On the Chernoff and the Kiefer–Weiss sequential testing problems

Mikhail Zhitlukhin

The Russian Academy of Sciences, Moscow
The Higher School of Economics, Moscow

Ulm – September, 2013

The talk is based on the following papers:

1. Zhitlukhin M. V., Muravlev A. A.
   On Chernoff's hypotheses testing problem about the drift of a Brownian motion. Theory of Probability and Its Applications, vol. 57, no. 4, pp. 778–788, 2012.

2. Zhitlukhin M. V., Muravlev A. A., Shiryaev A. N.
   An optimal decision rule in the Kiefer–Weiss problem for a Brownian motion. Russian Mathematical Surveys, vol. 68, no. 2, pp. 201-202, 2013.

1. Introduction

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

$$X_t = \mu t + B_t, \qquad t \geqslant 0,$$

where μ is an unknown drift coefficient.

We consider the problem of sequentially testing the hypothesis

$$H_+ : \mu \geqslant 0 \text{ and } H_- : \mu < 0.$$

The two settings will be studied: the Chernoff problem (Bayesian) and the Kiefer–Weiss problem (minimax).

By definition, a decision rule is a pair (τ, d) consisting of

– a stopping time τ of the filtration $(\mathscr{F}_t)_{t\geqslant 0}$, $\mathscr{F}_t = \sigma(X_s; s \leqslant t)$

– an $\mathscr{F}_\tau$-measurable function d taking values ± 1

The moment τ represents the moment of stopping of the observation, and the value of d corresponds to the hypothesis accepted.

Chernoff's problem

Assume that μ is an $\mathscr{N}(\mu_0, \sigma_0^2)$ random variable with known parameters μ_0, σ_0 and independent of B; $c, k > 0$ are given real numbers.

The problem (Chernoff, 1961) is to find an optimal rule (τ^*, d^*):

$$\mathrm{E}[c\tau^* + k|\mu|\mathbf{I}(d^* \neq \operatorname{sgn}(\mu))] = \inf_{(\tau,d)} \mathrm{E}[c\tau + k|\mu|\mathbf{I}(d \neq \operatorname{sgn}(\mu))]$$

In other words, an optimal rule (τ^*, d^*) minimizes the average penalty consisting of the observation cost and the penalty for a wrong decision.

Without loss of generality we may assume $c = k = 1$ (Chernoff, 1961).

Kiefer-Weiss' problem

Assume that μ is an unknown real parameter and let $\varepsilon > 0$, $\alpha \in (1/2, 1)$ be given numbers. Let Δ_α be the class of decision rules (τ, d) such that

$$\mathrm{P}(d \neq \operatorname{sgn}(u) \mid \mu = u) \leqslant \alpha \text{ for any } |u| > \varepsilon.$$

The parameter ε specifies the indifference area: the decision "$\mu < 0$" is correct if $\mu \leqslant \varepsilon$, and "$\mu \geqslant 0$" is correct if $\mu \geqslant -\varepsilon$.

The parameter α specifies the maximal acceptable probability of error.

The problem (Kiefer, Weiss, 1957) is to find $(\tau^*, d^*) \in \Delta_\alpha$ such that

$$\sup_{u \in \mathbb{R}} \mathrm{E}(\tau^* \mid \mu = u) = \inf_{(\tau,d) \in \Delta_\alpha} \sup_{u \in \mathbb{R}} \mathrm{E}(\tau \mid \mu = u),$$

i.e. an optimal (τ^*, d^*) minimizes the maximal observation time over all decision rules in Δ_α.

Known results and the aim of the research

1. Chernoff and Breakwell (1961-63) showed that an optimal decision rule is of the form

$$\tau^* = \inf\{t \geqslant 0 : |X'_t| \geqslant a^*(t)\}, \qquad d^* = \operatorname{sgn}(X'_{\tau^*})$$

where X'_t is a process obtained from X_t by some transformation, and $a^*(t)$ is some function on $\mathbb{R}_+$ (independent of μ_0, σ_0).

They found the asymptotics of $a^*(t)$ when $t \to 0$ and $t \to \infty$, which corresponds to $\sigma_0 \to \infty$ and $\sigma_0 \to 0$.

2. Lai (1973) showed that in the Kiefer-Weiss problem

$$\tau^* = \inf\{t \geqslant 0 : |X_t| \geqslant b^*(t)\}, \qquad d^* = \operatorname{sgn}(X_{\tau^*})$$

He studied the asymptotics of $b^*(t)$ when $t \to \infty$.

3. Our aim is to find the boundaries $a^*(t)$ and $b^*(t)$.

2. Solution of the Chernoff problem

Recall that the problem is to find (τ^*, d^*) such that

$$\mathrm{E}[c\tau^* + k|\mu|\mathbf{I}(d^* \neq \operatorname{sgn}(\mu))] = \inf_{(\tau,d)} \mathrm{E}[c\tau + k|\mu|\mathbf{I}(d \neq \operatorname{sgn}(\mu))]$$

Fix the parameters (μ_0, σ_0) and introduce the process $W = (W_t)_{0\leqslant t\leqslant 1}$,

$$W_t = \sigma_0(1-t)X_{\frac{t}{\sigma_0^2(1-t)}} - t\mu_0/\sigma_0.$$

We check that W is a standard Brownian motion.

It turns out, (τ^*, d^*) can be found from the optimal stopping problem

$$V_{\mu_0,\sigma_0} = \inf_{\tau\in\mathfrak{M}_1^W} \mathrm{E}\left[\frac{2}{\sigma_0^3(1-\tau)} - \left|W_\tau + \mu_0/\sigma_0\right|\right],$$

where $\mathfrak{M}_1^W$ is the class of all stopping times $\tau \leqslant 1$ of the filtration $(\mathscr{F}_t^W)_{t\leqslant 1}$, $\mathscr{F}_t^W = \sigma(W_s; s\leqslant t)$.

Theorem

1) Let τ_W^* be an optimal stopping time in V_{μ_0,σ_0}. Then an optimal decision rule (τ^*, d^*) for testing H_+ and H_- is given by

$$\tau^* = \frac{\tau_W^*}{\sigma_0^2(1-\tau_W^*)}, \qquad d^* = \operatorname{sgn}(X_{\tau^*} + \mu_0/\sigma_0^2).$$

2) The moment $\tau_W^* = \tau_W^*(\mu_0, \sigma_0)$ is of the form

$$\tau_W^*(\mu_0, \sigma_0) = \inf\{0 \leqslant t \leqslant 1 : |W_t + \mu_0/\sigma_0| \geqslant a_{\sigma_0}^*(t)\},$$

where $a_{\sigma_0}^*(t)\colon [0,1] \to \mathbb{R}_+$ is a non-increasing continuous function being the unique solution of the equation (with some concrete function H)

$$(1-t)H(1-t, a(t)) = \int_t^1 \frac{1}{\sigma_0^3(1-s)^2}\left[\Phi\left(\frac{a(s)-a(t)}{\sqrt{s-t}}\right) - \Phi\left(\frac{-a(s)-a(t)}{\sqrt{s-t}}\right)\right] ds$$

in the class of continuous functions $a(t)$ satisfying the properties

$$0 < a(t) \leqslant \frac{\sigma_0^3}{4}(1-t) \text{ for } t < 1, \qquad a(1) = 0.$$

Remark

Chernoff and Breakwell showed that the optimal stopping time τ^* is given by

$$\tau^*(\mu_0, \sigma_0) = \{t \geqslant 0 : |X_t - \mu_0/\sigma_0^2| \geqslant b^*(t + 1/\sigma_0^2)\},$$

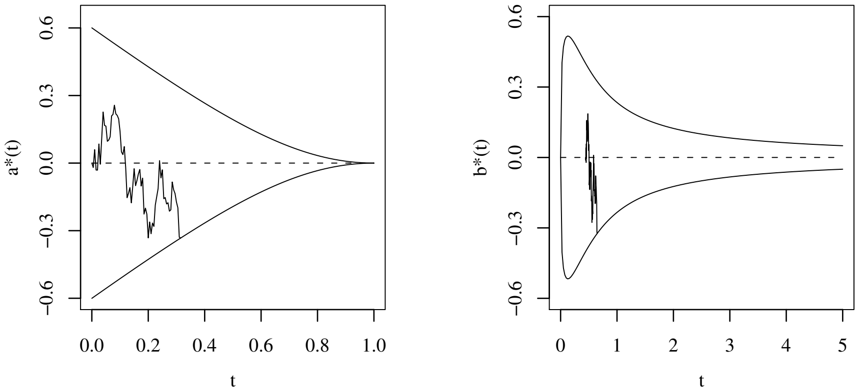
where $b^*(t)\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly positive function for $t > 0$, which does not depend on μ_0, σ_0.

As it follows from the structure of the processes W and X' the optimal stopping boundaries $a^*_{\sigma_0}$ and b^* are connected by

$$b^*(t) = \sigma_0 t \cdot a^*_{\sigma_0}\left(1 - \frac{1}{\sigma_0^2 t}\right), \qquad t \geqslant 1/\sigma_0^2.$$

The optimal stopping boundaries

Left: the boundary $a^*_{3/2}(t)$ for the process W.
Right: the boundary $b^*(t)$ for the process X'.



Outline of the proof

It is sufficient to consider only decision rules (τ, d) with $\mathrm{E}\tau < \infty$. For any such decision rule the average penalty is

$$R(\tau, d) = \mathrm{E}\big[\tau + E(\mu^- \mid \mathscr{F}_\tau)\mathbf{I}\{d = +1\} + \mathrm{E}(\mu^+ \mid \mathscr{F}_\tau)\mathbf{I}\{d = -1\}\big].$$

Thus the problem reduces to finding the stopping time τ^* minimizing

$$\mathscr{E}(\tau) = \mathrm{E}\big[\tau + \min\big\{\mathrm{E}(\mu^- \mid \mathscr{F}_\tau),\ \mathrm{E}(\mu^+ \mid \mathscr{F}_\tau)\big\}\big]$$

Using the Normal correlation theorem, we find

$$\mathscr{E}(\tau) = \mathrm{E}[\tau + H(\tau + 1/\sigma_0^2,\ X_\tau + \mu_0/\sigma_0^2)]$$

with the function

$$H(t, x) = \frac{1}{\sqrt{t}}\ \varphi(x/\sqrt{t}) - \frac{|x|}{t}\,\Phi(-|x|/\sqrt{t})).$$

Applying the Itô formula, we find that for any τ with $\mathrm{E}\tau < \infty$

$$\mathscr{E}(\tau) = \mathrm{E}\left[\tau - \frac{|X_\tau + m_0/\sigma_0^2|}{2(\tau + 1/\sigma_0^2)}\right] + H\left(\frac{1}{\sigma_0^2},\ \frac{m_0}{\sigma_0^2}\right) + \frac{|m_0|}{2}.$$

Then we check that the process

$$M_t = \frac{X_t + m_0/\sigma_0^2}{\sigma_0(t + 1/\sigma_0^2)} - \frac{m_0}{\sigma_0} \quad \text{is a martingale}.$$

Applying the change of time, we find that the process

$$W_t = M_{t/\sigma_0^2(1-t)} \quad \text{is a Brownian motion},$$

which reduces the Chernoff problem to the optimal stopping problem

$$V_{\mu_0,\sigma_0} = \inf_{\tau \leqslant 1} \mathrm{E}\left[\frac{2}{\sigma_0^3(1-\tau)} - \left|W_\tau + \mu_0/\sigma_0\right|\right]$$

It is solved using standard methods.

3. Solution of the Kiefer–Weiss problem

Recall the problem is to find (τ^*, d^*) such that

$$\sup_{u\in\mathbb{R}} \mathrm{E}(\tau^* \mid \mu = u) = \inf_{(\tau,d)\in\Delta_\alpha} \sup_{u\in\mathbb{R}} \mathrm{E}(\tau \mid \mu = u),$$

where Δ_α is the class of decision rules (τ, d) such that $\mathrm{P}(d \neq \mathrm{sgn}(u) \mid \mu = u) \leqslant \alpha$ for any $|u| > \varepsilon$.

The problem is reduced to the family of problems V_c, $c > 0$:

$$V_c = \inf_{(\tau,d)} \big[\mathrm{E}(\tau \mid \mu = 0) + c\{\mathrm{P}(d = -1 \mid \mu = \varepsilon) + \mathrm{P}(d = 1 \mid \mu = -\varepsilon)\}\big].$$

Namely, if for some $c > 0$ the optimal decision rule $\delta_c = (\tau_c, d_c)$ for V_c is such that

$$\mathrm{P}(d_c = -1 \mid \mu = \varepsilon) = \mathrm{P}(d_c = 1 \mid \mu = -\varepsilon) = \alpha,$$

then δ^c is optimal in the Kiefer–Weiss problem.

Using that $d(\mathrm{P}_t \mid \mu = u)/d(\mathrm{P}_t \mid \mu = u) = \exp(uX_t - u^2 t/2)$, we obtain

$$V_c = \inf_{(\tau,d)} \mathrm{E}\big[\tau + c\big(e^{\varepsilon X_\tau - \varepsilon^2 \tau/2}\mathbf{I}\{d = -1\} + e^{-\varepsilon X_\tau - \varepsilon^2 \tau/2}\mathbf{I}\{d = 1\}\big) \mid \mu = 0\big].$$

This implies that the optimal decision rule (τ_c, d_c) is such that

$$d_c = \operatorname{sgn} X_{\tau_c}$$

and τ_c solves the optimal stopping problem

$$V_c = \inf_{\tau} \mathrm{E}\big[\tau + ce^{-\varepsilon|X_\tau| - \varepsilon^2\tau/2} \mid \mu = 0\big].$$

The constant $c = c(\alpha)$ is found from the condition

$$\mathrm{P}(X_{\tau_c} > 0 \mid \mu = -\varepsilon) = \alpha.$$

Theorem

An optimal stopping time (τ^*, d^*) in the Kiefer–Weiss problem is of the form

$$\tau^* = \inf\{t \geqslant 0 : |X_t| \geqslant a^*(t + t_0)\}, \qquad d^* = \operatorname{sgn} X_{\tau^*},$$

where $a^*(t) > 0$ is a non-increasing function on $\mathbb{R}$, being the unique solution of the integral equation

$$\exp(-\varepsilon a(t) - \varepsilon^2 t/2) = \int_0^\infty \big[\Phi_s(a(s+t) - a(t)) - \Phi_s(-a(s+t) - a(t))\big]\, ds$$

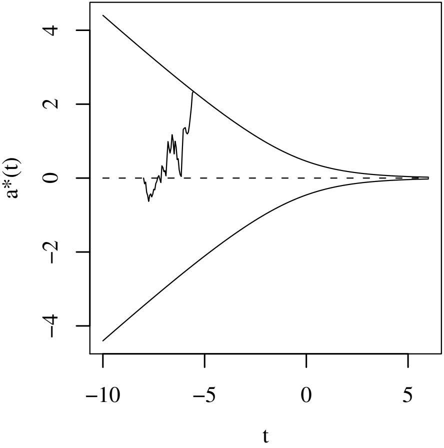
in the class of continuous function $a(t)$ on $\mathbb{R}$, satisfying the inequality

$$0 < a(t) \leqslant \varepsilon e^{-\varepsilon^2 t/2}/2, \qquad t \in \mathbb{R}.$$

The quantity $t_0 = t_0(\alpha)$ is found from the equation $\mathrm{P}(d^*(t_0) = 1 \mid \mu = -\varepsilon) = \alpha$.

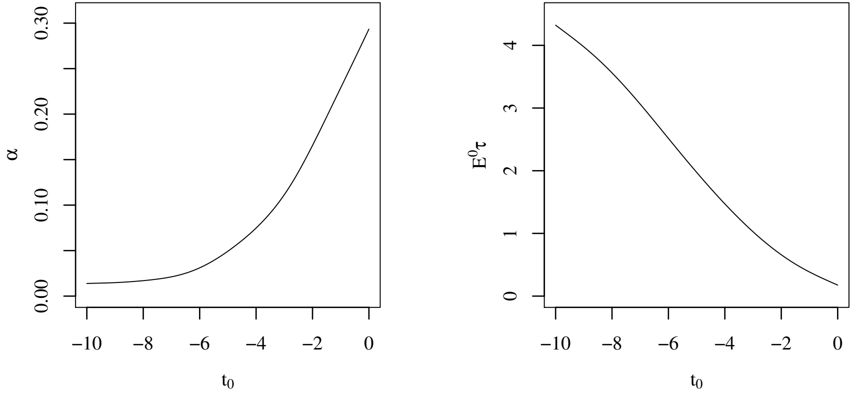
Numerical results

The optimal stopping boundary $a^*(t)$.



Left: the dependence of the probability α of a wrong decision on the value of t_0.

Right: the dependence of the maximal average observation time $\mathrm{E}^0\tau^*$ on t_0.



Thank you for your attention!