Albert N. Shiryaev
(Steklov Mathematical Institute
and Moscow State University)

ESSENTIALS of the ARBITRAGE THEORY

Part I. Basic notions and theorems of the "Arbitrage Theory"
Part II. Martingale measures and their constructions
Part III. Quickest detection of the appearing of the arbitrage possibilities

Appendix I. A. N. Shiryaev, A. S. Cherny. "Vector Stochastic Integrals and the Fundamental Theory of Asset Pricing".
Appendix II. J. Kallsen, A. N. Shiryaev. "The Cumulant Process and Esscher's Change of Measure".

Lectures
in IPAM (Institute for Pure
and Applied Mathematics),
UCLA; 3–4–5 January 2001

Albert N. Shiryaev
(lectures in UCLA)

Part III

QUICKEST DETECTION OF APPEARING OF THE ARBITRAGE POSSIBILITIES

1 Setting of the Problem

Suppose that we are observing a random process $X = (X_t)$ on an interval $[0, T]$. The objects θ and τ introduced below are essential throughout the paper:

θ – a *parameter* or a *random variable*; this is the time at which the observed process $X = (X_t)_{t \geq 0}$ changes its probability *characteristics*;

τ – a *stopping* (*Markov*) *time* which serves as the time of "alarm"; it warns of the coming of the time θ.

In connection with the technical analysis of the financial data, it is of interest to consider the schemes in which θ is interpreted as the time of the appearance of an *arbitrage* ("transition from a martingale to a submartingale", for example) or as the time of the appearance of a *change-point*.

We will concentrate our attention on the model in which the observed process X has the following form (see Figure 1):

$$X_t = r(t - \theta)^+ + \sigma B_t,$$

i.e.

$$dX_t = \begin{cases} \sigma\, dB_t, & t < \theta, \\ r\, dt + \sigma\, dB_t, & t \geq \theta. \end{cases}$$

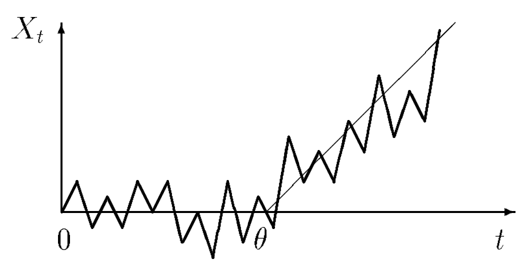


Figure 1

One can associate the following two events with a time τ (recall that τ is interpreted as the time of "alarm"):

$$\{\tau < \theta\} \text{ and } \{\tau \geq \theta\}.$$

The first event ($\{\tau < \theta\}$) corresponds to the "false alarm": the alarm time τ comes before the time θ. The second event ($\{\tau \geq \theta\}$) corresponds to the case where the alarm is raised in due time, i.e. after the time θ.

The above reasoning leads to several formulations of the quickest detection problem for the time θ. These formulations are given below.

Suppose that $\theta = \theta(\omega)$ is a *random variable* ($\theta \geq 0$). Then the first formulation (Variant **A**) of the *quickest detection problem* for θ is as follows. (This formulation can be called *conditionally-extremal*).

Variant A. For a given $\alpha \in (0,1)$, find a stopping time τ^*_α such that

$$\mathbb{A}(\alpha) = \inf_{\tau \in \mathfrak{M}_\alpha} \mathsf{E}(\tau - \theta \,|\, \tau \geq \theta) = \mathsf{E}(\tau^*_\alpha - \theta \,|\, \tau^*_\alpha \geq \theta),$$

where

$$\mathfrak{M}_\alpha = \{\tau : \mathsf{P}(\tau < \theta) \leq \alpha\}.$$

Variant B. For a given $c > 0$, find

$$\mathbb{B}(c) = \inf_\tau \big\{\mathsf{P}(\tau < \theta) + c\,\mathsf{E}(\tau - \theta)^+\big\}$$

and the corresponding optimal (Bayes) stopping time. (Note that

$$\mathsf{E}(\tau - \theta)^+ = \mathsf{E}(\tau - \theta \,|\, \tau \geq \theta)\,\mathsf{P}(\tau \geq \theta)\ ).$$

In the following Variants **C** and **D**, θ is an *unknown parameter.*

Variant C. Find

$$\mathbb{C}(T) = \inf_{\tau \in \mathfrak{M}^T} \sup_\theta \operatorname*{ess\,sup}_\omega \mathsf{E}_\theta\big((\tau - \theta)^+ \,|\, \mathcal{F}_\theta\big)(\omega)$$

and the corresponding optimal stopping time. Here,

$$\mathfrak{M}^T = \{\tau : \mathsf{E}_\infty \tau = T\},$$
$$\mathsf{P}_\theta(\,\cdot\,) = \mathrm{Law}(\,\cdot\,|\theta).$$

Variant D. Find

$$\mathbb{D}(T) = \inf_{\tau \in \mathfrak{M}^T} \sup_\theta \mathsf{E}_\theta(\tau - \theta \,|\, \tau \geq \theta)$$

and the corresponding optimal stopping time.

The following Variant **E** of the quickest detection problem is interesting due to the unusual assumptions made on the nature of θ and on the character of the observation procedure.

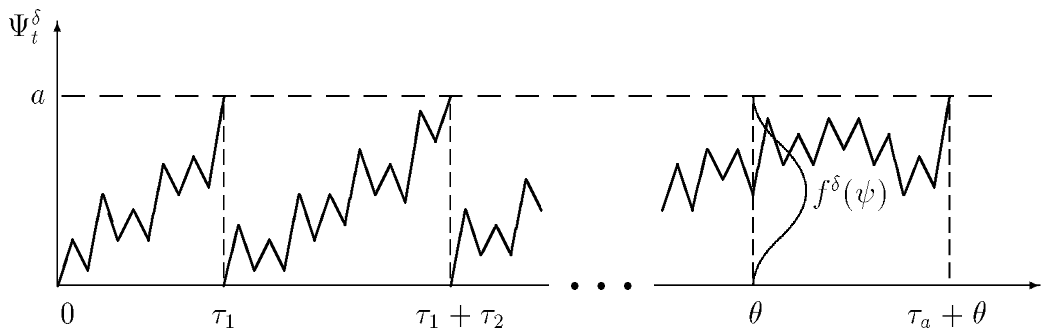
First, the assumption that θ is a random variable (or an unknown parameter) is replaced by the assumption that θ appears after the *stationary regime of the observations* is established. (Of course, the time θ is preceded by a long period of observations that contains many alarms of the appearance of "change-points"). Second, we suppose that the observation procedure in Variant **E** is *multistage.*

Informally, Variant **E** (to be more precise, its particular case; a more general formulation, Variant **E**′, is given in Subsection 9 of Section 3 below) is formulated as follows.

Let $\Psi_t = \Psi(t; X_s,\, s \le t)$ be a functional of the observations with $\Psi_0 = 0,\ t \ge 0$. Suppose that the alarm of the appearance of a "change-point" is based on the observation of Ψ and the alarm procedure has the following form.

We observe the process $(X_t)_{t\ge 0}$ and use it to construct a process $\Psi = (\Psi_t)_{t\ge 0}$. Once this process has reached a level $a > 0$, we raise an alarm of the appearance of a "change-point". Let us call this time τ_1. After this time, the process Ψ is returned to zero, i.e. for $t > \tau_1$, we observe the process $\Psi(t-\tau_1; X_s - X_{\tau_1},\, \tau_1 < s \le t)$. The next alarm is raised at a time $\tau_1 + \tau_2$. In a similar way, this procedure (let us call it δ) is repeated after the time $\tau_1 + \tau_2$ with the alarm raised at a time $\tau_1 + \tau_2 + \tau_3$ and so on.

Let us denote the process constructed above (which is a "renewal process") by $\Psi^\delta = (\Psi^\delta_t)_{t\ge 0}$ (see Figure 2). We will suppose that this process has a limit distribution $F^\delta(\psi) = \lim_{t\to\infty} \mathsf{P}_\infty(\Psi^\delta_t \le \psi)$, where P_∞ denotes the distribution of the process X under the assumption that there is no change-point (i.e. for the case where $\theta = \infty$).



`Figure 2.` (The function $f^\delta = f^\delta(\psi)$ is the density of the stationary distribution $F^\delta = F^\delta(\psi)$)

Let the mean time between two false alarms for the above observation procedure (it is determined by a functional Ψ and a level $a > 0$) be equal to T, i.e. $\mathsf{E}_\infty \tau_i = T$ $i = 1, 2, \ldots$. Here, E_∞ is the expectation taken with respect to the measure P_∞. Then the mean time of the delay after the appearance of θ is given by

$$R^\delta(T) = \int_0^a (\mathsf{E}^\psi_0 \tau_a)\, F^\delta(d\psi),$$

where $\mathsf{E}^\psi_0 \tau_a$ is the expectation of the hitting time of the level a by the process Ψ^δ under the assumption that $\Psi^\delta_0 = \psi$. Here, E_0 is the expectation with respect to the measure P_0.

The formulation of the quickest detection problem for the multistage observations and the stationary regime is as follows:

Variant E. Find among all the methods determined by a pair (Ψ, a) the infimum

$$\mathbb{E}(T) = \inf_{\{\Psi, a\}} R^\delta(T)$$

assuming that the mean time of delay between two false alarms is equal to $T > 0$.

Our aim is to describe the optimal or asymptotically optimal methods for Variants **A**, **B**, **C**, **D**, **E**.

2 Solution of the Problem

1. Suppose that we observe a continuous-time random process $X = (X_t)_{t\geq 0}$ defined as

$$X_t = r(t-\theta)^+ + \sigma B_t, \tag{1}$$

i.e.

$$dX_t = \begin{cases} \sigma\, dB_t, & t < \theta, \\ r\, dt + \sigma\, dB_t, & t \geq \theta, \end{cases}$$

where θ is the time of the appearance of the *arbitrage possibility* and $B = (B_t)_{t\geq 0}$ is a standard Brownian motion.

We will need the following notations:

$$\mathsf{P}_\theta = \mathrm{Law}(X|\theta),$$
$$L_t = \frac{d\mathsf{P}_0}{d\mathsf{P}_\infty}(t, X),$$

where

$$\frac{d\mathsf{P}_\theta}{d\mathsf{P}_\infty}(t, X) = \frac{d(\mathsf{P}_\theta \mid \mathcal{F}_t^X)}{d(\mathsf{P}_\infty \mid \mathcal{F}_t^X)}, \quad \theta \in [0, \infty].$$

For our model, we have

$$L_t = \frac{d\mathsf{P}_0}{d\mathsf{P}_\infty}(t, X) = \frac{d\mathsf{P}_0}{d\mathsf{P}_t}(t, X),$$
$$\frac{d\mathsf{P}_\theta}{d\mathsf{P}_\infty}(t, X) = \frac{d\mathsf{P}_\theta}{d\mathsf{P}_t}(t, X) = \frac{L_t}{L_\theta}, \quad \theta \leq t.$$

The following two statistics $\gamma = (\gamma_t)_{t\geq 0}$ and $\psi = (\psi_t)_{t\geq 0}$, defined as

$$\gamma_t = \max_{\theta \leq t} \frac{L_t}{L_\theta}$$

and

$$\psi_t = \int_0^t \frac{L_t}{L_\theta}\, d\theta$$

are essential in all the subsequent considerations.

We call $\gamma = (\gamma_t)_{t\geq 0}$ the *exponential CUSUM* (cumulative sum) process or the exponential *CUSUM-statistics*. For model (1), we have

$$L_t = e^{H_t}, \quad H_t = \frac{r}{\sigma^2} X_t - \frac{r^2}{2\sigma^2} t$$

and

$$\gamma_t = \exp\Big\{ H_t - \min_{\theta \leq t} H_\theta \Big\}. \tag{2}$$

In the statistical literature (see, for example, [8], [16]), the statistics $\psi = (\psi_t)_{t\geq 0}$ is called the *Shiryaev-Roberts statistics.*

For model (1), we have

$$\psi_t = \int_0^t e^{H_t - H_\theta} d\theta.$$

`Remark 1.` Let us consider the discrete-time case. We assume that the disorder (the change-point) can appear at the times $\theta = 0, 1 \dots$. If $\theta = 0$, then the observed sequence $x_0, x_1, \dots$ is a sequence of independent identically distributed random variables with the distribution density $f_1(x)$. If $\theta \geq 1$, then the observed sequence $x_0, \dots, x_{\theta-1}, x_\theta, x_{\theta+1}, \dots$ is again a sequence of independent random variables such that $x_0, \dots, x_{\theta-1}$ have the distribution density $f_0(x)$ and $x_\theta, \dots, x_{\theta+1}$ have the distribution density $f_1(x)$.

In this case,

$$L_n = \prod_{i=0}^n \frac{f_1(x)}{f_0(x)} = \exp\left\{ \sum_{i=0}^n \log \frac{f_1(x_i)}{f_0(x_i)} \right\}.$$

Let

$$S_n = \sum_{i=0}^n \log \frac{f_1(x_i)}{f_0(x_i)}, \quad n \geq 0.$$

Then $L_n = e^{S_n}$ and (compare with (2))

$$\gamma_n = \max_{0\leq\theta\leq n} \frac{L_n}{L_\theta} = \max_{0\leq\theta\leq n} \exp\{S_n - S_\theta\} = \exp\left\{ S_n - \min_{0\leq\theta\leq n} S_\theta \right\}. \tag{3}$$

Set

$$\widetilde{S}_n = S_n - \min_{0\leq\theta\leq n} S_\theta.$$

Obviously, $\widetilde{S}_0 = 0$ and, for $n \geq 1$, we have

$$\widetilde{S}_n = \max\left\{ 0, \widetilde{S}_{n-1} + \log \frac{f_1(x_n)}{f_0(x_n)} \right\}.$$

In the papers [6], [7], this recurrent relation was used to define the *cumulative sum* (CUSUM) process $\widetilde{S} = (\widetilde{S}_n)_{n\geq 0}$. This, together with representation (2), explains the above-mentioned name "exponential CUSUM-process" for the statistics $\gamma = (\gamma_t)_{t\geq 0}$. Here, γ_t has the form $\gamma_t = e^{\widetilde{H}_t}$ with $\widetilde{H}_t = H_t - \min_{0\leq\theta\leq t} H_\theta$ (compare with (3)).

The following statistics $\psi = (\psi_n)_{n\in\mathbb{N}}$ serves as a discrete-time analog of the statistics $\psi = (\psi_t)_{t\geq 0}$:

$$\psi_n = \sum_{\theta=1}^n \frac{L_n}{L_{\theta-1}}.$$

The statistics ψ defined by the above equality satisfies the recurrent relation

$$\psi_n = (1 + \psi_{n-1}) \frac{f_1(x_n)}{f_0(x_n)}, \quad n \geq 1$$

with $\psi_0 = 0$. □

`Remark 2.` Instead of the process $H_t - \min_{\theta\leq t} H_\theta$ that appears in (2), one can consider the process $\max_{\theta\leq t} H_\theta - H_t$.

If $H_t = B_t$, then, by P. Lévy's theorem,

$$\mathrm{Law}\Big(\max_{\theta\le t} B_\theta - B_t;\ t\ge 0\Big) = \mathrm{Law}(|B_t|;\ t\ge 0).$$

Due to this equality, the process $\max B - B$ is often called the *reflected* Brownian motion.

If $H_t = B_t^\mu$ with $B_t^\mu = \mu t + B_t$, then (see [2])

$$\mathrm{Law}\Big(\max_{\theta\le t} B_\theta^\mu - B_\theta^\mu;\ t\ge 0\Big) = \mathrm{Law}\big(|Y_t^\mu|;\ t\ge 0\big),$$

where $Y^\mu = (Y_t^\mu)_{t\ge 0}$ is the so-called "bang-bang process" defined as the solution of the stochastic differential equation

$$dY_t^\mu = -\mu\,\mathrm{sgn}\,Y_t^\mu\,dt + dB_t, \quad Y_0^\mu = 0.$$

□

2. We will now turn to the quickest detection problem in Variants **A** and **B**. First of all, we will make an assumption on the distribution of the random variable $\theta = \theta(\omega)$ (recall that θ corresponds to the "change-point" in the observed process (1)).

Let us assume that θ takes values in $[0,\infty)$ and it has the following distribution:

$$\mathsf{P}(\theta = 0) = \pi,$$

where $\pi \in [0,1]$, and

$$\mathsf{P}(\theta \ge t \,|\, \theta > 0) = e^{-\lambda t},$$

where $\lambda > 0$ is a known constant.

Let

$$\pi_t = \mathsf{P}(\theta \le t \,|\, \mathcal{F}_t^X)$$

be the posterior probability (constructed through the observation of the process X) that the change-point has appeared within the time-interval $[0,t]$ (here, $\mathcal{F}_t^X = \sigma(X_s;\ s\le t)$).

Set

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

Applying the Bayes formula, we get

$$\begin{aligned}\varphi_t = \frac{\mathsf{P}(\theta\le t\,|\,\mathcal{F}_t^X)}{\mathsf{P}(\theta> t\,|\,\mathcal{F}_t^X)} &= \frac{\pi}{1-\pi}\, e^{\lambda t}\,\frac{d\mathsf{P}_0}{d\mathsf{P}_\infty}(t,X) + e^{\lambda t}\int_0^t \frac{d\mathsf{P}_\theta}{d\mathsf{P}_\infty}(t,X)\,\lambda e^{-\lambda\theta}\,d\theta \\ &= \frac{\pi}{1-\pi}\, e^{\lambda t} L_t + \lambda e^{\lambda t}\int_0^t \frac{L_t}{L_\theta}\, e^{-\lambda t} d\theta.\end{aligned} \tag{4}$$

Set

$$U_t = e^{\lambda t} L_t\,\frac{\pi}{1-\pi},$$

$$V_t = \lambda e^{\lambda t}\int_0^t \frac{L_t}{L_\theta}\, e^{-\lambda\theta} d\theta.$$

Applying Itô's formula and taking the equality $dL_t = \frac{r}{\sigma^2} L_t dX_t$ into account, we get

$$dU_t = \lambda\, U_t\, dt + \frac{r}{\sigma^2}\, U_t\, dX_t, \quad U_0 = \frac{\pi}{1-\pi},$$
$$dV_t = \lambda(1+V_t)\, dt + \frac{r}{\sigma^2}\, V_t\, dX_t, \quad V_0 = 0.$$

Therefore, the process $\varphi_t = U_t + V_t$ satisfies the equation (see [11], [12])

$$d\varphi_t = \lambda(1+\varphi_t)\, dt + \frac{r}{\sigma^2}\, \varphi_t\, dX_t, \quad \varphi_0 = \frac{\pi}{1-\pi}.$$

Applying once again Itô's formula, we conclude that the posterior probability $(\pi_t)_{t\geq 0}$ satisfies the following stochastic differential equation (see [11], [12]):

$$d\pi_t = \Big(\lambda - \frac{r^2}{\sigma^2}\pi_t^2\Big)\,(1-\pi_t)\, dt + \frac{r}{\sigma^2}\, \pi_t(1-\pi_t)\, dX_t, \quad \pi_0 = \pi. \tag{5}$$

Let us denote φ_t by $\varphi_t(\lambda)$ in order to emphasize the dependence of φ_t on λ. Applying (4) with $\pi = 0$, we get

$$\frac{\varphi_t(\lambda)}{\lambda} = e^{\lambda t} \int_0^t \frac{L_t}{L_\theta}\, e^{-\lambda\theta}\, d\theta.$$

Consequently,

$$\lim_{\lambda\to 0} \frac{\varphi_t(\lambda)}{\lambda} = \psi_t \bigg(= \int_0^t \frac{L_t}{L_\theta}\, d\theta \bigg), \tag{6}$$

and the statistics ψ satisfies the following stochastic differential equation (it was obtained by the author in [11], [12]):

$$d\psi_t = dt + \frac{r}{\sigma^2}\, \psi_t\, dX_t, \quad \psi_0 = 0. \tag{7}$$

3. We will first solve the quickest detection problem in the Bayes formulation (**Variant B**). Set

$$\mathbb{B}(c;\pi) = \inf_\tau \big\{ \mathsf{P}_\pi(\tau < \theta) + c\mathsf{E}_\pi(\tau-\theta)^+ \big\}, \tag{8}$$

where the subscript π indicates that the prior probability of the event $\{\theta = 0\}$ equals π.

Using (8), we get

$$\mathbb{B}(c;\pi) = \inf_\tau \mathsf{E}_\pi \bigg\{ (1-\pi_\tau) + c\int_0^\tau \pi_s\, ds \bigg\} \ \big(= \rho_*(\pi)\big). \tag{9}$$

It is useful to introduce the *innovation* representation of the process X:

$$dX_t = r\pi_t\, dt + \sigma\, d\overline{B}_t,$$

where $\overline{B} = (\overline{B}_t)_{t\geq 0}$ is again a Brownian motion (with respect to the filtration $(\mathcal{F}_t^X)_{t\geq 0}$). See [3] for details.

In view of this representation, stochastic differential equation (5) takes the following form:

$$d\pi_t = \lambda(1-\pi_t)\, dt + \frac{r}{\sigma}\, \pi_t(1-\pi_t)\, d\overline{B}_t. \tag{10}$$

Consequently, the infinitesimal generator of the diffusion process $(\pi_t)_{t\geq 0}$ has the form

$$L = a(\pi)\,\frac{d}{d\pi} + \frac{1}{2}\,b^2(\pi)\,\frac{d^2}{d\pi^2}, \tag{11}$$

where

$$a(\pi) = \lambda(1-\pi),$$
$$b^2(\pi) = \Big(\frac{r}{\sigma}\Big)^2 \pi^2(1-\pi)^2.$$

`Remark 3.` It follows from (10) that

$$\int_0^t \pi_s\,ds = \frac{\pi_0 - \pi_t}{\lambda} + \frac{1}{\lambda}\frac{r}{\sigma}\int_0^t \pi_s(1-\pi_s)\,d\overline{B}_s + t.$$

Consequently, at least for τ satisfying the inequality $\mathsf{E}_\pi\tau < \infty$, we derive from (9) that $\mathbb{B}(c;\pi)$ can be represented as

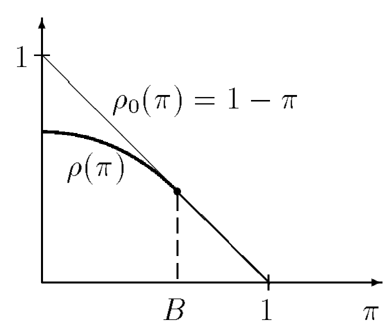
$$\mathbb{B}(c;\pi) = \inf_\tau \mathsf{E}_\pi\Big\{\Big(1+\frac{c}{\lambda}\,\pi\Big) - \Big(1+\frac{c}{\lambda}\Big)\,\pi_\tau + c\tau\Big\}\ \big(=\rho_*(\pi)\big). \tag{12}$$

□

Following the standard scheme of solving the quickest detection problems of the type (9) or (12) (see 16) consider the corresponding free-boundary (Stefan) problem (see Figure 3):

$$L\rho(\pi) = -c\pi, \qquad \pi \in [0,B), \tag{13}$$
$$\rho(B) = 1-B, \quad \pi \in [B,1], \qquad \text{(instantaneous stopping)}, \tag{14}$$
$$\rho'(B) = -1 \qquad \text{(smooth fit)}, \tag{15}$$
$$\rho'(0) = 0, \tag{16}$$

where L is the infinitesimal generator defined by (11).



`Figure 3.` The line $\rho_0(\pi) = 1-\pi$ corresponds to the risk of the instantaneous stopping.

A general solution of equation (13) includes two arbitrary constants. An additional unknown constant is the point B. Thus, we have *three* unknown parameters that can be found using conditions (14) for $\pi = B$, (15) and (16).

These considerations lead to the following solution of problems (13)–(16):

$$\rho(\pi) = \begin{cases} (1-B_*) - \displaystyle\int_\pi^{B_*} y_*(x)\,dx, & \pi \in [0, B_*], \\ 1-\pi, & \pi \in [B_*, 1] \end{cases}$$

with

$$y_*(x) = -C \int_0^x e^{-\Lambda[G(x)-G(y)]} \frac{dy}{y\,(1-y)^2},$$

$$G(y) = \log \frac{y}{1-y} - \frac{1}{y},$$

$$\Lambda = \lambda \Big/ \frac{r^2}{2\sigma^2}, \quad C = c \Big/ \frac{r^2}{2\sigma^2}.$$

The parameter B_* is defined as the root of the equation

$$C \int_0^{B_*} e^{-\Lambda[G(B_*)-G(y)]} \frac{dy}{y\,(1-y)^2} = 1. \tag{17}$$

(For more details, see [11]–[13]).

The standard technique of the "verification theorems" (see, for example, [15; p. 756]) shows that the obtained solution $\rho(\pi)$ is equal to $\rho_*(\pi)$ and the time $\tau_* = \tau_*(B_*)$ with

$$\tau_*(B_*) = \inf\{t : \pi_t \geq B_*\}$$

is optimal for any $0 \leq \pi \leq 1$. Namely, for any $\theta = \theta(\omega)$ with $\mathsf{P}(\theta = 0) = \pi$ and $\mathsf{P}(\theta \geq t \,|\, \theta > 0) = e^{-\lambda t}$ (here, $\pi \in [0,1]$ and $\lambda > 0$), we have

$$\rho_*(\pi) = \mathsf{E}_\pi\Big\{(1-\pi_{\tau_*}) + c\int_0^{\tau_*} \pi_s\,ds\Big\}$$

and

$$\mathbb{B}(c;\pi) = \mathsf{P}_\pi(\tau_* < \theta) + c\mathsf{E}_\pi(\tau_* - \theta)^+.$$

4. In order to solve the conditionally-extremal problem in **Variant A**, we use the method of the Lagrange multipliers and the obtained solution for Variant **B**.

Let $B_* = B_*(\lambda; c)$ be the value defined from equation (17) and $\mathfrak{M}_\alpha = \{\tau : \mathsf{P}_\pi(\tau < \theta) \leq \alpha\}$, where π is a *fixed* value in $[0,1]$ and α is a constant that defines the upper boundary for the probability of the false alarm $\mathsf{P}_\pi(\tau < \theta)$.

One can show that there exists a value c_α such that

$$B_*(\lambda, c_\alpha) = 1 - \alpha.$$

Then the stopping time $\tau_*(B_*) = \tau_*(B_*(\lambda, c_\alpha))$ belongs to $\mathfrak{M}_\alpha$ and we have

$$\begin{aligned} \mathbb{B}(c_\alpha, \pi) &= \inf_\tau \big\{\mathsf{P}_\pi(\tau < \theta) + c_\alpha \mathsf{E}_\pi(\tau - \theta)^+\big\} \\ &= \mathsf{P}_\pi\big(\tau_*(B_*) < \theta\big) + c_\alpha \mathsf{E}_\pi\big(\tau_*(B_*) - \theta\big)^+ \\ &= \mathsf{E}_\pi\big(1 - \pi_{\tau_*(B_*)}\big) + c_\alpha \mathsf{E}_\pi\big(\tau_*(B_*) - \theta\big)^+ \\ &= \alpha + c_\alpha \mathsf{E}_\pi\big(\tau_*(B_*) - \theta \,\big|\, \tau_*(B_*) \geq \theta\big)(1-\alpha). \end{aligned}$$

Since the time $\tau_*(B_*(\lambda, c_\lambda))$ is optimal for Variant **B**, this time is also optimal in the class $\mathfrak{M}_\alpha$. Furthermore, the mean delay in discovering the change point, i.e.

$$\mathbb{R}(\alpha,\lambda) = \mathsf{E}_\pi\big(\tau_*(B_*) - \theta \,\big|\, \tau_*(B_*) \ge \theta\big),$$

is given by the following formula (see [12], [13]):

$$\mathbb{R}(\alpha,\lambda) = \frac{\int_0^{1-\alpha}\Big[\int_0^x e^{-\frac{\lambda}{\nu}[G(x)-G(y)]}\,\frac{dy}{y\,(1-y)^2}\Big]\,dx}{(1-\alpha)\,\nu}$$

with $\nu = \frac{r^2}{2\sigma^2}$.

Now, let $\lambda \to 0$, $\alpha \to 1$ in such a way that

$$\frac{1-\alpha}{\lambda} \to T, \tag{18}$$

where $T > 0$ is a constant. Then

$$\mathbb{R}(T) \equiv \lim \mathbb{R}(\alpha,\lambda) = \frac{1}{\nu}\bigg\{e^b(-Ei(-b)) - 1 + b\int_0^\infty e^{-bz}\,\frac{\log(1+z)}{z}\,dz\bigg\},$$

where

$$b = \frac{1}{\nu T}, \quad -Ei(-y) = \int_y^\infty \frac{e^{-z}}{z}\,dz. \tag{19}$$

Using the above formula, we get for $\nu = 1$:

$$\mathbb{R}(T) = \begin{cases} \log T - 1 - \mathbb{C} + O(\frac{1}{T}), & T \to \infty, \\ \frac{T}{2} + O(T^2), & T \to 0, \end{cases} \tag{20}$$

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

Let us mention the following important property in connection with the passage to the limit $\lambda \to 0$, $\alpha \to 1$ (together with condition (18)). Suppose that $\pi = 0$ and $\mathsf{P}(\theta \ge t) = e^{-\lambda t}$. Let $(a,b) \in (A,B)$. Then

$$\mathsf{P}\big(\theta \in (a,b) \,|\, \theta \in (A,B)\big) \longrightarrow \frac{b-a}{B-A}.$$

In other words, using the limit procedure as $\lambda \to \infty$, we get from the exponential distribution on $[0,\infty)$ a *generalized distribution* on $[0,\infty)$ that is *conditionally uniform* in the following sense: the conditional distribution of θ under the condition that θ appears within (A,B) is the uniform distribution on (A,B).

This assumption on the distribution of θ is rather natural in the case where *there is no information* on the distribution of θ. Thus, the use of the exponential distribution, combined with the passage to the limit as $\lambda \to 0$, can be regarded as a useful *technical means* for the study of the schemes with the conditionally uniform distribution.

5. The above reasoning shows that the time $\tau_\alpha^* = \tau_*(B_*(\lambda, c_\alpha))$, i.e. the time

$$\tau_\alpha^* = \inf\{T : \pi_t \ge 1-\alpha\}, \tag{21}$$

is optimal in the class $\mathfrak{M}_\alpha$:

$$\inf_{\tau\in M_\alpha} \mathsf{E}(\tau-\theta\,|\,\tau\ge\theta)=\mathsf{E}(\tau_\alpha^*-\theta\,|\,\tau_\alpha^*\ge\theta).$$

Using the above notation $\varphi_t=\frac{\pi_t}{1-\pi_t}$, we can express (21) as

$$\tau_\alpha^*=\inf\Big\{t:\frac{\varphi_t(\lambda)}{\lambda}\ge\frac{1-\alpha}{\alpha\lambda}\Big\}.$$

For $\psi_t=\lim_{\lambda\to\infty}\frac{\varphi_t(\lambda)}{\lambda}$ (see (6), (7)), we find that, in the case $\pi=0$,

$$d\psi_t=dt+\frac{r^2}{\sigma^2}\,\psi_t\,dX_t,\quad \psi_0=0$$

or, equivalently,

$$\psi_t=t+\frac{r}{\sigma^2}\int_0^t\psi_s\,dX_s.$$

If $\theta=\infty$, then

$$\psi_t=t+\frac{r}{\sigma}\int_0^t\psi_s\,dB_s.$$

Passing to the limit $\lambda\to 0$, $\alpha\to 1$ in such a way that

$$\frac{1-\alpha}{\lambda}\to T,$$

we find that

$$\tau_\alpha^*\to\tau^*(T)=\inf\{t:\psi_t\ge T\}.$$

Consequently,

$$\psi_{\tau^*(T)}=\tau^*(T)+\frac{r}{\sigma}\int_0^{\tau^*(T)}\psi_s\,dB_s.$$

It follows that

$$T=\mathsf{E}_\infty\psi_{\tau^*(T)}=\mathsf{E}_\infty\tau^*(T).$$

In other words, the constant T has a simple intuitive meaning: it is the *mean time up to the false alarm.* The value

$$\mathbb{R}(T)=\lim_{\left\{\lambda\to 0,\alpha\to 1,\frac{1-\alpha}{\lambda}\to T\right\}}\mathbb{R}(\alpha,\lambda)$$

is the *mean delay time* in discovering the change-point (arbitrage). It follows from (20) that, for large T, this time has order $\log T$ while, for small T, this time has order $T/2$ (for the case $\nu=\frac{r^2}{2\sigma^2}=1$).

6. We will now turn to the parametric *minimax* formulations of the quickest detection problems. In these formulations, θ is supposed to be a parameter with values in $[0,\infty)$.

The formulation in **Variant C** is to find

$$\mathbb{C}(T)=\inf_{\tau\in\mathfrak{M}^T}\sup_\theta\operatorname*{esssup}_\omega\mathsf{E}_\theta\big((\tau-\theta)^+\,|\,\mathcal{F}_\theta\big)(\omega)$$

and the corresponding optimal stopping time. Here,

$$\mathfrak{M}^T = \{\tau : \mathsf{E}_\infty \tau = T\}$$

and $\mathcal{F}_\theta = \mathcal{F}_\theta^X = \sigma(X_s;\, s \le \theta)$.

For $\tau \in \mathfrak{M}^T$, we set

$$C(\tau) = \sup_\theta \operatorname*{esssup}_\omega \mathsf{E}_\theta\big((\tau-\theta)^+ \,|\, \mathcal{F}_\theta\big)(\omega).$$

The main idea in finding the optimal stopping time in Variant **C** is to give a *lower* estimate for $C(\tau)$ (τ belongs to $\mathfrak{M}^T$).

The author showed in [14] that the following estimate is true (see details below in Subsection 7):

$$C(\tau) \ge \frac{\mathsf{E}_\infty \int_0^\tau \gamma_t\, dt}{\mathsf{E}_\infty \gamma_\tau}, \tag{22}$$

where $\gamma = (\gamma_t)_{t\ge 0}$ is the exponential CUSUM-statistics introduced above:

$$\gamma_t = \sup_{\theta \le t} \frac{L_t}{L_\theta}.$$

Here,

$$L_t = \frac{d\mathsf{P}_0}{d\mathsf{P}_t}(t, X) = \exp\Big\{\frac{r}{\sigma^2} X_t - \frac{r^2}{2\sigma^2} t\Big\}.$$

Suppose that $\nu = \frac{r^2}{2\sigma^2} = 1$ and set

$$\tau^*(B) = \inf\{t : \gamma_t \ge B\}, \quad \gamma_0 = 1.$$

Using the Markov property of the process $\gamma = (\gamma_t)_{t\ge 0}$, one can show that

$$C(\tau^*(B)) = \mathsf{E}_0 \tau^*(B)$$

(the proof of this property is given in Subsection 7 below).

Let us now find $B = B_T$ such that

$$\tau^*(B) \in \mathfrak{M}^T = \{\tau : \mathsf{E}_\infty \tau = T\}.$$

Using the standard method of the differential backward Kolmogorov equation and taking equality (27) (it is given below) into account, we deduce that

$$\mathsf{E}_\infty \tau^*(B) = B - 1 - \log B,$$

and consequently, B_T is defined from the equation

$$B_T - 1 - \log B_T = T.$$

By the Markov property of the process γ, we get by analogy with $\mathsf{E}_\infty \tau^*(B)$,

$$\mathsf{E}_0 \tau^*(B_T) = \frac{B_T \log B_T + 1 - B_T}{B_T}. \tag{23}$$

Using the inequalities

$$\mathsf{E}_0\tau^*(B_T) \ge C(\tau^*(B_T)),$$

$$C(\tau) \ge \frac{\mathsf{E}_\infty \int_0^\tau \gamma_t\,dt}{\mathsf{E}_\infty\gamma_\tau},$$

we get

$$\begin{aligned}\mathsf{E}_0\tau^*(B_T) &\ge C(\tau^*(B_T)) \ge \inf_{\tau\in\mathfrak{M}^T} C(\tau) \ge \frac{\inf\limits_{\tau\in\mathfrak{M}^T} \mathsf{E}_\infty \int_0^\tau \gamma_t\,dt}{\sup\limits_{\tau\in\mathfrak{M}^T} \mathsf{E}_\infty\gamma_\tau} \\ &= \frac{\inf\limits_{\tau\in\mathfrak{M}^T} \mathsf{E}_\infty \int_0^\tau \gamma_t\,dt}{B_T} = \frac{\mathsf{E}_\infty \int_0^{\tau^*(B_T)} \gamma_t\,dt}{B_T} = \frac{B_T\log B_T + 1 - B_T}{B_T}.\end{aligned} \tag{24}$$

From (23) and (24), it follows that $\tau^*(B_T)$ is the *optimal* stopping time in the class $\mathfrak{M}^T = \{\tau : \mathsf{E}_\infty\tau = T\}$:

$$\mathbb{C}(T) = C(\tau^*(B_T))$$

and

$$\begin{aligned}\mathbb{C}(T) &= \frac{B_T\log B_T + 1 - B_T}{B_T} \\ &\sim \begin{cases}\log T - 1 + O\big(\frac{1}{T}\big), & T\to\infty, \\ \frac{T}{2} + O(T^2), & T \to 0.\end{cases}\end{aligned}$$

Thus, the exponential CUSUM-process $\gamma = (\gamma_t)_{t\ge 0}$ defined as

$$\gamma_t = \sup_{\theta\le t}\frac{L_t}{L_\theta}$$

is the *optimal* statistics in Variant **C**.

It is interesting to note that, for Variant **C**, the statistics $\psi = (\psi_t)_{t\ge0}$ is *asymptotically optimal*.

Indeed, take

$$\psi_t = \int_0^t \frac{L_t}{L_\theta}\,d\theta$$

and set

$$\sigma_T^* = \inf\{t : \psi_t \ge T\}.$$

Then $\mathsf{E}_\infty\sigma_T^* = T$. Using (7) and (1), we find that

$$\mathsf{E}_0\sigma_T^* = e^b(-Ei(-b)), \quad b = \frac{1}{T}$$

(the function Ei is defined in (19)). By the arguments similar to those used to obtain (24) (see also Section 7 below), we can show that

$$\mathsf{E}_0\sigma_T^* \ge C(\sigma_T^*) \ge \frac{1}{T}\,\mathsf{E}_\infty\int_0^{\sigma_T^*}\psi_s\,ds.$$

For $b \to 0$, we have

$$e^b(-Ei(-b)) = -\mathbb{C} - \log b + O(b),$$

where $\mathbb{C} = 0.577\ldots$ is the Euler constant. Thus, for $T \to \infty$, we have

$$\mathsf{E}_0 \sigma_T^* = \log T - \mathbb{C} + O\Big(\frac{1}{T}\Big). \tag{25}$$

It can also be shown that

$$\frac{1}{T}\mathsf{E}_\infty \int_0^{\sigma_T^*} \psi_s\, ds = \log T - \mathbb{C} - 1 + O\Big(\frac{1}{T}\Big). \tag{26}$$

Consequently,

$$\log T - \mathbb{C} + O\Big(\frac{1}{T}\Big) \geq C(\sigma_T^*) \geq \log T - \mathbb{C} - 1 + O\Big(\frac{1}{T}\Big).$$

On the other hand, as we have already seen,

$$\mathbb{C}(T) = \log T - 1 + O\Big(\frac{1}{T}\Big).$$

Thus, the statistics $\psi = (\psi_t)_{t\geq 0}$ is *asymptotically* $(T \to \infty)$ optimal in Variant **C**.

7. We now turn to the proof of the basic inequality (22) that was used to prove that the exponential CUSUM-statistics is optimal in Variant **C**.

Set

$$C_\theta(\tau;\omega) = \mathsf{E}_\theta\big((\tau-\theta)^+ \mid \mathcal{F}_\theta\big)(\omega).$$

We have

$$\gamma_t = \sup_{\theta\leq t} \frac{L_t}{L_\theta} = \frac{L_t}{\inf_{\theta\leq t} L_\theta} = \frac{L_t}{N_t},$$

$$d\gamma_t = d\Big(\frac{L_t}{N_t}\Big) = \frac{dL_t}{N_t} - \frac{L_t\, dN_t}{(N_t)^2} = \frac{r}{\sigma^2}\,\gamma_t\, dX_t - \gamma_t\,\frac{dN_t}{N_t}.$$

Note that $\gamma_t \geq 1$ and N changes its values only on the set $\{(t,\omega) : \gamma_t = 1\}$. Hence,

$$d\gamma_t = \frac{r}{\sigma^2}\,\gamma_t\, dX_t - \gamma_t\, I(\gamma_t = 1)\,\frac{dN_t}{N_t}.$$

Denote

$$H_t = -\int_0^t \gamma_s\, I(\gamma_s = 1)\,\frac{dN_s}{N_s}.$$

Then

$$d\gamma_t = dH_t + \frac{r}{\sigma^2}\,\gamma_t\, dX_t, \quad \gamma_0 = 1, \tag{27}$$

and therefore,

$$\gamma_t = L_t\Big[1 + \int_0^t \frac{dH_\theta}{L_\theta}\Big].$$

We have

$$C(\tau)\,\mathsf{E}_\infty H_\tau = \mathsf{E}_\infty\big(C(\tau)H_\tau\big) = \mathsf{E}_\infty\Big(C(\tau)\int_0^\infty I(\theta\le\tau(\omega))\,dH_\theta(\omega)\Big)$$
$$\ge \mathsf{E}_\infty\int_0^\infty C_\theta(\tau;\omega)\,I(\theta\le\tau(\omega))\,dH_\theta(\omega),$$

where $C_\theta(\tau;\omega)$ is defined above and $C(\tau)$ is defined in Subsection 6.

Since

$$(\tau-\theta)^+ = \int_\theta^\infty I(u\le\tau)\,du,$$

we get

$$C_\theta(\tau;\omega) = \int_\theta^\infty \mathsf{E}_\theta\big(I(u\le\tau)\,|\,\mathcal{F}_\theta\big)(\omega)\,du$$
$$= \int_\theta^\infty \mathsf{E}_\infty\Big(\frac{L_u}{L_\theta}\,I(u\le\tau)\,\Big|\,\mathcal{F}_\theta\Big)(\omega)\,du = \mathsf{E}_\infty\Big(\int_\theta^\tau \frac{L_u}{L_\theta}\,du\,\Big|\,\mathcal{F}_\theta\Big)(\omega).$$

We used here the fact that $\xi = I(u\le\tau)$ is $\mathcal{F}_u$-measurable and

$$\mathsf{E}_\theta(\xi\,|\,\mathcal{F}_\theta) = \mathsf{E}_\infty\Big(\xi\,\frac{L_u}{L_\theta}\,\Big|\,\mathcal{F}_\theta\Big).$$

As a result,

$$C(\tau)\,\mathsf{E}_\infty H_\tau \ge \mathsf{E}_\infty\int_0^\infty I(\theta\le\tau(\omega))\,\mathsf{E}_\infty\Big(\int_\theta^\tau \frac{L_u}{L_\theta}\,du\,\Big|\,\mathcal{F}_\theta\Big)\,dH_\theta$$
$$= \mathsf{E}_\infty\int_0^\tau \mathsf{E}_\infty\Big(\int_\theta^\tau \frac{L_u}{L_\theta}\,du\,\Big|\,\mathcal{F}_\theta\Big)\,dH_\theta.$$

Set

$$\widetilde{H}_\theta = \int_0^\theta \frac{dH_s}{L_s}, \quad \xi_\theta = \int_0^\theta L_s\,ds.$$

Then

$$C(\tau)\,\mathsf{E}_\infty H_\tau \ge \mathsf{E}_\infty\int_0^\tau \mathsf{E}_\infty\Big(\int_\theta^\tau \frac{L_u}{L_\theta}\,du\,\Big|\,\mathcal{F}_\theta\Big)\,dH_\theta = \mathsf{E}_\infty\int_0^\tau \mathsf{E}_\infty\Big(\int_\theta^\tau L_u\,du\,\Big|\,\mathcal{F}_\theta\Big)\,d\widetilde{H}_\theta$$
$$= \mathsf{E}_\infty\int_0^\tau \mathsf{E}_\infty\Big(\int_0^\tau L_u\,du\,\Big|\,\mathcal{F}_\theta\Big)\,d\widetilde{H}_\theta - \mathsf{E}_\infty\int_0^\tau \mathsf{E}_\infty\Big(\int_0^\theta L_u\,du\,\Big|\,\mathcal{F}_\theta\Big)\,d\widetilde{H}_\theta$$
$$= \mathsf{E}_\infty\int_0^\tau \mathsf{E}_\infty(\xi_\tau\,|\,\mathcal{F}_\theta)\,d\widetilde{H}_\theta - \mathsf{E}_\infty\int_0^\tau \xi_\theta\,d\widetilde{H}_\theta.$$

The process $M_\theta = \mathsf{E}_\infty(\xi_\tau\,|\,\mathcal{F}_\theta)$ is a P_∞-martingale and

$$\mathsf{E}_\infty\int_0^\tau M_\theta\,d\widetilde{H}_\theta = \mathsf{E}_\infty M_\tau\widetilde{H}_\tau - \mathsf{E}_\infty M_0\widetilde{H}_0 = \mathsf{E}_\infty\xi_\tau\widetilde{H}_\tau.$$

Thus,

$$C(\tau)\,\mathsf{E}_\infty H_\tau \ge \mathsf{E}_\infty \xi_\tau \widetilde{H}_\tau - \mathsf{E}_\infty \int_0^\tau \xi_\theta\, d\widetilde{H}_\theta = \mathsf{E}_\infty \int_0^\tau \widetilde{H}_\theta\, d\xi_\theta$$
$$= \mathsf{E}_\infty \int_0^\tau \widetilde{H}_\theta\, L_\theta\, d\theta = \mathsf{E}_\infty \int_0^\tau L_\theta \biggl(\int_0^\theta \frac{dH_s}{L_s}\biggr)\, d\theta$$
$$= \mathsf{E}_\infty \int_0^\tau \biggl[\int_0^\theta \frac{L_\theta}{L_s}\, dH_s\biggr]\, d\theta = \mathsf{E}_\infty \int_0^\tau [\gamma_\theta - L_\theta]\, d\theta$$
$$= \mathsf{E}_\infty \int_0^\tau \gamma_\theta\, d\theta - \mathsf{E}_\infty \int_0^\tau L_\theta\, d\theta.$$

We have

$$\gamma_t = 1 + H_t + \int_0^t \frac{r}{\sigma^2}\, \gamma_s\, dX_s.$$

Hence,

$$\mathsf{E}_\infty \gamma_\tau = 1 + \mathsf{E}_\infty H_\tau,$$

and, from the obtained inequality

$$C(\tau)\,\mathsf{E}_\infty H_\tau \ge \mathsf{E}_\infty \int_0^\tau \gamma_\theta\, d\theta - \mathsf{E}_\infty \int_0^\tau L_\theta\, d\theta,$$

we get

$$C(\tau)[\mathsf{E}_\infty \gamma_\tau - 1] \ge \mathsf{E}_\infty \int_0^\tau \gamma_\theta\, d\theta - \mathsf{E}_\infty \int_0^\tau L_\theta\, d\theta. \tag{28}$$

Now, note that, at least for bounded stopping times τ,

$$C(\tau) = \sup_\theta \operatorname*{ess\,sup}_\omega \mathsf{E}_\theta\bigl((\tau-\theta)^+ \mid \mathcal{F}_\theta\bigr)(\omega) \ge \mathsf{E}_0 \tau = \mathsf{E}_\infty(\tau L_\tau) = \mathsf{E}_\infty \int_0^\tau L_\theta\, d\theta \tag{29}$$

as $d(tL_t) = L_t\, dt + t\, dL_t$. In the general case of finite τ, we have

$$C(\tau) \ge \mathsf{E}_0 \tau \ge \mathsf{E}_0 (\tau \wedge N)\, L_{\tau\wedge N} = \mathsf{E}_\infty \int_0^{\tau\wedge N} L_\theta\, d\theta \;\uparrow\; \mathsf{E}_\infty \int_0^\tau L_\theta\, d\theta.$$

From (28) and (29), we get

$$C(\tau)\,\mathsf{E}_\infty \gamma_\tau \ge \mathsf{E}_\infty \int_0^\tau \gamma_\theta\, d\theta.$$

Thus,

$$C(\tau) \ge \frac{\mathsf{E}_\infty \int_0^\tau \gamma_\theta\, d\theta}{\mathsf{E}_\infty \gamma_\tau}$$

and

$$\mathbb{C}(T) = \inf_{\tau\in\mathfrak{M}^T} C(\tau) \ge \frac{\inf_{\tau\in\mathfrak{M}^T} \mathsf{E}_\infty \int_0^\tau \gamma_\theta\, d\theta}{\sup_{\tau\in\mathfrak{M}^T} \mathsf{E}_\infty \gamma_\tau},$$

where

$$\mathfrak{M}^T = \{\tau : \mathsf{E}_\infty \tau = T\}.$$

We will now give a proof of the property

$$\mathsf{E}_0\tau^*(B) = C(\tau^*(B)),$$

where

$$\tau^*(B) = \inf\{t : \gamma_t \ge B\},$$
$$C(\tau^*(B)) = \sup_{\theta} \operatorname{esssup}_{\omega} \mathsf{E}_\theta\big((\tau^*(B)-\theta)^+ \,|\, \mathcal{F}_\theta\big)(\omega).$$

On the set $\{\omega : \tau^*(B) \le \theta\}$, we have:

$$\mathsf{E}_\theta\big((\tau^*(B)-\theta)^+ \,|\, \mathcal{F}_\theta\big)(\omega) = 0.$$

Consider the set

$$\{\omega : \tau^*(B) > \theta\} = \{\omega : \gamma_u < B,\, u \le \theta\}.$$

By the Markov property of the process $\gamma = (\gamma_t)$, we have on the set

$$\{\omega : \tau^*(B) > \theta\} \cap \{\gamma_\theta = x\}$$

that

$$\mathsf{E}_\theta\big((\tau^*(B)-\theta)^+ \,|\, \mathcal{F}_\theta\big)(\omega) = f(x),$$

i.e. this conditional expectation is the function only of the value x of γ_θ. It is clear that $\max_{1\le x\le B} f(x) = f(1)$ and

$$f(1) = \mathsf{E}_0\tau^*(B).$$

Hence,

$$C(\tau^*(B)) = \mathsf{E}_0\tau^*(B).$$

8. Let us now consider the quickest detection problem in **Variant D**. One should find

$$\mathbb{D}(T) = \inf_{\tau\in\mathfrak{M}^T} \sup_{\theta} \mathsf{E}_\theta(\tau-\theta \,|\, \tau \ge \theta)$$

and the optimal stopping time in the class $\mathfrak{M}^T$.

We will follow the same scheme as that used in Variant **C**. Let us first prove that

$$\mathsf{E}_0\sigma_T^* \ge \mathbb{D}(T) \ge \inf_{\tau\in\mathfrak{M}^T} \frac{\mathsf{E}_\infty \int_0^\tau \psi_s\,ds}{T}, \tag{30}$$

where

$$\sigma_T^* = \inf\{t \ge 0 : \psi_t \ge T\}.$$

We have

$$D_\theta(\tau) = \mathsf{E}_\theta(\tau-\theta \,|\, \tau\ge\theta) = \mathsf{E}_\theta\big((\tau-\theta)^+ \,|\, \tau\ge\theta\big) = \int_\theta^\infty \mathsf{E}_\theta\big(I(u\le\tau) \,|\, \tau\ge\theta\big)\,du$$
$$= \int_\theta^\infty \frac{\mathsf{E}_\theta I(u\le\tau)}{\mathsf{E}_\theta I(\theta\le\tau)}\,du = \int_\theta^\infty \frac{\mathsf{E}_\theta I(u\le\tau)}{\mathsf{E}_\infty I(\theta\le\tau)}\,du = \int_\theta^\infty \frac{\mathsf{E}_\infty\big(\frac{L_u}{L_\theta} I(u\le\tau)\big)}{\mathsf{E}_\infty I(\theta\le\tau)}\,du.$$

Thus,

$$\mathbb{D}(T)\,\mathsf{E}_\infty I(\tau\le\theta) \ge D_\theta(\tau)\,\mathsf{E}_\infty I(\theta\le\tau) = \int_\theta^\infty \mathsf{E}_\infty\left(\frac{L_u}{L_\theta}\, I(u\le\tau)\right)du = \mathsf{E}_\infty\int_\theta^\tau \frac{L_u}{L_\theta}\,du.$$

Integrating on θ (from 0 to ∞), we get

$$\mathbb{D}(T)\,\mathsf{E}_\infty\tau \ge \mathsf{E}_\infty\int_0^\infty\left(\int_\theta^\tau\frac{L_u}{L_\theta}\,du\right)d\theta = \mathsf{E}_\infty\int_0^\tau\left(\int_0^\theta\frac{L_\theta}{L_u}\,du\right)d\theta = \mathsf{E}_\infty\int_0^\tau\psi_\theta\,d\theta.$$

As a result,

$$\begin{aligned}\mathsf{E}_0\sigma_T^* &= \sup_\theta \mathsf{E}_\theta(\sigma_T^*-\theta\,|\,\sigma_T^*\ge\theta) \ge \inf_{\tau\in\mathfrak{M}^T}\sup_\theta \mathsf{E}_\theta(\tau-\theta\,|\,\tau\ge\theta)\\ &\ge \inf_{\tau\in\mathfrak{M}^T}\frac{\mathsf{E}_\infty\int_0^\tau\psi_s\,ds}{\mathsf{E}_\infty\tau} = \inf_{\tau\in\mathfrak{M}^T}\frac{\mathsf{E}_\infty\int_0^\tau\psi_s\,ds}{T},\end{aligned} \tag{31}$$

where $\mathfrak{M}^T=\{\tau : \mathsf{E}_\infty\tau = T\}$.

Similarly to Variant **C**, we have here

$$\sup_\theta \mathsf{E}_\theta(\sigma_T^*-\theta\,|\,\sigma_T^*\ge\theta) = \mathsf{E}_0\sigma_T^*.$$

This, together with (31), leads to the desired inequalities (30).

It has already been mentioned in Subsection 6 that, for large T, formulas (25) and (26) are true. Combining these formulas with inequalities (30), we deduce that the statistics $\psi=(\psi_t)_{t\ge0}$ is asymptotically ($T\to\infty$) optimal.

9. Variant E. In Section 1, we considered a particular case of the observation procedure for which the observations are multistage, i.e. the observations are continued after each alarm. It is also important that the change-point should be preceded by a long period of observations and a stationary regime of observations should be established within this period. The change-point can appear only after this stationary regime has been established.

The analysis of the observation procedure that was described in Section 1 shows that any observation procedure is eventually described by a sequence of stopping times $\delta = (\tau_1,\tau_2,\dots)$ such that $\tau_1,\tau_2,\dots$ are independent identically distributed random variables with respect to the measure P_∞. For this procedure, the alarms are raised at the times $\tau_1, \tau_1+\tau_2,\dots$:

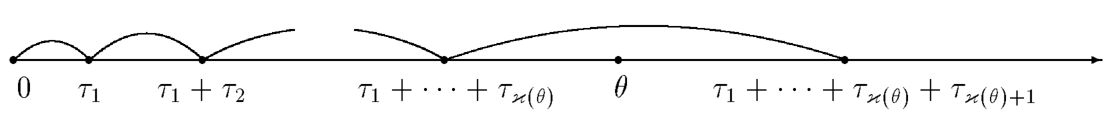


Figure 4

For $\theta>0$, define $\varkappa(\theta)$ (it takes values in $0,1,\dots$) from the inequality

$$\tau_0+\tau_1+\dots+\tau_{\varkappa(\theta)} < \theta \le \tau_0+\tau_1+\dots+\tau_{\varkappa(\theta)}+\tau_{\varkappa(\theta)+1},$$

where $\tau_0=0$.

For a fixed θ and an observation procedure $\delta = (\tau_1, \tau_2, \dots)$, the mean delay time in discovering the time θ equals

$$R^\delta_\theta(T) = \mathsf{E}_\theta(\tau_1 + \dots + \tau_{\varkappa(\theta)+1} - \theta)$$

(we assume here that $\mathsf{E}_\infty \tau_i = T > 0$).

Let

$$F^\delta_\theta(u) = \mathsf{P}_\theta\big(\theta - (\tau_1 + \dots + \tau_{\varkappa(\theta)}) \le u\big).$$

Then

$$\begin{aligned} R^\delta_\theta(T) &= \int_0^\infty \mathsf{E}_\theta\big(\tau_1 + \dots + \tau_{\varkappa(\theta)+1} - \theta \,|\, \theta - (\tau_1 + \dots + \tau_{\varkappa(\theta)}) = u\big)\, dF^\delta_\theta(u) \\ &= \int_0^\infty \mathsf{E}_u(\tau_1 - u \,|\, \tau_1 > u)\, dF^\delta_\theta(u). \end{aligned}$$

We will suppose that the observation procedure $\delta = (\tau_1, \tau_2, \dots)$ is such that the distribution of $\mathrm{Law}(\tau_i | \mathsf{P}_\infty)$ is nonlattice. Then the general renewal theory guarantees that there exists a limit distribution

$$F^\delta_\infty(u) = w\text{-}\lim_{\theta \to \infty} F^\delta_\theta(u).$$

By the well-known *Basic Renewal Theorem*,

$$F^\delta_\infty(u) = \frac{1}{T} \int_0^u (1 - F(x))\, dx,$$

where $F(x) = \mathsf{P}_\infty(\tau_1 \le x)$.

Suppose now that the method δ is such that, for $\theta \to \infty$,

$$\begin{aligned} R^\delta_\theta(T) &= \int_0^\infty \mathsf{E}_u(\tau_1 - u | \tau_1 \ge u)\, dF^\delta_\theta(u) \\ &\to \int_0^\infty \mathsf{E}_u(\tau_1 - u | \tau_1 \ge u)\, dF^\delta_\infty(u) \\ &= \frac{1}{T} \int_0^\infty \mathsf{E}_u(\tau_1 - u | \tau_1 \ge u)\, \mathsf{P}_\infty(\tau_1 \ge u)\, du \\ &= \frac{1}{T} \int_0^\infty \mathsf{E}_u(\tau_1 - u | \tau_1 \ge u)\, \mathsf{P}_u(\tau_1 \ge u)\, du \\ &= \frac{1}{T} \int_0^\infty \mathsf{E}_u(\tau_1 - u)^+\, du \ \big(= R^\delta(T)\big). \end{aligned}$$

For the Bayes setting with $\mathrm{Law}(\theta) = \exp(\lambda)$, we have

$$\mathsf{E}(\tau - \theta)^+ = \lambda \int_0^\infty e^{-\lambda u}\, \mathsf{E}_u(\tau_1 - u)^+\, du.$$

Fix $T > 0$ and let $\lambda \to 0$, $\alpha \to 1$ in such a way that $\frac{1-\alpha}{\alpha\lambda} = T$. Then

$$\frac{1}{T}\int_0^\infty \mathsf{E}_u(\tau_1 - u)^+ \, du =$$

$$= \lim_{\{\lambda\to 0,\, \alpha\to 1,\, \frac{1-\alpha}{\alpha\lambda}=T\}} \frac{\lambda\int_0^\infty e^{-\lambda u}\,\mathsf{E}_u(\tau_1 - u)^+ \, du}{(1-\alpha)/\alpha}$$

$$= \lim_{\{\lambda\to 0,\, \alpha\to 1,\, \frac{1-\alpha}{\alpha\lambda}=T\}} \frac{\mathsf{E}(\tau_1 - \theta)^+}{1-\alpha}$$

$$= \lim_{\{\lambda\to 0,\, \alpha\to 1,\, \frac{1-\alpha}{\alpha\lambda}=T\}} \mathsf{E}(\tau_1 - \theta | \tau_1 \geq \theta).$$

From these formulas and the asymptotic ($\lambda \to 0$, $\alpha \to 1$, $\frac{1-\alpha}{\lambda} \to T$) optimality of the statistics $\psi = (\psi_t)_{t\geq 0}$, we conclude that this statistics is *optimal* for

Variant E′. Find

$$\mathbb{E}'(T) = \inf_{\tau\in\mathfrak{M}^T} \frac{1}{T}\int_0^\infty \mathsf{E}_\theta(\tau - \theta)^+ d\theta. \tag{32}$$

Note that, in this formulation, the parameter θ can be regarded as a *generalized* random variable with the "uniform distribution on $[0,\infty)$". Formulation (32) deals only with one stopping time (in other words, with *one stage* of observations). However, this formulation is directly related to the multistage procedure described above because $\mathbb{E}'(T) = \lim_{\theta\to\infty} R^\delta_\theta(T)$.

10. Table 1 sums up the results described above. These results are related to the optimality and the asymptotic optimality of the statistics $(\pi_t)_{t\geq 0}$, $(\gamma_t)_{t\geq 0}$ and $(\psi_t)_{t\geq 0}$ in various formulations of the quickest detection problem of the change-point (or arbitrage).

3 Some Comments

1. The researchers working in the field of finance as well as the market operators can be grouped, regarding their approach to the analysis of the dynamics of prices, as follows:

{1} "fundamentalists";

{2} "quantitative analysts";

{3} "technicians".

"Fundamentalists" make their decisions by looking at the state of the "economy at large" or by analyzing some of its sectors. The development prospects are of particular interest to them. The basis of their analysis is the assumption that the actions of the market operators are "rational". The second group ("quantitative analysts") emerged in the 1950s as the followers of L. Bachelier. This group is closer to the "fundamentalists" than to the "technicians" due to the fact that the "quantitative analysts" attach more significance to the rational aspects of the investors' behaviour than to the tones of the market.

It should be pointed out that the theoretical basis for groups {1} and {2} is currently stronger compared to that of group {3}. We believe that the problems considered in this paper are directly related to the technical analysis of the financial data and to the decision-making procedures.

θ	**Variants**	**Optimality of statistics**
r.v. $\exp(\lambda)$	**A:** $\inf\limits_{\tau\in\mathfrak{M}_\alpha} \mathsf{E}(\tau-\theta\mid\tau\ge\theta)$	$\pi=(\pi_t)$ is optimal $\pi_t=\mathsf{P}(\theta\le t\mid\mathcal{F}_t^X)$
r.v. $\exp(\lambda)$	**B:** $\inf\limits_{\tau}\big\{\mathsf{P}(\tau<\theta)+c\,\mathsf{E}(\tau-\theta)^+\big\}$	$\pi=(\pi_t)$ is optimal
parameter $\theta\in\mathbb{R}_+$	**C:** $\inf\limits_{\tau\in\mathfrak{M}^T}\sup\limits_{\theta}\operatorname*{esssup}\limits_{\omega}\mathsf{E}_\theta\big((\tau-\theta)^+\mid\mathcal{F}_\theta\big)$	$\gamma=(\gamma_t)$ is optimal $\gamma_t=\sup\limits_{\theta\le t}\dfrac{L_t}{L_\theta}$ (exponential CUSUM)
parameter $\theta\in\mathbb{R}_+$	**D:** $\inf\limits_{\tau\in\mathfrak{M}^T}\sup\limits_{\theta}\mathsf{E}_\theta(\tau-\theta\mid\tau\ge\theta)$	$\psi=(\psi_t)$ is asymptotically ($T\to\infty$) optimal $\gamma=(\gamma_t)$ is asymptotically ($T\to\infty$) optimal
θ is a generalized r.v.	**E′:** $\inf\limits_{\tau\in\mathfrak{M}^T}\dfrac{1}{T}\displaystyle\int_0^\infty\mathsf{E}_\theta(\tau-\theta)^+d\theta$	$\psi=(\psi_t)$ is optimal $\psi_t=\displaystyle\int_0^t\frac{L_t}{L_\theta}\,d\theta$

$$\mathfrak{M}_\alpha=\{\tau:\mathsf{P}(\tau\le\theta)\le\alpha\},$$

$$\mathfrak{M}^T=\{\tau:\mathsf{E}_\infty\tau=T\}.$$

Table 1. Optimal and asymptotically optimal statistics for Variants **A**–**E′**

2. The formulations of the quickest detection problem have a long history. Variants **A** and **B** were considered by the author in [10], [11], [12], [13]. The formulation of Variant **C** was proposed in [4]. In this paper, the asymptotic optimality of the statistics $\gamma = (\gamma_n)_{n \geq 0}$ was proved for the discrete-time case. The proof of the asymptotic optimality of $\gamma = (\gamma_t)_{t \geq 0}$ in the continuous-time case was given in our paper [14]. In particular, this paper contains the lower estimate (22) that is essential in the proof of the asymptotic optimality. For the discrete-time case, the corresponding estimate was proved in the paper [5].

The formulation of Variant **D** was considered in many papers. See, for instance, [8] and the collection of works [1].

It seems that our approach to the proof of the asymptotic optimality of the statistics $\psi = (\psi_t)_{t \geq 0}$ and $\gamma = (\gamma_t)_{t \geq 0}$ (based on the lower estimate (30)) has not previously been considered.

Finally, the optimality of the statistics $\psi = (\psi_t)_{t \geq 0}$ in Variants **E** and **E′** was proved in the author's papers [11] and [12].

References

[1] *E. Carlstein, H.-G. Müller, D. Siegmund (eds.* Change-point problems. Institute of Mathematical Statistics. Lecture Notes Monograph. Ser. 23, 1994.

[2] *S.E. Graversen, A.N. Shiryaev.* An extension of P. Lévy's distributional properties to the case of a Brownian motion with a drift.// Bernoulli, **6** (2000), No. 4, p. 615–620.

[3] *R.S. Liptser, A.N. Shiryaev.* Statistics of random processes, I and II. Springer-Verlag, 1977 and 1978.

[4] *G. Lorden.* Procedures for reacting to a change of distributions.// The Annals of Mathematical Statistics, **42** (1971), No. 6, p. 1897–1908.

[5] *G.V. Moustakides.* Optimal stopping times for detecting changes in distributions.// The Annals of Statistics, **14** (1986), No. 4, p. 1379–1387.

[6] *E.S. Page.* Continuous inspection schemes.// Biometrika, **41** (1954), p. 100–115.

[7] *J.L. Pedersen.* Optimal prediction of the ultimate maximum of Brownian motion. In thesis "Some results on optimal stopping and Skorokhod embedding with applications". University of Aarhus, January, 2000.

[8] *M. Pollak, D. Siegmund.* A diffusion process and its application to detecting a change in the drift of Brownian motion.// Biometrika, **72** (1985), No. 2, p. 267–280.

[9] *D. Revuz, M. Yor.* Continuous martingales and Brownian motion. Springer-Verlag, 1998.

[10] *A.N. Shiryaev.* Detection of spontaneously occurring effects.// Dokl. Acad. Nauk SSSR, **138** (1961), p. 799–801.

[11] *A.N. Shiryaev.* The problem of quickest detection of the destruction of a stationary regime.// Dokl. Acad. Nauk. SSSR, **138** (1961), p. 1039–1042.

[12] *A.N. Shiryaev.* On optimum methods in the quickest detection problems.// Theory of Probability and its Applications, **8** (1963), p. 22–46.

[13] *A.N. Shiryaev.* Optimal stopping rules. Springer-Verlag, 1978.

[14] *A.N. Shiryaev.* Minimax optimality of the method of Cumulative Sums (CUSUM) in the case of continuous time.// Russian Mathematical Surveys, **51** (1996), p. 750–751.

[15] *A.N. Shiryaev.* Essentials of stochastic finance. World Scientific, 1999.

[16] *D. Siegmund.* Sequential analysis. Springer-Verlag, 1985.