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.

2. Zhitlukhin M. V., Muravlev A. A., Shiryaev A. N.
1. Introduction

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

$$X_t = \mu t + B_t, \quad t \geq 0,$$

where $\mu$ is an unknown drift coefficient.

We consider the problem of sequentially testing the hypothesis

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

The two settings will be studied: the Chernoff problem (Bayesian) and the Kiefer–Weiss problem (minimax).
By definition, a decision rule is a pair $(\tau, d)$ consisting of
- a stopping time $\tau$ of the filtration $(\mathcal{F}_t)_{t \geq 0}$, $\mathcal{F}_t = \sigma(X_s; s \leq t)$
- an $\mathcal{F}_\tau$-measurable function $d$ taking values $\pm 1$

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

**Chernoff’s problem**

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

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

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

In other words, an optimal rule $(\tau^*, d^*)$ 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 $\mu$ is an unknown real parameter and let $\varepsilon > 0$, $\alpha \in (1/2, 1)$ be given numbers. Let $\Delta_{\alpha}$ be the class of decision rules $(\tau, d)$ such that

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

The parameter $\varepsilon$ specifies the indifference area: the decision “$\mu < 0$” is correct if $\mu \leq \varepsilon$, and “$\mu \geq 0$” is correct if $\mu \geq -\varepsilon$.

The parameter $\alpha$ specifies the maximal acceptable probability of error.

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

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

i.e. an optimal $(\tau^*, d^*)$ minimizes the maximal observation time over all decision rules in $\Delta_{\alpha}$. 

Known results and the aim of the research

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

$$\tau^* = \inf\{t \geq 0 : |X'_t| \geq a^*(t)\}, \quad d^* = \text{sgn}(X'_{\tau^*})$$

where $X'_t$ is a process obtained from $X_t$ by some transformation, and $a^*(t)$ is some function on $\mathbb{R}_+$ (independent of $\mu_0, \sigma_0$).

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

$$\tau^* = \inf\{t \geq 0 : |X_t| \geq b^*(t)\}, \quad d^* = \text{sgn}(X_{\tau^*})$$

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

3. Our aim is to find the boundaries $a^*(t)$ and $b^*(t)$.
Recall that the problem is to find \((\tau^*, d^*)\) such that
\[
E[c\tau^* + k|\mu|I(d^* \neq \text{sgn}(\mu))] = \inf_{(\tau, d)} E[c\tau + k|\mu|I(d \neq \text{sgn}(\mu))]
\]

Fix the parameters \((\mu_0, \sigma_0)\) and introduce the process \(W = (W_t)_{0 \leq t \leq 1}, \)
\[
W_t = \sigma_0(1 - t)X\frac{t}{\sigma^2_0(1-t)} - t\mu_0/\sigma_0.
\]
We check that \(W\) is a standard Brownian motion.

It turns out, \((\tau^*, d^*)\) can be found from the optimal stopping problem
\[
V_{\mu_0, \sigma_0} = \inf_{\tau \in \mathcal{M}^W_1} \mathbb{E} \left[ \frac{2}{\sigma^3_0(1 - \tau)} - |W_\tau + \mu_0/\sigma_0| \right],
\]
where \(\mathcal{M}^W_1\) is the class of all stopping times \(\tau \leq 1\) of the filtration \((\mathcal{F}^W_t)_{t \leq 1}, \mathcal{F}^W_t = \sigma(W_s; s \leq t).\)
Theorem

1) Let $\tau^*_W$ be an optimal stopping time in $V_{\mu_0,\sigma_0}$. Then an optimal decision rule $(\tau^*, d^*)$ for testing $H_+$ and $H_-$ is given by

$$\tau^* = \frac{\tau^*_W}{\sigma_0^2 (1 - \tau^*_W)}, \quad d^* = \text{sgn}(X_{\tau^*} + \mu_0/\sigma_0^2).$$

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

$$\tau^*_W(\mu_0, \sigma_0) = \inf\{0 \leq t \leq 1 : |W_t + \mu_0/\sigma_0| \geq a^*_0(t)\},$$

where $a^*_0(t) : [0, 1] \rightarrow \mathbb{R}_+$ 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

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

Chernoff and Breakwell showed that the optimal stopping time $\tau^*$ is given by

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

where $b^*(t) : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly positive function for $t > 0$, which does not depend on $\mu_0, \sigma_0$.

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

$$
b^*(t) = \sigma_0 t \cdot a^*_{\sigma_0} \left(1 - \frac{1}{\sigma_0^2 t}\right), \quad t \geq 1/\sigma_0^2.
$$
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 \((\tau, d)\) with \(E\tau < \infty\). For any such decision rule the average penalty is

\[
R(\tau, d) = E[\tau + E(\mu^- | \mathcal{F}_\tau)I\{d = +1\} + E(\mu^+ | \mathcal{F}_\tau)I\{d = -1\}].
\]

Thus the problem reduces to finding the stopping time \(\tau^*\) minimizing

\[
\mathcal{E}(\tau) = E[\tau + \min\{E(\mu^- | \mathcal{F}_\tau), E(\mu^+ | \mathcal{F}_\tau)\}]
\]

Using the Normal correlation theorem, we find

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

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 \( \tau \) with \( E\tau < \infty \)

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

Then we check that the process

\[
M_t = \frac{X_t + m_0/\sigma_0^2}{\sigma_0(t + 1/\sigma_0^2)} - \frac{m_0}{\sigma_0}
\]

is a martingale.

Applying the change of time, we find that the process

\[
W_t = M_{t/\sigma_0^2(1-t)}
\]

is a Brownian motion,

which reduces the Chernoff problem to the optimal stopping problem

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

It is solved using standard methods.
3. Solution of the Kiefer–Weiss problem

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

\[
\sup_{u \in \mathbb{R}} \mathbb{E}(\tau^* | \mu = u) = \inf_{(\tau, d) \in \Delta_\alpha} \sup_{u \in \mathbb{R}} \mathbb{E}(\tau | \mu = u),
\]

where \(\Delta_\alpha\) is the class of decision rules \((\tau, d)\) such that \(\mathbb{P}(d \neq \text{sgn}(u) | \mu = u) \leq \alpha\) for any \(|u| > \varepsilon\).

The problem is reduced to the family of problems \(V_c, c > 0\):

\[
V_c = \inf_{(\tau, d)} \left[ \mathbb{E}(\tau | \mu = 0) + c\{\mathbb{P}(d = -1 | \mu = \varepsilon) + \mathbb{P}(d = 1 | \mu = -\varepsilon)\} \right].
\]

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

\[
\mathbb{P}(d_c = -1 | \mu = \varepsilon) = \mathbb{P}(d_c = 1 | \mu = -\varepsilon) = \alpha,
\]

then \(\delta^c\) is optimal in the Kiefer–Weiss problem.
Using that \( d(P_t \mid \mu = u)/d(P_t \mid \mu = u) = \exp(u X_t - u^2 t/2) \), we obtain

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

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

\[
d_c = \text{sgn} \, X_{\tau_c}
\]

and \(\tau_c\) solves the optimal stopping problem

\[
V_c = \inf_{\tau} E[\tau + c e^{-\varepsilon |X_{\tau}| - \varepsilon^2 \tau/2} \mid \mu = 0].
\]

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

\[
P(X_{\tau_c} > 0 \mid \mu = -\varepsilon) = \alpha.
\]
Theorem

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

$\tau^* = \inf\{t \geq 0 : |X_t| \geq a^*(t + t_0)\}$, \hspace{1cm} d^* = \text{sgn} \ X_{\tau^*},$

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

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

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

$$0 < a(t) \leq \varepsilon e^{-\varepsilon^2 t/2}/2, \hspace{1cm} t \in \mathbb{R}.$$  

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

The optimal stopping boundary $a^*(t)$. 
Left: the dependence of the probability $\alpha$ of a wrong decision on the value of $t_0$.

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

\begin{align*}
\begin{array}{c}
\text{Left: } \alpha(t_0) = \frac{1}{\sqrt{1 + t_0^2}} \\
\text{Right: } E^0_{\tau^*}(t_0) = \frac{1}{1 + t_0^2}
\end{array}
\end{align*}
Thank you for your attention!