Sequential Detection of Changes: an overview

George V. Moustakides

Outline

- Sequential hypothesis testing and Sequential detection of changes
- The Sequential Probability Ratio Test (SPRT) for optimum hypothesis testing
- Performance criteria and optimum detection rules for sequential change detection
- Lorden's criterion and the CUSUM test
- Course-Layout

Sequential Hypothesis testing

In conventional (fixed sample size) binary hypothesis testing we are given a collection of observations $\xi_1,\ldots,\xi_t$ that satisfy one of the following two hypotheses

$$\mathbb{H}_0: \xi_1,\ldots,\xi_t \sim \mathbb{P}_0$$

$$\mathbb{H}_1: \xi_1,\ldots,\xi_t \sim \mathbb{P}_1$$

Using the available information (observations) we would like to decide which hypothesis is valid. This is achieved with the help of a

Decision rule $D(\xi_1,\ldots,\xi_t) \in \{0,1\}$

$$u_t = \log\left(\frac{d\mathbb{P}_1}{d\mathbb{P}_0}(\xi_1,\ldots,\xi_t)\right) \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtreqless}} t$$

In sequential hypothesis testing observations $\xi_1,\ldots,\xi_t,\ldots$ are supplied **sequentially**. With every new time instant we obtain a new observation. As before **all** observations satisfy one of the two hypotheses

$$\mathbb{H}_0\colon \xi_1,\ldots,\xi_t,\ldots \quad \sim \mathbb{P}_0$$

$$\mathbb{H}_1\colon \xi_1,\ldots,\xi_t,\ldots \quad \sim \mathbb{P}_1$$

Only now we like to reach a decision **as soon as possible.**

We first need a **Stopping rule** $T(\xi_1,\ldots,\xi_t)$ which at every time instant t consults the available information $\xi_1,\ldots,\xi_t$ and decides whether this information is *adequate to make a reliable decision between the two hypotheses or not*.

If YES, Stop and use **Decision rule** $D(\xi_1,\ldots,\xi_T)\in\{0,1\}$

If NO, ask for one more sample

We expect, in the average, to need (significantly) less samples to reach a decision as compared to the conventional fixed sample size method, for the same level of confidence.

The SPRT test (Wald 1947)

$$u_t = \log \left(\frac{d\mathbb{P}_1}{d\mathbb{P}_0}(\xi_1, \ldots, \xi_t) \right)$$

We define two thresholds $A < 0 < B$

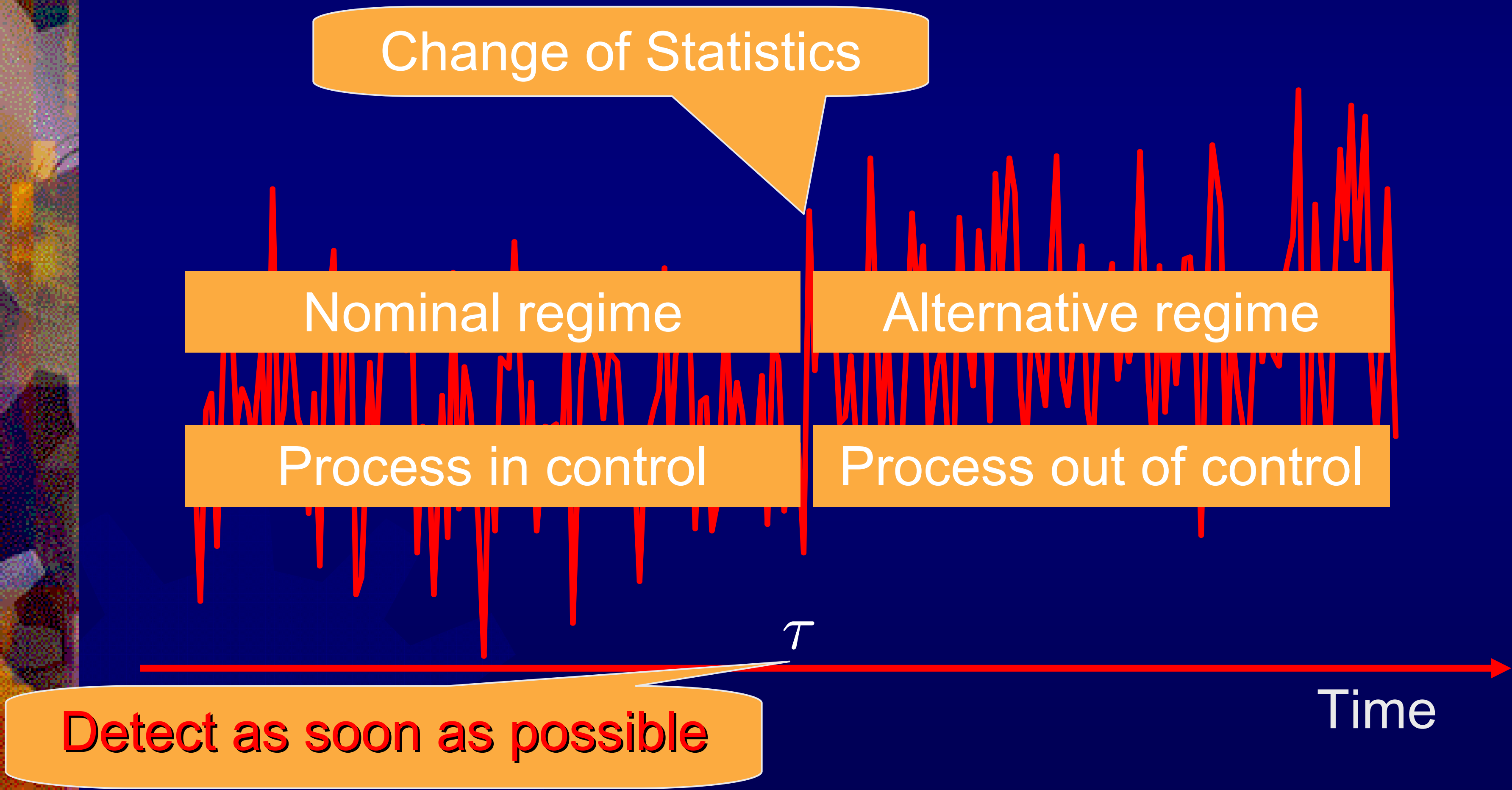
Stopping rule: $T = \inf_t \{t : u_t \notin (A, B)\}$

Decision rule: $D(\xi_1, \ldots, \xi_T) = \begin{cases} 1 & \text{if } u_T \geq B \\ 0 & \text{if } u_T \leq A \end{cases}$

Optimum for i.i.d. observations (Wald and Wolfowitz, 1948); for BM with constant drift (Shiryayev, 1967) and for homogeneous Poisson (Peskir, Shiryayev, 2000).

The Sequential change detection problem

Also known as the Disorder problem or the Change-Point problem or the Quickest Detection problem.



Applications

- Monitoring of quality of manufacturing process (1930's)
- Biomedical Engineering
- Electronic Communications
- Econometrics
- Seismology
- Speech & Image Processing
- Vibration monitoring
- Security monitoring (fraud detection)
- Spectrum monitoring
- Scene monitoring
- Network monitoring and diagnostics (router failures, intruder detection)
- Databases .....

Mathematical setup

We are observing sequentially a process $\{\xi_t\}$ with the following statistics:

$$\xi_t \sim \mathbb{P}_\infty \quad \text{for } 0 < t \leqslant \tau$$
$$\sim \mathbb{P}_0 \quad \text{for } \tau < t$$

Goal: Detect the change time τ **"as soon as possible"**

- Change time τ : deterministic (but unknown) or random
- Probability measures $\mathbb{P}_\infty$, $\mathbb{P}_0$: known

The observation process $\{\xi_t\}$ is available sequentially. This can be expressed through the filtration $\{\mathcal{F}_t\}$

$$\mathcal{F}_t = \sigma\{\xi_s : 0 < s \leqslant t\}.$$

- Interested in **sequential detection schemes**.
  At every time instant t we perform a test to decide whether to stop and declare an alarm or continue sampling. The test at time t must be based on the available information up to time t.
- Any sequential detection scheme can be represented by a **stopping time** T adapted to the filtration $\{\mathcal{F}_t\}$

Overview of existing results

$\mathbb{P}_\tau$: the probability measure induced, when the change takes place at time τ

$\mathbb{E}_\tau[.]$: the corresponding expectation

$\mathbb{P}_\infty$: all data under nominal regime

$\mathbb{P}_0$: all data under alternative regime

Optimality criteria

They must take into account two quantities:

- The detection delay T - τ
- The frequency of false alarms

Possible approaches: **Baysian and Min-max**

Baysian approach (Shiryayev 1978)

The change time τ is random with exponential prior.

$$\mathrm{Pro}[\tau = t] = (1\text{-}\varpi)\varpi^t$$

For any stopping time T define the criterion:

$$J(T) = c\mathbb{E}[\ (T - \tau)^+\ ] + \mathbb{P}[\ T \leqslant \tau\ ]$$

Optimization problem: $\inf_T J(T)$

Define the statistics: $\pi_t = \mathbb{P}[\ \tau \leqslant t \mid \mathcal{F}_t]$

Stopping rule: $T_S = \inf_t \{\ t\colon \pi_t \geqslant \nu\ \}$

- Discrete time: when $\{\xi_t\}$ is i.i.d. and there is a change in the pdf from $f_\infty(\xi)$ to $f_0(\xi)$.

$$\pi_t = \frac{\pi_{t-1} f_0(\xi_t)}{\pi_{t-1} f_0(\xi_t) + (1 - \pi_{t-1}) f_\infty(\xi_t)}$$

- Continuous time: when $\{\xi_t\}$ is a Brownian Motion and there is a change in the constant drift from μ_∞ to μ_0.

Min-max approach (Pollak, 1985)

The change time τ is deterministic but unknown.

For any stopping time T define the criterion:

$$J(T) = \sup_\tau \mathbb{E}_\tau[\ (T - \tau)^+ \mid T > \tau\ ]$$

Optimization problem: $\inf_T J(T)$;

subject to: $\mathbb{E}_\infty[\ T\ ] \geqslant \gamma$

Discrete time: when $\{\xi_t\}$ is i.i.d. and there is a change in the pdf from $f_\infty(\xi)$ to $f_0(\xi)$.

Compute the statistics: $S_t = (S_{t\text{-}1} + 1)\ \dfrac{f_0(\xi_n)}{f_\infty(\xi_n)}$.

Stopping rule: $T_\mathrm{P} = \inf_t \{\ t\colon S_t \geqslant \nu\ \}$ Mei (2006)

CUSUM test and Lorden's criterion

Page (1954) introduced CUSUM for i.i.d. observations.

Suppose we are given $\xi_1,\ldots,\xi_t$. Form a likelihood ratio test for the following two hypotheses:

$\mathbb{H}_0$: All observations are under the nominal regime

$\mathbb{H}_1$: There is a change at $\tau < t$

assuming τ known

$$\sup_{0\le\tau\le t} \sum_{n=\tau+1}^{t} \log\left(\frac{f_0(\xi_n)}{f_\infty(\xi_n)}\right) \ge \nu$$

$$\sum_{n=1}^{t} \log\left(\frac{f_0(\xi_n)}{f_\infty(\xi_n)}\right) - \inf_{0\le\tau\le t} \sum_{n=1}^{\tau} \log\left(\frac{f_0(\xi_n)}{f_\infty(\xi_n)}\right) \ge \nu$$

Define the CUSUM process y_t as follows:

$$y_t = u_t - m_t$$

where

$$u_t = \log\left(\frac{d\mathbb{P}_0}{d\mathbb{P}_\infty}(\mathcal{F}_t)\right)$$

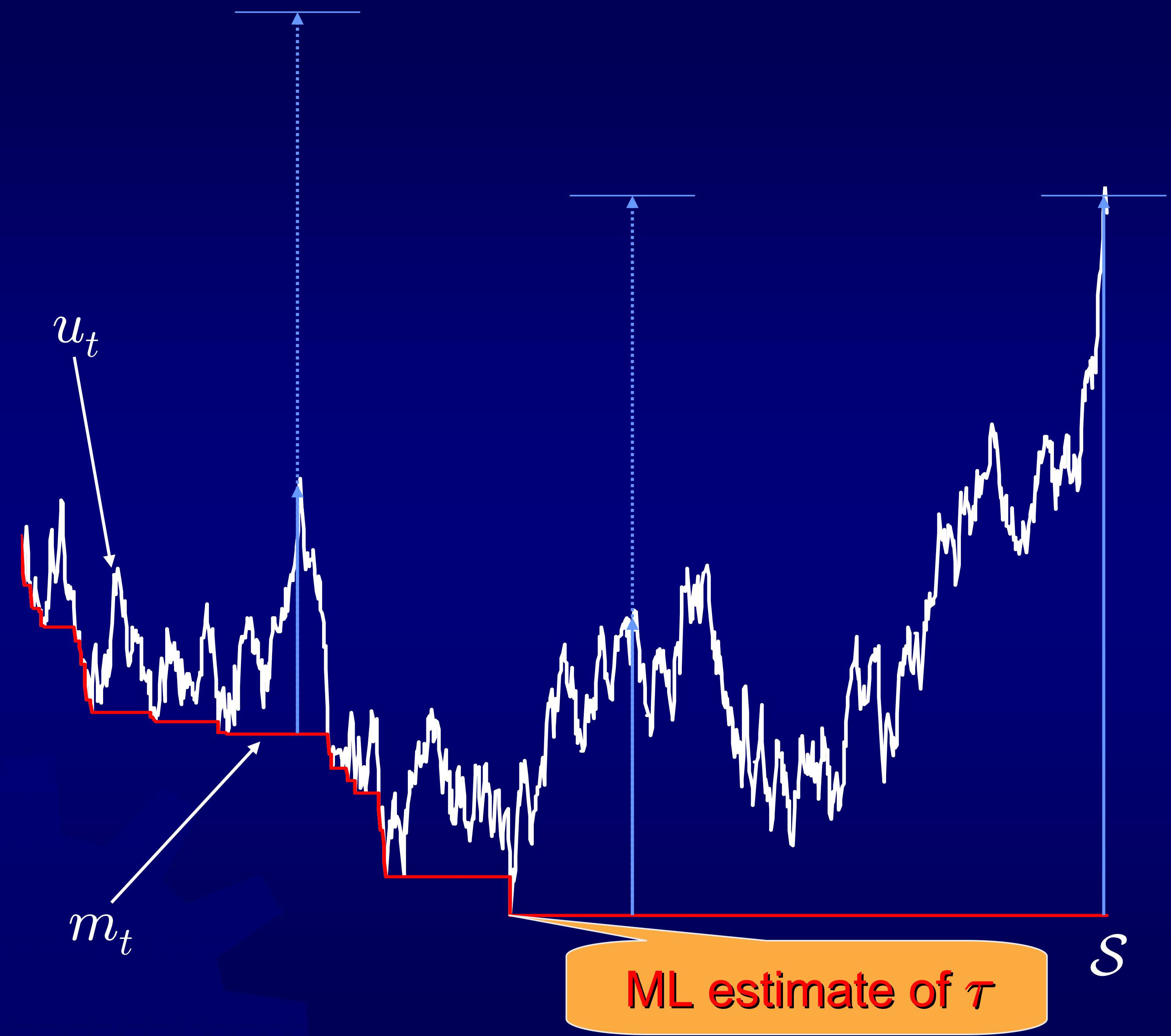
$$m_t = \inf_{0 \leqslant s \leqslant t} u_s \,.$$

The CUSUM stopping time:

$$\mathcal{S} = \inf_t \{ \, t \colon y_t \geqslant \nu \, \}$$

For the i.i.d. case we have a convenient recursion:

$$y_t = \left(y_{t-1} + \log\left(\frac{f_0(\xi_t)}{f_\infty(\xi_t)}\right) \right)^+$$


u_t
m_t
$\mathcal{S}$
ML estimate of τ

Min-max criterion (Lorden 1971):

$$J(T) = \sup_\tau \operatorname{essup} \mathbb{E}_\tau[\ (T - \tau)^+ \mid \mathcal{F}_\tau\ ]$$

Optimization problem: $\inf_T J(T)$;

subject to: $\mathbb{E}_\infty[\ T\ ] \geqslant \gamma$.

- Lorden (1971) proved asymptotic optimality in discrete time for i.i.d. before and after the change for $\gamma \to \infty$.
- Discrete time: when $\{\xi_t\}$ is i.i.d. before and after the change (Moustakides 1986, Ritov 1990).
- Continuous time: when $\{\xi_t\}$ is a Brownian Motion with constant drift before and after the change (Shiryayev 1996, Beibel 1996).

Course-layout

- Elements of optimal stopping (M 05/21)
- The SPRT test for i.i.d. and BM (W 05/23)

Sequential Detection of Changes

- Change-time models and comparison of criteria (M 05/28)
- CUSUM for BM (W 05/30)
- Shiryayev for BM and CUSUM for Ito process (M 06/04)
- CUSUM for Poisson (W 06/06)
- CUSUM for i.i.d. (M 06/11)
- Alternative Lorden-like criteria (W 06/13)&(M 06/18)
- Two sided CUSUM tests (W 06/20)