $[0, T]$.

When is it optimal to sell the stock in this case? — We need to develop a model of changing parameters and to find out how to detect changes.

Next we consider one discrete-time problem, and then consider the general theory for continuous time.

Random walk model of stock prices

Let $S_0, S_1, S_2, \dots$ be a random sequence representing the prices of stock at the moments of time $t = 0, 1, 2, \dots$.

Assume that the log-returns are normally distributed:

$$\log \frac{S_t}{S_{t-1}} = \mu + \sigma \xi_t \quad \Longleftrightarrow \quad S_t = S_0 \exp\Big(\mu t + \sigma \sum_{t=1}^{t} \xi_t\Big),$$

where $\xi_t \sim \mathcal{N}(0,1)$ is a sequence of independent random variables, and $\sigma > 0,\ \mu \in \mathbb{R}$ are known volatility and drift coefficients.

Suppose one wants to **sell the stock** at a time $\tau \leqslant T$ maximizing the expected utility $\mathrm{E}U(S_\tau)$. For $U_\alpha(x) = x^\alpha$ or $U_0(x) = \log(x)$ she should

- sell at $\tau = 0$ if $\mu < -\frac{\sigma^2}{2}\alpha$ (because $\mathrm{E}U_\alpha(S_u) \geqslant \mathrm{E}U_\alpha(S_t)$ for $u \leqslant t$);
- sell at $\tau = T$ if $\mu > -\frac{\sigma^2}{2}\alpha$ (here $\mathrm{E}U_\alpha(S_u) \leqslant \mathrm{E}U_\alpha(S_t)$ for $u \leqslant t$).

We will study the problem of optimal selling the stock when the parameters μ, σ **may change** before the time $t = T$.

Consider a random sequence $S_0, S_1, \ldots$ such that

$$\log \frac{S_t}{S_{t-1}} = \begin{cases} \mu_1 + \sigma_1 \xi_t, & t < \theta \\ \mu_2 + \sigma_2 \xi_t, & t \geqslant \theta \end{cases}$$

where
$\mu_1 > -\frac{\sigma_1^2}{2}\alpha$, $\mu_2 < -\frac{\sigma_2^2}{2}\alpha$ are known parameters,
$\theta \in \{1, 2, \ldots, T+1\}$ is the **moment of disorder** of the price sequence.

We assume that θ is an **unobservable** random variable independent of ξ_t with a known distribution $G(t) = \mathrm{P}(\theta \leqslant t) \;\Leftrightarrow\; p_t = \mathrm{P}(\theta = t)$.

(*Remark:* p_1 is the probability that $\mu = \mu_2$, $\sigma = \sigma_2$ from the beginning; p_{T+1} is the probability that $\mu = \mu_1$, $\sigma = \sigma_1$ until the time $t = T$.)

The question we study:

when is it optimal to sell the stock in the above model?

By definition, the moment τ when one sells the stock should be a **stopping time** of the sequence S, i. e.

$$\{\tau = t\} \in \sigma(S_u; u \leqslant t) \quad \text{for any } t = 1, 2, \ldots,$$

which means that a decision to sell the stock should be based only on the price history up to the present moment of time.

Let $\{U_\alpha(x)\}_\alpha$, $a \in (-\infty, 1]$ be the family of utility functions:

$$U_\alpha(x) = x^\alpha, \alpha \in (0,1], \quad U_0(x) = \log(x), \quad U_\alpha(x) = -x^\alpha, \alpha < 0.$$

We consider the following **optimal stopping problems** for $\alpha \leqslant 1$:

$$V^\alpha = \sup_{\tau \leqslant T} \mathrm{E} U_\alpha(S_\tau).$$

The problems consist in finding the stopping times τ^*_α at which the suprema are attained.

What can be called the solution of the problem V^{α}?

We are interested in obtaining a **Markov–type solution** of the problem: to find a sequence Z_t, such that Z_t is a function of $S_0, \ldots, S_t$ and the optimal stopping time τ^* is

$$\tau^* = \inf\{t \geqslant 0 : Z_t \in D(t)\},$$

where $D(t)$, $t = 0, \ldots, T$ are some sets in $\mathbb{R}$.

In fact, we will show that $D(t)$ are of the form

$$D(t) = \{x : x \geqslant a(t)\},$$

where $a(0), a(1), \ldots, a(T)$ define the **optimal stopping boundary**.

We provide an algorithm to find $a(t)$ numerically.

Literature review

The theory of **disorder detection** has been developed since the 1950s. Basic results was obtained by Page (1954, 1955), Roberts (1959), Shiryaev (1960, 1963) and others.

The problem we consider was proposed by Beibel & Lerche (1997) for geometric Brownian motion and later considered by Novikov & Shiryaev (2009), Ekstöm & Lindberg (2013).

They solved the problem for the **homogeneous case** – when θ is exponentially distributed on $[0, \infty)$ and only μ changes (σ remains constant).

The model we consider (θ is discrete and takes values in a finite set; both μ and σ may change) is **non-homogeneous** and more difficult.

The main result

Define

$$X_t = \log \frac{S_t}{S_{t-1}}, \qquad t = 1, \dots, T.$$

Introduce the **Shiryaev–Roberts statistic** $\psi = (\psi_t)_{t \geqslant 0}$:

$$\psi_0 = 0, \quad \psi_t = (p_t + \psi_{t-1}) \cdot \tfrac{\sigma^1}{\sigma^2} \exp\Big(\tfrac{(X_t - \mu_1)^2}{2\sigma_1^2} - \tfrac{(X_t - \mu_2)^2}{2\sigma_2^2} \Big).$$

Theorem. *The optimal stopping time in problem V^α is given by*

$$\tau_\alpha^* = \inf\{0 \leqslant t \leqslant T : \psi_t \geqslant a_\alpha(t)\},$$

where

$$a_\alpha(t) = \inf\{x \geqslant 0 : V_t^\alpha(x) = 0\}$$

for the family of functions $V_0^\alpha, V_1^\alpha, \dots, V_T^\alpha$, which are increasing, have unique positive roots, and can be found recurrently as follows:

$\dots\dots$

For $\alpha = 0$*:*

$$V_T^0(x) = 0 \textit{ for all } x \geqslant 0;$$

$$V_t^0(x) = \max\big\{0,\ \mu_1(1 - G(t+1)) + \mu_2(x + p_{t+1}) + f^0(t,x)\big\},$$

$$\textit{where} \quad f^0(t,x) = \int_{\mathbb{R}} V_{t+1}^0 \Big[\Big(p_{t+1} + x\Big) \cdot \frac{\sigma_1}{\sigma_2} \exp\Big(\tfrac{(z-\mu_1)^2}{2\sigma_1^2} - \tfrac{(z-\mu_2)^2}{2\sigma_2^2}\Big)\Big]$$

$$\times \frac{1}{\sigma_1\sqrt{2\pi}} \exp\Big(-\tfrac{(z-\mu_1)^2}{2\sigma_1^2}\Big) dz$$

For $\alpha \neq 0$*:*

$$V_T^\alpha(x) = 0 \textit{ for all } x \geqslant 0;$$

$$V_t^\alpha(x) = \max\big\{0,\ \mathrm{sgn}(\alpha) \cdot \beta^t \big[(\beta - 1)(1 - G(t+1)) + (\gamma - 1)(p_{t+1} + x)\big] + f^\alpha(t,x)\big\},$$

$$\textit{where} \quad f^\alpha(t,x) = \int_{\mathbb{R}} V_{t+1}^\alpha \Big[\Big(p_{t+1} + x\Big) \cdot \frac{\sigma_1}{\sigma_2} \exp\Big(\tfrac{(z-\mu_1)^2}{2\sigma_1^2} - \tfrac{(z-\mu_2)^2}{2\sigma_2^2}\Big)\Big]$$

$$\times \frac{1}{\sigma_1\sqrt{2\pi}} \exp\Big(-\tfrac{(z-\mu_1-\alpha\sigma_1^2)^2}{2\sigma_1^2}\Big) dz$$

$\cdots\cdots$

with the constants

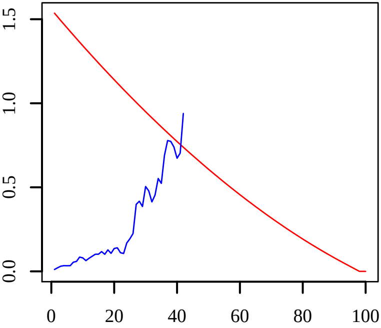
$$\beta = \exp\left(\alpha\mu_1 + \frac{\alpha^2\sigma_1^2}{2}\right), \quad \gamma = \exp\left(\frac{\alpha^2}{2}(\sigma_2^2 - \sigma_1^2) + \alpha(\mu_2 - \mu_1)\right).$$

A numerical example

Let $T = 100$, $\mu_1 = -\mu_2 = 1$, $\sigma_{1,2} = 1$ and θ be uniformly distributed.

The graphs below presents the solution of the problem V^0 when $\theta = 30$: the left graph — $\log S_t$; the right graph — $a_0(t)$ and ψ_t.

The optimal stopping time $\tau^* = 42$.



Overview of the proof

Recall, we have the sequence $S = \{S_t\}_{t \geqslant 0}$ such that

$$\log \frac{S_t}{S_{t-1}} = \begin{cases} \mu_1 + \sigma_1 \xi_t, & t < \theta, \\ \mu_2 + \sigma_2 \xi_t, & t \geqslant \theta, \end{cases} \qquad \xi_t \sim \mathcal{N}(0,1),$$

and we look for the solution of the optimal stopping problem

$$V^\alpha = \sup_{\tau \leqslant T} \mathrm{E} U_\alpha(S_\tau).$$

The problem will be solved in 2 steps:

Step 1. Reduce the problem to an optimal stopping problem without unobservable parameters.

Step 2. Prove that the solution is of the Markov type and find the stopping boundary $a(t)$ by backward induction.

Reduction to a fully observable optimal stopping problem

On the measure space $(\Omega,\ \mathscr{F}^X)$, where $\mathscr{F}^S = \sigma(X_t; t \leqslant T)$, introduce the measures P^α, $\alpha \leqslant 1$, such that

$$X_t \overset{\mathrm{P}^\alpha}{\sim} \mathcal{N}(\mu_1 + \alpha\sigma_1^2,\ \sigma_1^2).$$

The explicit formula for the density is given by the formula

$$\frac{d\mathrm{P}}{d\mathrm{P}^\alpha} = (\psi_T + p_{T+1}) \cdot \exp\Big(-\alpha \sum_{t=1}^{T} X_t + \big(\alpha\mu_1 + \tfrac{\alpha^2\sigma_1^2}{2}\big)T\Big).$$

Lemma. *For any stopping time $t \leqslant T$ it holds that*

$$\mathrm{E}U_\alpha(S_\tau) = \mathrm{E}^\alpha \left[\sum_{t=1}^{\tau} \beta^{t-1} [(\beta - \gamma)\psi_t + (\beta - 1)(1 - G(t)) \right] + 1, \ \alpha > 0$$

$$\mathrm{E}U_0(S_\tau) = \mathrm{E}^0 \left[\sum_{t=1}^{\tau} [\mu_1(1 - G(t)) + \mu_2 \psi_t] \right]$$

$$\mathrm{E}U_\alpha(S_\tau) = \mathrm{E}^\alpha \left[\sum_{t=1}^{\tau} \beta^{t-1} [(\gamma - \beta)\psi_t + (1 - \beta)(1 - G(t)) \right] - 1, \ \alpha < 0,$$

where E^α *is the expectation w.r.t.* P^α *and*

$$\beta = \exp\left(\alpha\mu_1 + \frac{\alpha^2\sigma_1^2}{2} \right), \quad \gamma = \exp\left(\frac{\alpha^2}{2}(\sigma_2^2 - \sigma_1^2) + \alpha(\mu_2 - \mu_1) \right).$$

The lemma reduces the optimal stopping problems V_α, which contain the **unobservable** random variable θ, to the optimal stopping problems for the sequence ψ_t **without unobservable** elements.

Solution of the optimal stopping problems for ψ_t

Consider the Markov setting of the optimal stopping problems V^α and introduce the value functions $V_t^\alpha(x)$ for $t = 0, \ldots, T$

$$V_t^\alpha(x) = \sup_{\tau \leqslant T-t} \mathrm{E}_{t,x}^\alpha \Big[\sum_{u=1}^{\tau} \beta^{t+u-1} [(\beta - \gamma)\psi_u + (\beta - 1)(1 - G(t+u))\Big]$$

for $\alpha > 0$,

$$V_t^0(x) = \sup_{\tau \leqslant T-t} \mathrm{E}_{t,x}^0 \Big[\sum_{u=1}^{\tau} [\mu_1(1 - G(t+u)) + \mu_2 \psi_u]\Big] \qquad \text{for } \alpha = 0,$$

$$V_t^\alpha(x) = \sup_{\tau \leqslant T-t} \mathrm{E}_{t,x}^\alpha \Big[\sum_{u=1}^{\tau} \beta^{t+u-1} [(\gamma - \beta)\psi_u + (1 - \beta)(1 - G(t+u))\Big]$$

for $\alpha < 0$,

where w.r.t. $\mathrm{E}_{t,x}^\alpha$ the sequence ψ_t satisfies the recurrent relation

$$\psi_0 = x, \qquad \psi_u = (p_{u+t} + \psi_{u-1}) \cdot \frac{\sigma_1}{\sigma_2} \exp\Big(\tfrac{(X_u - \mu_1)^2}{2\sigma_1^2} - \tfrac{(X_u - \mu_2)^2}{2\sigma_2^2} \Big)$$

with $X_1, X_2, \ldots$ being i.i.d. $\mathcal{N}(\mu_1 + \alpha\sigma_1^2, \sigma_1^2)$ random variables.

From the general optimal stopping theory, it follows that

$$\tau_\alpha^* = \inf\{0 \leqslant T : V_t^\alpha(\psi_t) = 0\},$$

where 0 is the gain from *instantaneous stopping*.

Then we show that $V_t^\alpha(x)$ increases for $x \geqslant 0$ and has a unique root $a_\alpha(t)$.

The recurrent relation for V_t^α follows from properties of conditional expectations.

Applications to stock markets

We apply the results obtained to the problem of choosing the optimal moment of time to sell stock based on real market data.

We consider the following examples:

- Apple Inc. prices in 2012;
- The NASDAQ-100 index in 1998-2004;

The model of choosing a moment to sell stock

1. We observe a sequence of stock prices (or index values) $S_0, S_1, \dots, S_T$, which, we believe, has a positive trend initially.
2. It is expected that the trend will become negative by time T.
3. For a buying time $t_0 < T$ we need to find the selling time τ maximizing the expected utility from selling.

In order to apply the disoder detection rule, we represent the prices by a random walk with a disorder:

1. We assume
$$\log \frac{S_t}{S_{t-1}} = \begin{cases} \mu_1 + \sigma_1 \xi_t, & t < \theta \\ \mu_2 + \sigma_2 \xi_t, & t \geqslant \theta, \end{cases}$$
where $\theta \in \{t_0, \dots, T\}$ is a random variable.
In the examples below, we will consider *daily* prices, so S_t and S_{t+1} correspond to two consecutive trading days.
2. The parameters μ_1, σ are estimated using the data $S_0, \dots, S_{t_0}$.

   The choice of μ_2, σ_2 and the distribution of θ is subjective. In the numerical examples below we take $\mu_2 = -\mu_1$, $\theta \sim U\{t_0, \dots, T\}$, which, as we found empirically, gives good results.
3. Then we choose the stopping time maximizing $\mathrm{E} U_\alpha(S_\tau)$. In the examples below, $U_\alpha(x) = x$.

How to estimate T?

A difficult problem is to estimate the final time T, until which, we believe, a disorder will happen.

One model for **predicting** market crashes is the bond-stock earnings yield differential (**BSEYD**) model, see Ziemba, *The stochastic programming approach to asset liability and wealth management*, 2003.

The BSEYD model relates the yield on stocks, measured by the ratio of earnings to stock prices, to the yield on nominal Treasury bonds. When the bond yield is too high, there is a shift out of stocks into bonds. If the adjustment is large, it causes an equity market correction.

The model is based on observing the earnings/price ratio (**E/P**) and the bond yield (**B**).

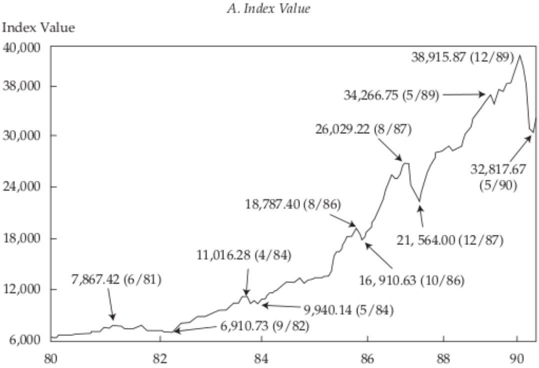
When the difference

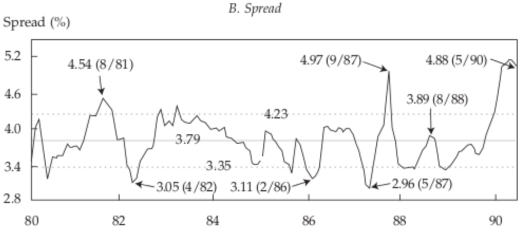
$$\mathsf{B} - \frac{\mathsf{E}}{\mathsf{P}}$$

exceeds some threshold, a decline in stock prices is expected.

As an example, let us consider the NIKKEI index in 1980-1990; the data are taken from Ziemba, 2003.

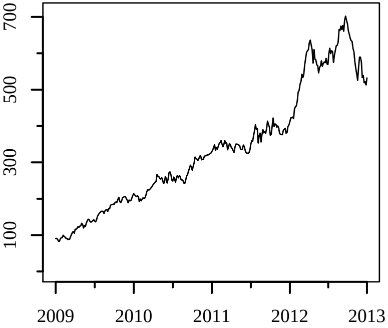
Figure 2.4. Bond–Stock Yield Differential Model for the Nikkei Stock Average, 1980–90



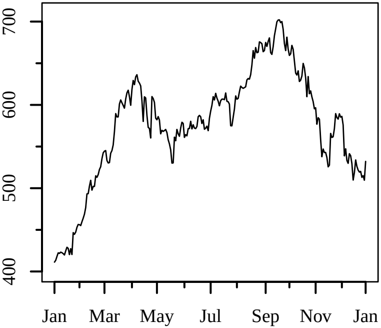


Example 1: Apple Inc

During 2009-2012, Apple's stock price increased almost 9 times, from \$82.33 (6-Mar-09), to \$705.07 (21-Sep-12). By the end of 2012 it fell to \$532.17.



Apple's stock in 2009-2012



Apple's stock in 2012 .

Choosing the optimal time to sell Apple

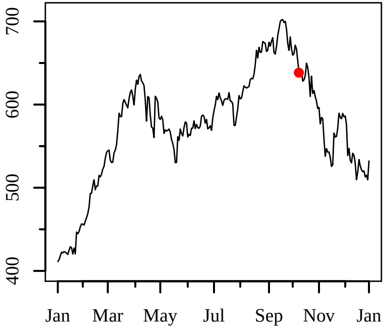
We assume $T \sim$ 31 Dec. 2012..

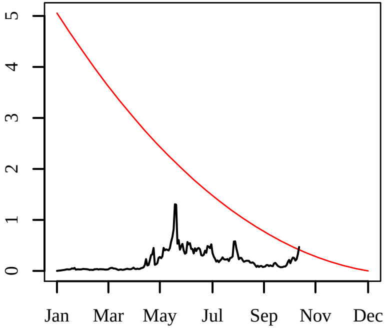
Buy	Sell	% of max.	Return*, %
3-Jan-11 ($ 329.57)	9-Oct-12 ($ 635.85)	90.56	37.13
1-Jul-11 ($ 343.26)	8-Oct-12 ($ 638.17)	90.89	48.83
3-Jan-12 ($ 411.23)	8-Oct-12 ($ 638.17)	90.89	57.38
1-May-12 ($ 582.13)	9-Oct-12 ($ 635.85)	90.56	19.86
3-Jul-12 ($ 599.41)	9-Oct-12 ($ 635.85)	90.56	21.87
1-Aug-12 ($ 606.81)	11-Oct-12 ($ 628.10)	89.46	17.38

* Return = average annual return from date n_0 to date τ^*.

On the graphs – the result of applying the method when buying on January 3, 2012.

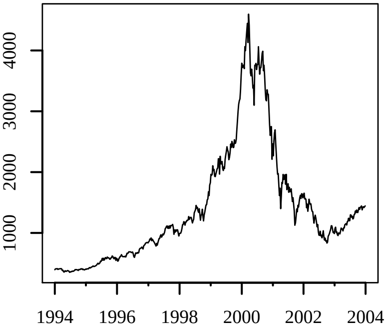
Left – the graph of the price (the red point is the selling price).
Right – the process ψ and the optimal stopping boundary.





Example 2: NASDAQ-100

From the beginning of 1994, by March 2000 the NASDAQ-100 increased more than 12 times, from 395 to 4816, and then fell to 795 by October 2002.

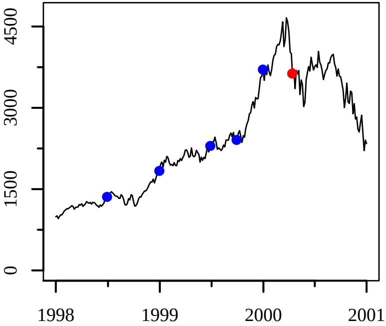


Choosing the optimal time to sell NASDAQ-100

Buy	Sell	% of max.	Return, %
2-Jul-98 ($ 1332.53)	12-Apr-00 ($ 3633.63)	77.23	56.30
4-Jan-99 ($ 1854.39)	13-Apr-00 ($ 3553.81)	75.54	50.75
1-Jul-99 ($ 2322.32)	13-Apr-00 ($ 3553.81)	75.54	53.88
1-Oct-99 ($ 2404.45)	14-Apr-00 ($ 3207.96)	68.19	53.42
3-Jan-00 ($ 3790.55)	14-Apr-00 ($ 3207.96)	68.19	-22.89

The assumption: $T \sim$ Dec. 31, 2001.

On the graph, the buying dates are marked by the **blue** points, and April 13, 2000 (one of the selling dates) is marked by the **red** point.



Concluding remarks

The solution we obtained should not be thought of as the only true rule for choosing the moment to sell stock:

- Real stock prices do not exactly follow Gaussian random walk (or geometric Brownian motion);
- It is difficult to estimate the parameters μ, σ, and the prior distribution of θ;
- There may be many disorders rather than only one.

However, the optimal criteria we find can be used as indicators of trend changes together with other known indicators.

The advantage of the result we obtain is that we develop a strict mathematical model and find the mathematically optimal criteria. We also show that these criteria are applicable to real market data.

Disorder detection for Brownian motion

Let $B = (B_t)_{t\geqslant 0}$ be a Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathrm{P})$.

Suppose we sequentially observe the process $X = (X_t)_{t\geqslant 0}$

$$X_t = \mu(t-\theta)^+ + B_t \quad \Longleftrightarrow \quad dX_t = \mu\mathbf{I}(t \geqslant \theta)dt + dB_t,$$

where $\mu \neq 0$ is a known constant, and $\theta \geqslant 0$ is an unknown moment of the appearance of a drift (a **moment of disorder**).

Each **disorder detection rule** is identified with a **stopping time** τ of the filtration $(\mathscr{F}_t^X)$ and is interpreted as the time when we raise an alarm that a disorder has occurred.

Generally, we want to find τ, which is in some sense close to θ.

Variant A

Suppose θ is a random variable with values in $[0, \infty]$ and $\mathfrak{M}$ is the class of Markov times w.r.t the filtration $(\mathscr{F}_t^X)_{t \geqslant 0}$.

- **Bayesian formulation.**
  For a given $c > 0$, to find $\tau^* \in \mathfrak{M}$ minimizing

$$\inf_{\tau} \big[\mathrm{P}(\tau < \theta) + c\mathrm{E}(\tau - \theta)^+ \big].$$

- **Conditionally variational formulation.**
  In the class $\mathfrak{M}_\alpha = \{\tau \in \mathfrak{M} : \mathrm{P}(\tau < \theta) \leqslant \alpha\}$, where $\alpha \in (0, 1)$, to find τ_α^* minimizing

$$\inf_{\tau \in \mathfrak{M}_\alpha} \mathrm{E}(\tau - \theta \mid \tau \geqslant \theta).$$

- **Absolute formulation.**
  To find $\tau^* \in \mathfrak{M}$ minimizing

$$\inf_{\tau \in \mathfrak{M}} \mathrm{E}|\tau - \theta|.$$

Variant B

Let $\mathfrak{M}_T = \{\tau \in \mathfrak{M} : \mathrm{E}^\infty \tau = T\}$ be the class of stopping times τ with the mean time $\mathrm{E}^\infty \tau$ under the assumption of no disorder, equals T.

The problem is to find $\tau_T^* \in \mathfrak{M}_T$ minimizing

$$\inf_{\tau \in \mathfrak{M}_T} \frac{1}{T} \int_0^T \mathrm{E}^\theta (\tau - \theta)^+ d\theta,$$

where E^θ is the expectation under the assumption that the disorder occurs at time θ.

We call this variant

generalized Bayesian setting

because the integration w.r.t $d\theta$ can be considered as the integration w.r.t the "generalized uniform" distribution on $\mathbb{R}_+$.

Solutions of the problems of Variant A

We will assume that θ is an exponentially distributed random variable:

$$\mathrm{P}(\theta = 0) = \pi, \qquad \mathrm{P}(\theta > t \mid \theta > 0) = e^{-\lambda t},$$

where $\pi \in [0,1)$ and $\lambda > 0$ are known.

Introduce the **a posteriori** probability process $\pi = (\pi_t)_{t \geqslant 0}$

$$\pi_t = \mathrm{P}(\theta \leqslant t \mid \mathscr{F}_t^X), \qquad \pi_0 = \pi.$$

Then we find that for any stopping time τ with $\mathrm{E}\tau < \infty$

$$\mathrm{P}(\tau < \theta) + c\mathrm{E}(\tau - \theta)^+ = \mathrm{E}_\pi \Big[1 - \pi_\tau + c \int_0^\tau \pi_t dt \Big],$$

where E_π stands for the expectation w.r.t. the distribution P_π of the process X with $\pi_0 = \pi$. The process $(\pi_t)_{t \geqslant 0}$ has the stochastic differential

$$d\pi_t = \big(\lambda - \mu^2 \pi_t^2\big)(1 - \pi_t)dt + \mu \pi_t (1 - \pi_t) dX_t.$$

First, we solve the **Bayesian formulation** of the problem.

The process $X = (X_t)_{t \geqslant 0}$ admits the **innovation representation**

$$X_t = \mu \int_0^t \pi_s ds + \widetilde{B}_t$$

where $\widetilde{B}_t = X_t - \mu \int_0^t \pi_s ds$ is a Brownian motion w.r.t $(\mathscr{F}_t^X)_{t \geqslant 0}$.

Consequently, π_t admits the stochastic differential

$$d\pi_t = \lambda(1 - \pi_t)dt + \mu\pi_t(1 - \pi_t)d\widetilde{B}_t, \qquad \pi_0 = \pi,$$

and is a Markov process.

Thus we can formulate the Markovian optimal stopping problem

$$V(\pi) = \inf_{\tau} \mathrm{E}_\pi \Big[1 - \pi_t + c \int_0^\tau \pi_s ds \Big].$$

To solve this problem, we consider the corresponding **free-boundary problem**

$$\begin{cases} V(\pi) = 1 - \pi, & \pi \geqslant A, \\ \mathscr{L}_\pi V(\pi) = -c\pi, & \pi < A, \end{cases}$$

where $\mathscr{L}_\pi$ is the infinitesimal operator of the process π_t:

$$\mathscr{L}_\pi = \lambda(1-\pi)\frac{d}{d\pi} + \frac{\mu^2}{2}\pi^2(1-\pi)^2\frac{d}{d\pi^2}$$

The general solution of the first equation contains two arbitrary constants C_1, C_2. Thus, in order to find unknown C_1, C_2, A we use the following three boundary conditions

$$\begin{cases} V(A) = 1 - A, \\ V'(A) = -1 \quad \text{(smooth fit)} \\ V'(0) = 0. \end{cases}$$

Using the above conditions we find

$$V(\pi) = \begin{cases} (1 - A_*) - \int_\pi^{A_*} y(s)dx, & \pi \in [0, A_*), \\ 1 - \pi, & \pi \in [A_*, 1], \end{cases}$$

where

$$y(x) = -\frac{2c}{\mu^2} \int_0^x \frac{e^{2\lambda(G(x)-G(u))/\mu^2}}{u(1-u)^2} du, \qquad G(u) = \log \frac{u}{1-u} - \frac{1}{u}.$$

The optimal stopping point $A_* = A_*(c)$ can be found from the equation

$$\frac{2c}{\mu^2} \int_0^{A_*} \frac{e^{2\lambda(G(A_*)-G(u))/\mu^2}}{u(1-u)^2} du = 1.$$

Then the optimal stopping time $\tau^* = \tau^*(c)$ is given by

$$\tau^* = \inf\{t \geqslant 0 : \pi_t \geqslant A_*(c)\}.$$

In the **conditionally variational formulation**

$$\inf_{\tau \in \mathfrak{M}_\alpha} \mathrm{E}_\pi(\tau - \theta \mid \tau \geqslant \theta)$$

the optimal stopping time is of the very simple structure:

$$\tau_\alpha^* = \inf\{t \geqslant 0 : \pi_t \geqslant 1 - \alpha\}.$$

Indeed, for any stopping time $\tau \not\equiv 0$ we have

$$\mathrm{E}_\pi(\tau - \theta \mid \tau \geqslant \theta) = \frac{\mathrm{E}_\pi(\tau - \theta)^+}{\mathrm{P}_\pi(\tau \geqslant \theta)}.$$

Using that $\mathrm{P}_\pi(\tau < \theta) = \mathrm{E}_\pi(1 - \pi_\tau)$ if $\pi \leqslant 1 - \alpha$, and the process π_t is continuous, we see that we must have $1 - \pi_{\tau_\alpha^*} = \alpha$, or $\pi_{\tau_\alpha^*} = 1 - \alpha$.

(Note that if $\pi \geqslant 1 - \alpha$ then $\tau_\alpha^* = 0$.)

The solution of the **absolute formulation** can be obtained from the Bayesian one (in the case θ is exponential).

Indeed, we have

$$\mathrm{E}|\tau-\theta| = \mathrm{E}[\theta-\tau+2(\tau-\theta)^+] = \frac{1}{\lambda} - \mathrm{E}\Big[\tau + 2\int_0^\tau \pi_s ds\Big],$$

where it was established above that $\mathrm{E}(\tau-\theta)^+ = \mathrm{E}\int_0^\tau \pi_s ds$.

Since $d\pi_t = \lambda(1-\pi_t)dt + \mu\pi_t(1-\pi_t)d\widetilde{B}_t$, we find

$$\mathrm{E}\pi_\tau = \lambda\Big[\int_0^\tau (1-\pi_s)ds\Big],$$

from where we get that $\mathrm{E}\tau = \mathrm{E}\pi_\tau/\lambda + \mathrm{E}\int_0^\tau \pi_s ds$, and finally

$$\mathrm{E}|\tau-\theta| = \frac{1}{\lambda}\mathrm{E}\Big[1-\pi_\tau + \int_0^\tau \pi_s ds\Big]$$

so the optimal τ^* solves the **Bayesian problem** with $c=\lambda$.

Variant B

We want to solve the following optimal stopping problem: to find

$$\mathbb{B}(T) = \inf_{\tau \in \mathfrak{M}_T} \frac{1}{T} \int_0^\infty \mathrm{E}^\theta (\tau - \theta)^+ d\theta,$$

where θ is a parameter with values in $\mathbb{R}_+$ and $\mathfrak{M}_T = \{\tau : \mathrm{E}^\infty \tau = T\}$.

The key point is the following representation:

$$\int_0^\infty \mathrm{E}^\theta (\tau - \theta)^+ d\theta = \mathrm{E}^\infty \int_0^\tau \psi_u du,$$

where $d\psi_u = du + \mu \psi_u dX_u$.

To prove this representation, we note first of all that $(\tau - \theta)^+ = \int_\theta^\infty \mathbf{I}(u \leqslant \tau) du$.

Using **change of measure**, we get

$$\mathrm{E}^{\theta}(\tau-\theta)^{+}=\int\limits_{\theta}^{\infty}\mathrm{E}^{\theta}\mathbf{I}(u\leqslant\tau)du=\int\limits_{\theta}^{\infty}\mathrm{E}^{\infty}\frac{L_u}{L_\theta}\mathbf{I}(u\leqslant\tau)du=\mathrm{E}^{\infty}\int_0^{\tau}\frac{L_u}{L_\theta}du,$$

where $L_t=d\mathrm{P}_t^0/d\mathrm{P}_t^{\infty}$, and

$$\begin{aligned}\int_0^{\infty}\mathrm{E}^{\theta}(\tau-\theta)^{+}d\theta &= \mathrm{E}^{\infty}\int_0^{\infty}\left[\int_0^{\tau}\frac{L_u}{L_\theta}du\right]d\theta\\ &= \mathrm{E}^{\infty}\int_0^{\tau}\left[\int_0^{\infty}\frac{L_u}{L_\theta}d\theta\right]du\\ &= \mathrm{E}^{\infty}\int_0^{\tau}\psi_u du.\end{aligned}$$

The process $(\psi_t)_{t\geqslant 0}$ is a P^{∞}-diffusion Markov process with the differential $d\psi_t=dt+\mu\psi_t dB_t$. We see that

$$\inf_{\tau\in\mathfrak{M}_T}\int_0^{\infty}\mathrm{E}^{\theta}(\tau-\theta)^{+}d\theta=\inf_{\tau\in\mathfrak{M}_T}\mathrm{E}^{\infty}\int_0^{\tau}\psi_u du.$$

From the general theory of optimal stopping for Markov processes it follows that an optimal stopping time in the problem

$$\tau \rightsquigarrow \inf_{\tau \in \mathfrak{M}_T} \mathrm{E}^\infty \int_0^\tau \psi_u du$$

has the following form:

$$\tau_T^* = \inf\{t \geqslant 0 : \psi_t \geqslant b(T)\},$$

where $b(T)$ is such that $\mathrm{E}^\infty \tau_T^* = T$. Since $\psi_t = t + \mu \int_0^t \psi_u dB_u$, we find that

$$\mathrm{E}^\infty \psi_{\tau_T^*} = \mathrm{E}^\infty \tau_T^*.$$

But $\psi_{\tau_T^*} = b(T)$, so that $b(T) = \mathrm{E}^\infty \tau_T^* = T$. We have got, for optimal stopping time τ_T^* in **Variant B**, the very simple formula:

$$\tau_T^* = \inf\{t \geqslant 0 : \psi_t \geqslant T\}.$$

For this stopping time τ_T^* the quantity $\mathrm{E}^\infty \int_0^{\tau_T^*} \psi_u du$ is easy to find. Indeed, consider the process $(\psi_t)_{t\geqslant 0}$ with $\psi_0 = x \geqslant 0$. The corresponding function

$$\boxed{U(x) = \mathrm{E}_x^\infty \int_0^{\tau_T^*} \psi_u du,}$$

E_x^∞ stands for averaging w.r.t. the P_x^∞-distribution of $(\psi_t)_{t\geqslant 0}$ when $\psi_0 = x$

satisfies the backward equation

$$\boxed{\mathscr{L}^\infty U(x) = -x, \text{ where } \mathscr{L}^\infty \equiv \frac{\partial}{\partial x} + \rho x^2 \frac{\partial^2}{\partial x^2} = -x, \quad \rho = \frac{\mu^2}{2}.}$$

Put for simplicity $\rho = 1$, then it is easy to find that

$$\boxed{U(x) = G\left(\frac{1}{T}\right) - G\left(\frac{1}{x}\right),}$$

where
$$\begin{aligned} G(x) &= \int_x^\infty \boldsymbol{F(u)}\, u^{-2} du, \\ \boldsymbol{F(u)} &= e^u(\boldsymbol{-\mathrm{Ei}(-u)}), \\ \boldsymbol{-\mathrm{Ei}(-u)} &\equiv \int_u^\infty \frac{e^{-t}}{t} dt. \end{aligned}$$

These formulae imply that

$$\mathbb{B}(\boldsymbol{T}) = \inf_{\tau\in\mathfrak{M}_T} \frac{1}{T}\int_0^\infty \mathrm{E}^\theta(\tau-\theta)^+ d\theta = \inf_{\tau\in\mathfrak{M}_T} \frac{1}{T}\mathrm{E}^\infty \int_0^\tau \psi_u du,$$

$$= \frac{1}{T}\mathrm{E}^\infty \int_0^{\tau_T^*} \psi_u du = \frac{1}{T}U(0) = \frac{1}{T}G\left(\frac{1}{T}\right) =$$

$$\boldsymbol{= F\left(\frac{1}{T}\right) - \Delta\left(\frac{1}{T}\right)}, \text{ where } \Delta(b) = 1 - b\int_0^\infty e^{-bu}\frac{\log(1+u)}{u}du.$$

Thus, $\mathbb{B}(T) = \frac{1}{T}G\left(\frac{1}{T}\right) = F\left(\frac{1}{T}\right) - \Delta\left(\frac{1}{T}\right)$ and we have the following asymptotics for small and large T:

$$\boxed{\mathbb{B}(T) = \begin{cases} \frac{T}{2} + O(T^2), & T \to 0, \\ \log T - (1+\mathrm{C}) + O(T^{-1}\log^2 T), & T \to \infty, \end{cases}}$$

where $\mathrm{C} = 0.577\ldots$ is the Euler constant.

7. Maximal inequalities

Let $B = (B_t)_{t \geqslant 0}$ be a standard Brownian motion defined on a probability space $(\Omega, \mathscr{F}, \mathrm{P})$.

Our aim is to prove the following **maximal inequalities** for B.

Theorem. *For any stopping time τ of the filtration $(\mathscr{F}_t^B)_{t \geqslant 0}$ the following inequalities hold:*

$$\mathrm{E} \max_{s \leqslant \tau} B_s \leqslant \sqrt{\mathrm{E}\tau}, \tag{1}$$

$$\mathrm{E} \max_{s \leqslant \tau} |B_s| \leqslant \sqrt{2\mathrm{E}\tau}, \tag{2}$$

$$\mathrm{E}\Big[\max_{s \leqslant \tau} B_s - \min_{s \leqslant \tau} B_s\Big] \leqslant \sqrt{3\mathrm{E}\tau}. \tag{3}$$

---

(1), (2): Dubins, Shepp, Shiryaev, *Theory Probab. Appl. 38:2* (1993).

(3): Dubins, Gilat, Meilijson, *Ann. Prob. 37:1* (2009); Zhitlukhin, *Statistics & Decision 27* (2009).

Remark 1. We also show that these inequalities are **strict** in the following sense:
For any $T > 0$ *there exist stopping times* τ_1, τ_2, τ_3 *with* $\mathrm{E}\tau_i = T$ *such that inequalities* (1), (2), (3) *turn into equalities for* τ_1, τ_2, τ_3 *respectively.*

Remark 2. For any **non-random time** $t \geqslant 0$ we have

$$\mathrm{E}\max_{s\leqslant t} B_s = \sqrt{\frac{2}{\pi}t},$$

$$\mathrm{E}\max_{s\leqslant t} |B_s| = \sqrt{\frac{\pi}{2}t}.$$

The maximal inequality for $\max B$

Let us consider the following *auxiliary* optimal stopping problem with the parameter $c \geqslant 0$:

$$V_c^1 = \sup_{\tau \geqslant 0} \mathrm{E}\big[\max_{s \leqslant \tau} B_s - c\tau\big].$$

We will use the solutions of this problem to prove the maximal inequality for $\max B$. We will find $V_c^1 = 1/(4c)$, so for any τ

$$\mathrm{E}\max_{s \leqslant \tau} B_s \leqslant \inf_{c \geqslant 0}[V_c^1 + c\mathrm{E}\tau] = \inf_{c \geqslant 0}[1/(4c) + c\mathrm{E}\tau] = \sqrt{\mathrm{E}\tau}.$$

Moreover, for any $T > 0$ and $c = 1/(2\sqrt{T})$ the optimal $\tau_1^*(c)$ is such that $\mathrm{E}\tau_1^*(c) = T$ and $\mathrm{E}\max\limits_{s \leqslant \tau_1^*(c)} B_s = \sqrt{T}$, i. e. inequality (1) is strict.

Solution of problem V_c^1

$$\boxed{V_c^1 = \sup_{\tau \geqslant 0} \mathrm{E}\big[\max_{s \leqslant \tau} B_s - c\tau\big]}$$

It is possible to apply the general theory of optimal stopping to this problem, but we prefer to give a simpler, but "tricky" solution.

Obviously, we can consider only stopping times τ with $\mathrm{E}\tau < \infty$. For any such stopping time we have $\mathrm{E}B_\tau = 0$, so

$$\mathrm{E}\big[\max_{s \leqslant \tau} B_s - c\tau\big] = \mathrm{E}\big[\max_{s \leqslant \tau} B_s - B_\tau - c\tau\big]$$

By **Lévy's theorem**, $\mathrm{Law}(\max B - B) = \mathrm{Law}(|W|)$, where W is a Brownian motion. So we have

$$\mathrm{E}\big[|W_\tau| - c\tau\big] = \mathrm{E}\big[|W_\tau| - cW_\tau^2\big],$$

where we used the **Wald identity**: $\mathrm{E}W_\tau^2 = \mathrm{E}\tau$.

It is easy to check that

$$|x| - cx^2 \leqslant \frac{1}{4c} \text{ for any } x \geqslant 0, \qquad |x| - cx^2 = \frac{1}{4c} \text{ for } x = \frac{1}{2c}.$$

Then the optimal stopping time for V_c^1 is given by

$$\tau_c^{1*} = \inf\{t \geqslant 0 : |W_t| = 1/(2c)\}$$

or, recalling that $|W_t| = \max\limits_{s \leqslant t} B_s - B_t$,

$$\boxed{\tau_c^{1*} = \inf\{t \geqslant 0 : \max_{s \leqslant t} B_s - B_t = 1/(2c)\}}$$

and we have

$$\boxed{V_c^1 = \frac{1}{4c}, \qquad \mathrm{E}\tau_c^{1*} = 1/(4c^2).}$$

Moreover, it is possible to extend the maximal inequality for $\max B$ to any **continuous local martingale** $M = (M_t)_{t\geqslant 0}$ with $M_0 = 0$. By changing the time we get

$$\mathrm{E}\max_{s\leqslant\tau} M_s = \mathrm{E}\max_{s\leqslant\tau} B_{\langle M\rangle_s} = \mathrm{E}\max_{s\leqslant\langle M\rangle_\tau} B_s \leqslant \sqrt{\mathrm{E}\langle M\rangle_\tau}.$$

The maximal inequality for $\max|B|$

We have

$$
\begin{aligned}
\mathrm{E}\max_{t\leqslant\tau}|B_t| &= \mathrm{E}\max_{t\geqslant 0}|B_{t\wedge\tau}| \leqslant \mathrm{E}\Big\{\max_{t\geqslant 0}\mathrm{E}\big[|B_\tau|\,\big|\,\mathscr{F}_{t\wedge\tau}\big]\Big\} \\
&= \mathrm{E}\Big\{\max_{t\geqslant 0}\mathrm{E}\big[|B_\tau|-\mathrm{E}|B_\tau|\,\big|\,\mathscr{F}_{\tau\wedge t}\big]\Big\}+\mathrm{E}|B_\tau| \\
&\overset{(*)}{\leqslant} \sqrt{\mathrm{E}(|B_\tau|-\mathrm{E}|B_\tau|)^2}+\mathrm{E}|B_\tau| \\
&\leqslant \sqrt{\mathrm{E}\tau-(\mathrm{E}|B_\tau|)^2}+\mathrm{E}|B_\tau| \overset{(**)}{\leqslant} \sqrt{2\mathrm{E}_\tau}
\end{aligned}
$$

Where in $(*)$ we used that $\mathrm{E}[|B_\tau|-\mathrm{E}|B_\tau| \mid \mathscr{F}_{t\wedge\tau}]$ is a continuous martingale, and in $(**)$ we used the inequality $\sqrt{A-x^2}+x\leqslant\sqrt{2A}$ valid for any $0\leqslant x\leqslant\sqrt{A}$.

The maximal inequality is attained at the stopping times

$$\tau_2^* = \inf\Big\{t\geqslant 0 : \max_{s\leqslant t}|B_s|-|B_t|\geqslant a\Big\} \qquad \text{for any } a>0.$$

Indeed, for any fixed $a > 0$ denote $\sigma_a = \inf\{t \geqslant 0 : |B_t| = a\}$. Then

$$\tau_2^* = \sigma_a + \inf\Big\{t \geqslant \sigma_a : \max_{s \leqslant t} |B_s| - |B_t| \geqslant a\Big\} - \sigma_a.$$

From the strong Markov property,

$$\inf\Big\{t \geqslant \sigma_a : \max_{s \leqslant t} |B_s| - |B_t| \geqslant a\Big\} - \sigma_a \overset{Law}{=} \inf\Big\{t \geqslant 0 : \max_{s \leqslant t} B_s - B_t \geqslant a\Big\}.$$

Thus

$$\mathrm{E}\tau_2^* = 2a^2, \qquad \mathrm{E}\max_{s \leqslant \tau_2^*} |B_s| = 2a,$$

which gives

$$\mathrm{E}\max_{s \leqslant \tau_2^*} |B_s| = \sqrt{2\mathrm{E}\tau_2^*}.$$

Remark. Another approach to prove this maximal inequality is to solve the optimal stopping problem

$$V_c^2 = \sup_{\tau \geqslant 0} \mathrm{E}\Big[\max_{s \leqslant \tau} |B_s| - c\tau\Big].$$

Its solution is given by

$$\tau_c^{2*} = \inf\{t \geqslant 0 : \max_{s \leqslant t} |B_s| - |B_t| \geqslant 1/(2c)\}, \qquad V_c^2 = \frac{1}{2c}.$$

The maximal inequality for $\max B - \min B$

The inequality is proved by solving the optimal stopping problem

$$V_c^3 = \sup_{\tau \geqslant 0} \mathrm{E}\Big[\max_{s \leqslant \tau} B_s - \min_{s \leqslant \tau} B_s - c\tau\Big].$$

We provide only the answer:

$$\tau_c^{3*} = \inf\{t \geqslant 0 : (\max_{s \leqslant t} B_s - B_t) \wedge (B_t - \min_{s \leqslant t} B_s) \geqslant 1/(2c)\}$$

$$V_c^3 = \frac{3}{4c}$$

Thus one needs to stop when B_t deviates more than $1/(2c)$ from both its current maximum and minimum.

The proof can be carried in the same way as for V_c^2, but is more complicated. Another proof, based on the **martingale approach** can be found in [Dubins, Gilat, Meilijson, *Ann. Probab. 37:1* (2009)].