Change-time models
and
Performance criteria
for
Sequential change detection

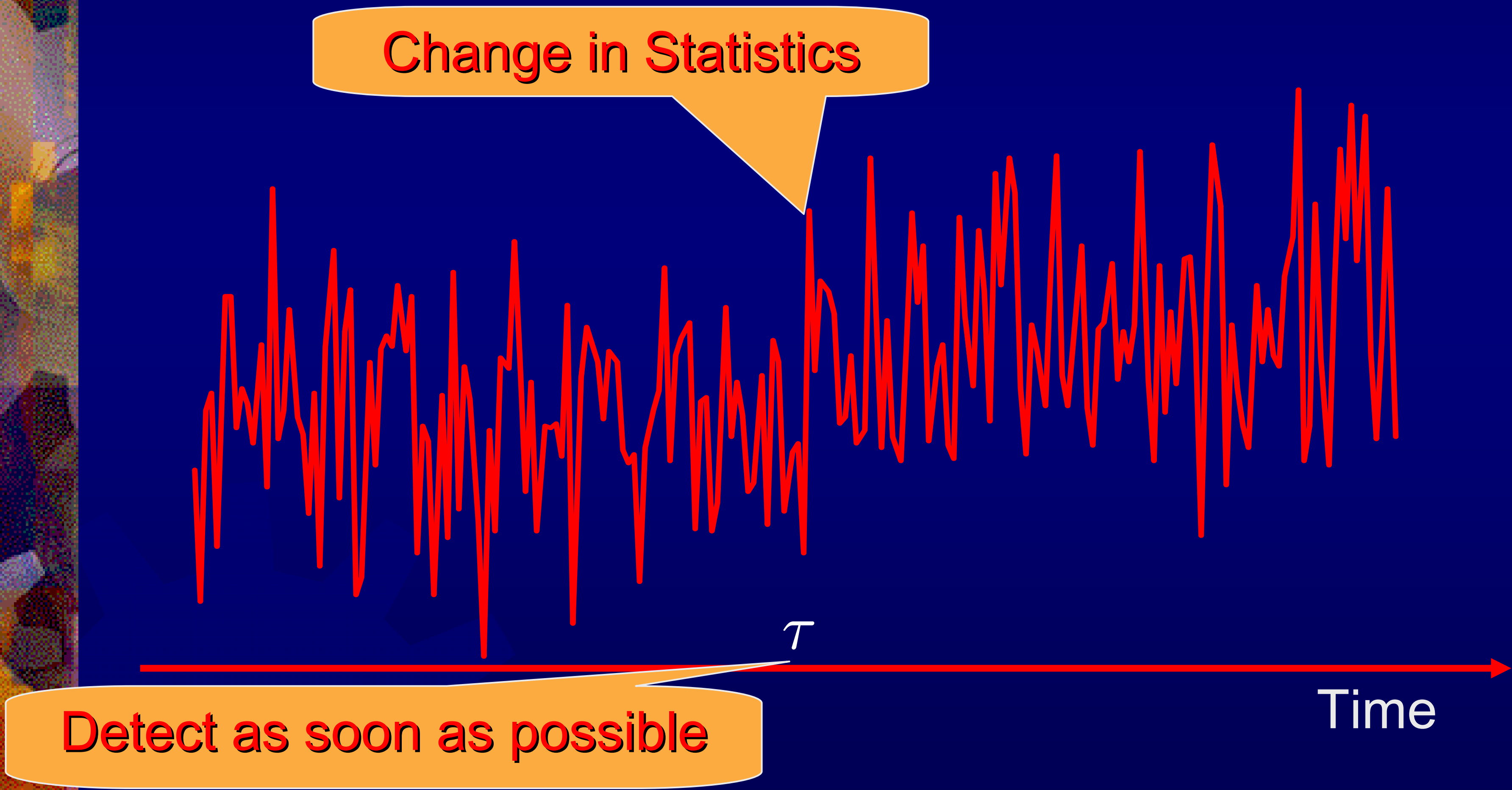
George V. Moustakides,
University of Patras, Greece

Outline

- Sequential change detection: existing criteria and optimum tests
- Models for the probability measure induced by the change
- Models for the change-time mechanism
- A general performance measure and the associated optimization problem
- New performance measures; for a particular we offer the optimum test.

Sequential Change Detection

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:

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

- 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. We have a time increasing information expressed with the help of the filtration $\{\mathcal{F}_t\}$ where

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

At each time t we would like to consult the available observations and infer whether a change took place or not.

A sequential detection scheme can therefore be regarded as a **Stopping Time** T adapted to $\{\mathcal{F}_t\}$

$\mathbb{P}_\tau$: 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 must take into account:

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

Baysian approach (Shiryayev 1961)

The change time τ is random with **exponential prior**.

For any stopping time T define the criterion:

$$J_{\mathrm{S}}(T) = \mathbb{E}[\ (T - \tau)^{+} \mid T > \tau\ ]$$

Optimization problem: $\inf_T J_{\mathrm{S}}(T)$

subject to: $\mathbb{P}[\ T \leqslant \tau\ ] \leqslant \alpha$

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

and stop: $T_{\mathrm{S}} = \inf_t \{\ t\colon \pi_t \geqslant \nu\ \}$

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

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

Optimization problem: $\inf_{T} J_{\mathrm{P}}(T)$;

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

<u>Discrete time</u>: 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: $R_t = (R_{t\text{-}1} + 1)\ \dfrac{f_0(\xi_t)}{f_{\infty}(\xi_t)}$.

and stop (Pollak 1985): $T_{\mathrm{SRP}} = \inf_t \{\ t\text{:}\ R_t \geqslant \nu\ \}$

Lorden's criterion (1971)

The change time τ is deterministic and unknown.
For any stopping time T define the criterion:

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

Optimization problem: $\inf_T J_{\mathrm{L}}(T)$;
subject to: $\mathbb{E}_\infty[\ T\ ] \geqslant \gamma$.

The test closely related to Lorden's criterion and being the most popular test for the change detection problem in practice, is the **Cumulative Sum** (CUSUM) test.

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

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

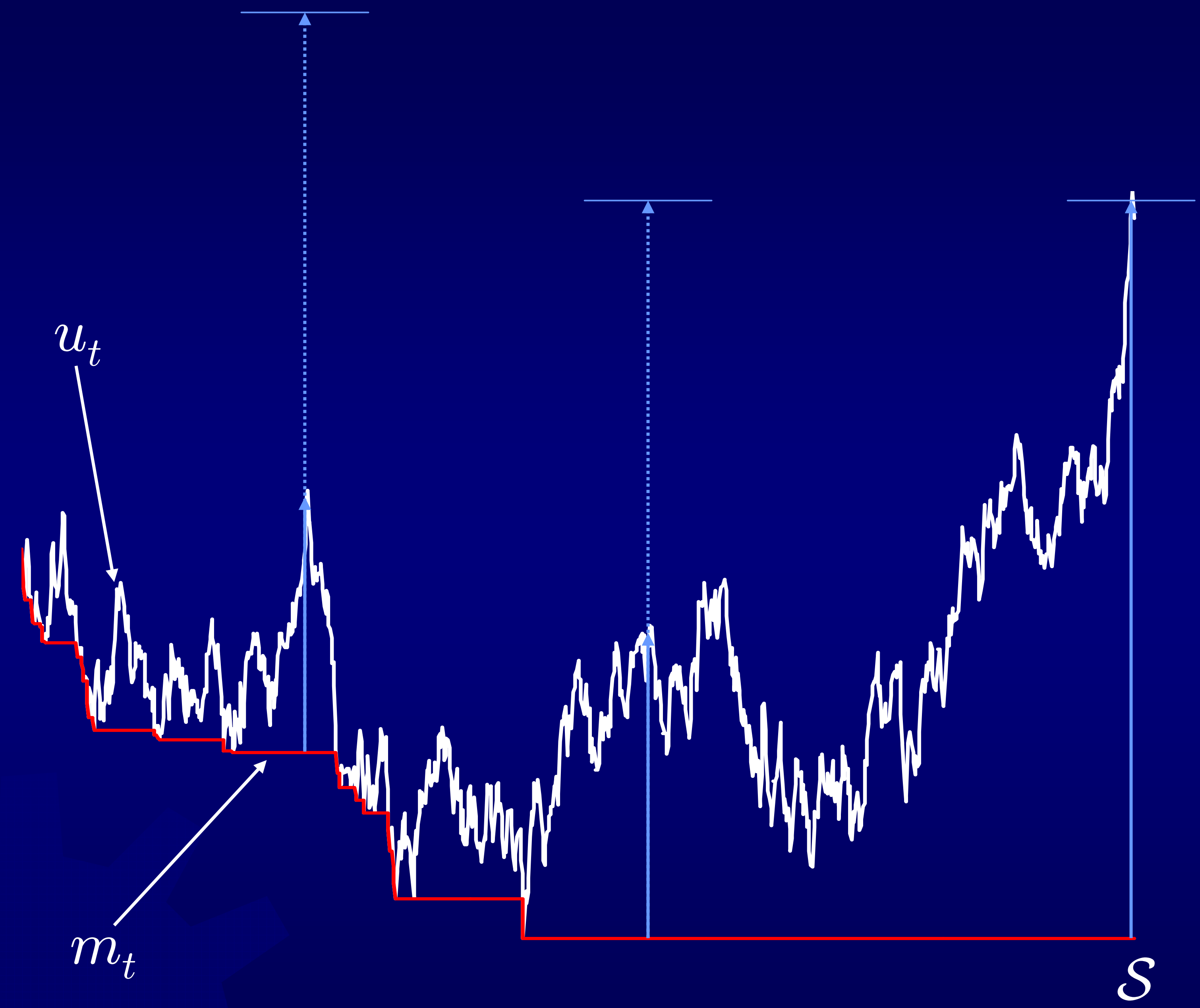
Define the CUSUM process y_t as follows:

$$y_t = u_t - m_t$$

The CUSUM stopping time (Page 1954):

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

- Discrete time: when $\{\xi_t\}$ is i.i.d. before and after the change (Lorden 1971, 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). Ito process (Moustakides 2004).


u_t
m_t
S

$$J_{\mathrm{S}}(T) = \mathbb{E}[\ (T - \tau)^{+} \mid T > \tau\ ]$$

$$\wedge\!\!\backslash$$

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

$$\wedge\!\!\backslash$$

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

Overly Pessimistic ???

Models for probability measures

Assume change at given time τ.

$$\mathbb{P}_\infty: \quad f_\infty(\xi_1,\ldots,\xi_\tau,\xi_{\tau+1},\ldots) = f_\infty(\xi_1,\ldots,\xi_\tau) f_\infty(\xi_{\tau+1},\ldots \mid \xi_1,\ldots,\xi_\tau)$$

$$\mathbb{P}_0: \quad f_0(\xi_1,\ldots,\xi_\tau,\xi_{\tau+1},\ldots) = f_0(\xi_1,\ldots,\xi_\tau) f_0(\xi_{\tau+1},\ldots \mid \xi_1,\ldots,\xi_\tau)$$

$$\mathbb{P}_\tau: \quad f_\infty(\xi_1,\ldots,\xi_\tau) f_0(\xi_{\tau+1},\ldots \mid \xi_1,\ldots,\xi_\tau)$$

The data before the change influence the data after the change

Consider X any $\mathcal{F}_\infty$ - measurable rv, then

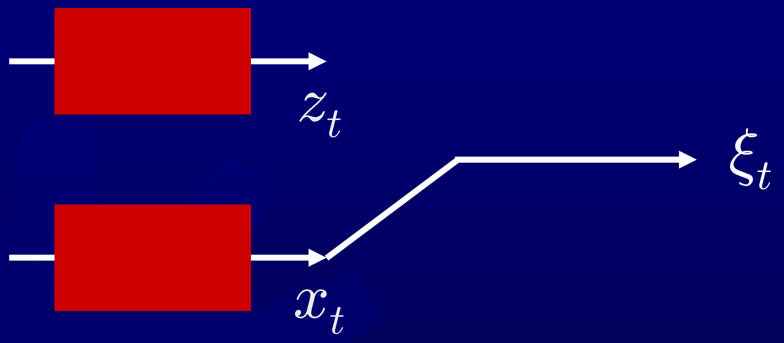
$$\mathbb{E}_\tau[X] = \mathbb{E}_\infty\Big[\mathbb{E}_0[X \mid \mathcal{F}_\tau]\Big]$$

$$\mathbb{P}_\infty: \quad f_\infty(\xi_1, \ldots, \xi_\tau, \xi_{\tau+1}, \ldots)$$
$$= f_\infty(\xi_1, \ldots, \xi_\tau) f_\infty(\xi_{\tau+1}, \ldots \mid \xi_1, \ldots, \xi_\tau)$$

$$\mathbb{P}_0: \quad f_0(\xi_1, \ldots, \xi_\tau, \xi_{\tau+1}, \ldots)$$
$$= f_0(\xi_1, \ldots, \xi_\tau | \xi_{\tau+1}, \ldots) f_0(\xi_{\tau+1}, \ldots)$$

$$\mathbb{P}_\tau: \quad f_\infty(\xi_1, \ldots, \xi_\tau) f_0(\xi_{\tau+1}, \ldots)$$

The two parts are independent



Next we consider only the previous model!!

Change-time mechanisms

$$X_t = AX_{t-1} + BW_t; \qquad t \leq \tau$$

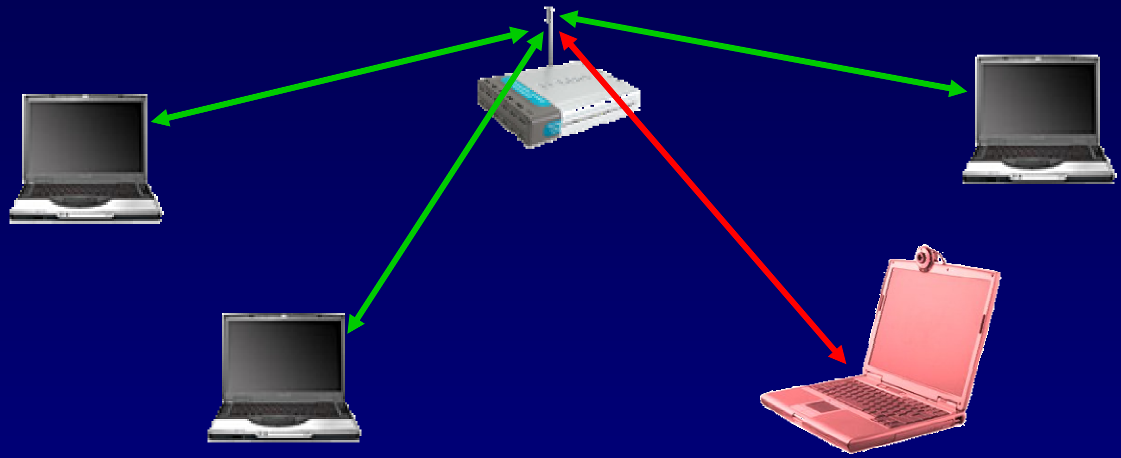
$$X_t = \mathcal{A}X_{t-1} + BW_t; \qquad t > \tau$$

There is a change-time mechanism that consults the observation history in order to impose the change

How can a change happen?

In the case of vibration monitoring in structures we have a change in the structure when the amplitude of the vibrations is very large for a period of time T.

A change can be imposed by an external independent factor (bomb) **not being related to the observations.**



 Legitimate User (follows the wireless protocol)

 Illegitimate User (DOES NOT follow the wireless protocol).

The illegitimate user can switch to the illegitimate mode f.e. after a fixed time interval T.

Or by observing the traffic he can look for the **right moment** to make the switching.

The observation process $\{\xi_t\}$ is available sequentially. We have a time increasing information expressed with the help of the filtration $\{\mathcal{F}_t\}$ where

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

We model the change-time mechanism as a **random time** τ for which the probability to impose the change at any specific time t depends on the observed history up to time t.

$$\{\pi_t\} : \quad \pi_t = \mathbb{P}[\tau = t | \mathcal{F}_t]$$

If $\{\mathcal{X}_t\}$ is a sequence of nonnegative, $\mathcal{F}_\infty$ –measurable rv. We are interested in finding $\mathbb{E}_\tau[\mathcal{X}_\tau]$

$$\mathbb{E}_\tau[\mathcal{X}_\tau] = \sum_{t=1}^{\infty} \mathbb{E}_t[\mathcal{X}_t \pi_t] = \sum_{t=1}^{\infty} \mathbb{E}_\infty[\mathbb{E}_0[\mathcal{X}_t | \mathcal{F}_t] \pi_t]$$

A general performance measure

If T is an $\{\mathcal{F}_t\}$–adapted stopping time used to detect the change. Then we define the following performance measure

$$J(T) = \mathbb{E}_\tau[(T-\tau)^+ | T > \tau] = \frac{\mathbb{E}_\tau[(T-\tau)^+]}{\mathbb{P}_\tau[T > \tau]}$$

$$J(T) = \frac{\sum_{t=0}^{\infty} \mathbb{E}_\infty[\pi_t \mathbb{E}_0[(T-t)^+ | \mathcal{F}_t]]}{\sum_{t=0}^{\infty} \mathbb{E}_\infty[\pi_t 1_{\{T>t\}}]}$$

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

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

In principle it can be solved **IF** $\{\pi_t\}$ is known.

When $\{\pi_t\}$ is **not** known, or **partially** known then this defines an uncertainty class $\mathcal{T}$ for the change-time τ. The performance measure, by following a **worst-case** approach, can then be modified to

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

and the optimization problem for finding the optimum stopping time T is transformed into a min-max

$$\inf_T \sup_{\tau \in \mathcal{T}} \mathbb{E}_\tau[(T-\tau)^+ | T > \tau]$$

subject to

$$\mathbb{E}_\infty[T] \geq \gamma$$

Special cases and uncertainty classes

Let us decompose π_t as follows:

$$\pi_t = \varpi_t \ p_t$$

where

$$\varpi_t = \mathbb{E}_\infty[\pi_t]; \qquad p_t = \frac{\pi_t}{\varpi_t}$$

Deterministic

$\{\mathcal{F}_t\}$–adapted $\mathbb{E}_\infty[p_t] = 1$

Quantity ϖ_t expresses the total probability that we will have a change at time t whereas p_t how this probability is distributed among the events that can occur up to time t.

$$J(T) = \frac{\sum_{t=0}^{\infty} \varpi_t \mathbb{E}_\infty[p_t \mathbb{E}_0[(T-t)^+ | \mathcal{F}_t]]}{\sum_{t=0}^{\infty} \varpi_t \mathbb{E}_\infty[p_t 1_{\{T>t\}}]}$$

Case $p_t = 1$ (no dependence on the history)

$$J(T) = \frac{\sum_{t=0}^{\infty} \varpi_t \mathbb{E}_t[(T-t)^+]}{\sum_{t=0}^{\infty} \varpi_t \mathbb{P}_\infty[T > t]}$$

For ϖ_t exponential we have Shiryayev's measure.

If ϖ_t unknown and we follow a worst case scenario:

$$\begin{aligned}
\mathcal{J}(T) &= \sup_{\{\varpi_t\}} \frac{\sum_{t=0}^{\infty} \varpi_t \mathbb{E}_t[(T-t)^+]}{\sum_{t=0}^{\infty} \varpi_t \mathbb{P}_\infty[T > t]} \\
&= \sup_t \frac{\mathbb{E}_t[(T-t)^+]}{\mathbb{P}_\infty[T > t]} \\
&= \sup_t \mathbb{E}_t[(T-t)^+ | T > t]
\end{aligned}$$

We recover Pollak's measure.

Case $p_t \neq 1$ (dependence on the history)

$$
\begin{aligned}
\mathcal{J}(T) &= \sup_{\{p_t\}} \sup_{\{\varpi_t\}} \frac{\sum_{t=0}^{\infty} \varpi_t \mathbb{E}_\infty[p_t \mathbb{E}_0[(T-t)^+ | \mathcal{F}_t]]}{\sum_{t=0}^{\infty} \varpi_t \mathbb{E}_\infty[p_t 1_{\{T>t\}}]} \\
&= \sup_{\{p_t\}} \sup_{t} \frac{\mathbb{E}_\infty[p_t \mathbb{E}_0[(T-t)^+ | \mathcal{F}_t]]}{\mathbb{E}_\infty[p_t 1_{\{T>t\}}]} \\
&= \sup_{t} \sup_{p_t} \frac{\mathbb{E}_\infty[p_t \mathbb{E}_0[(T-t)^+ | \mathcal{F}_t]]}{\mathbb{E}_\infty[p_t 1_{\{T>t\}}]} \\
&= \sup_{t} \operatorname{essup} \mathbb{E}_0[(T-t)^+ | \mathcal{F}_t] \\
&= \sup_{t} \operatorname{essup} \mathbb{E}_t[(T-t)^+ | \mathcal{F}_t]
\end{aligned}
$$

We recover Lorden’s measure.

For Lorden’s measure **WE ASSUME DEPENDENCE on the HISTORY!!!!**

New performance measures

In addition to the observation process $\{\xi_t\}$ we **observe** a sequence of **random** times $\{\tau_n\}$ with $\tau_n \to \infty$ a.s. We assume that for some n

$$\begin{aligned} \xi_t \quad &\sim \quad \mathbb{P}_\infty \qquad \text{for } 0 \leqslant t \leqslant \tau_n \\ &\sim \quad \mathbb{P}_0 \qquad \text{for } \tau_n < t \end{aligned}$$

Earthquake damage detection in structures.

Define $\quad \mathcal{N}_t = \sup_n \{\tau_n \leqslant t\}$

$$\mathcal{F}_t = \sigma\{\xi_s \,, \mathcal{N}_s : 0 \leqslant s \leqslant t\}$$

then τ_n becomes a s.t. adapted to $\{\mathcal{F}_t\}$

$$J(T) = \sup_{n \geqslant 0} \text{ essup } \mathbb{E}_{\tau_n}[\,(T - \tau_n)^+ \mid \mathcal{F}_{\tau_n}]$$

Optimization problem: $\inf_T J(T)$

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

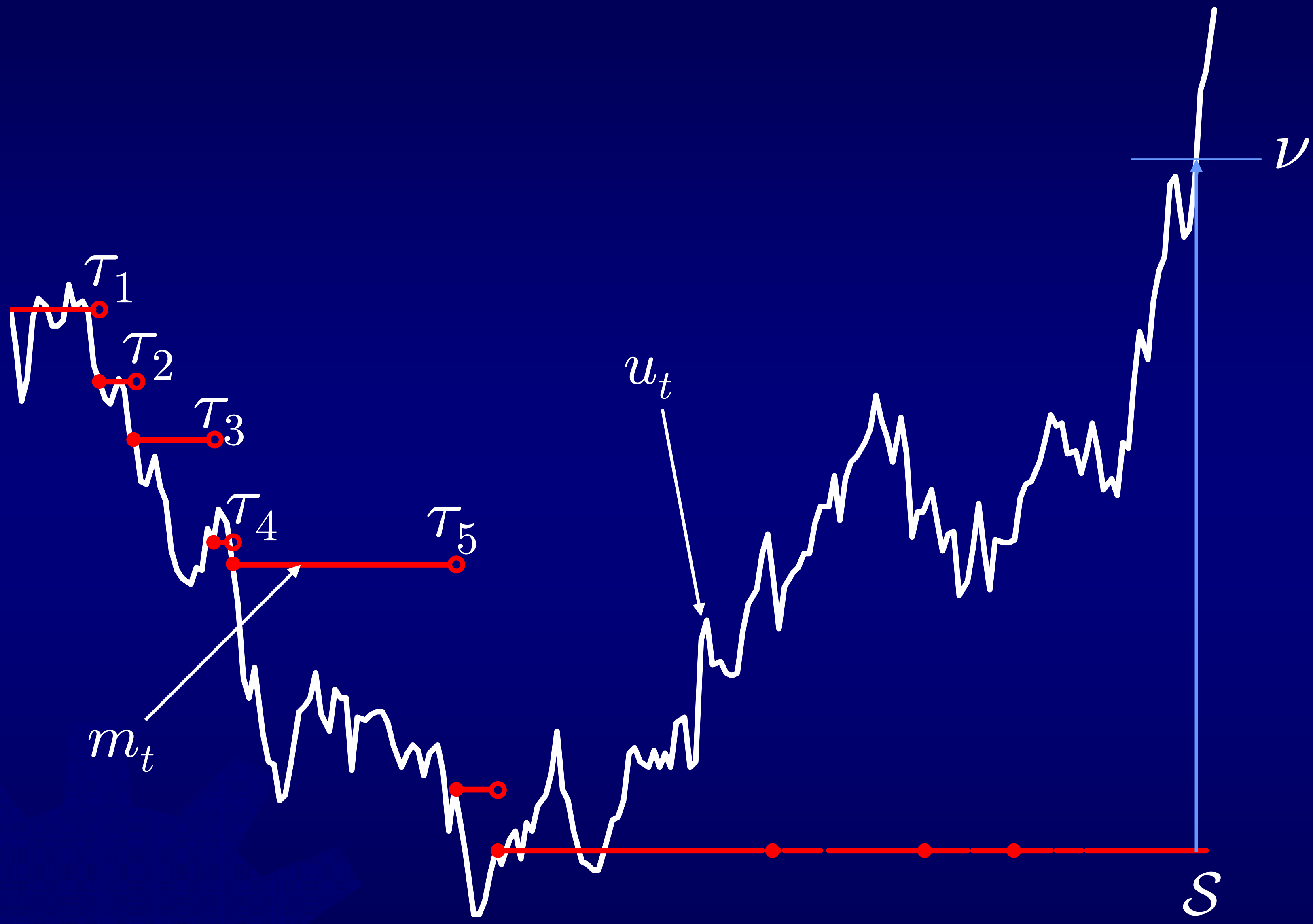
Extended CUSUM (ECUSUM)

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

$$m_t = \inf_{1 \leq n \leq \mathcal{N}_t} u_{\tau_n}$$

$$y_t = u_t - m_t$$

$$\mathcal{S} = \inf_{t \geq 0}\{t : \; y_t \geq \nu\}$$

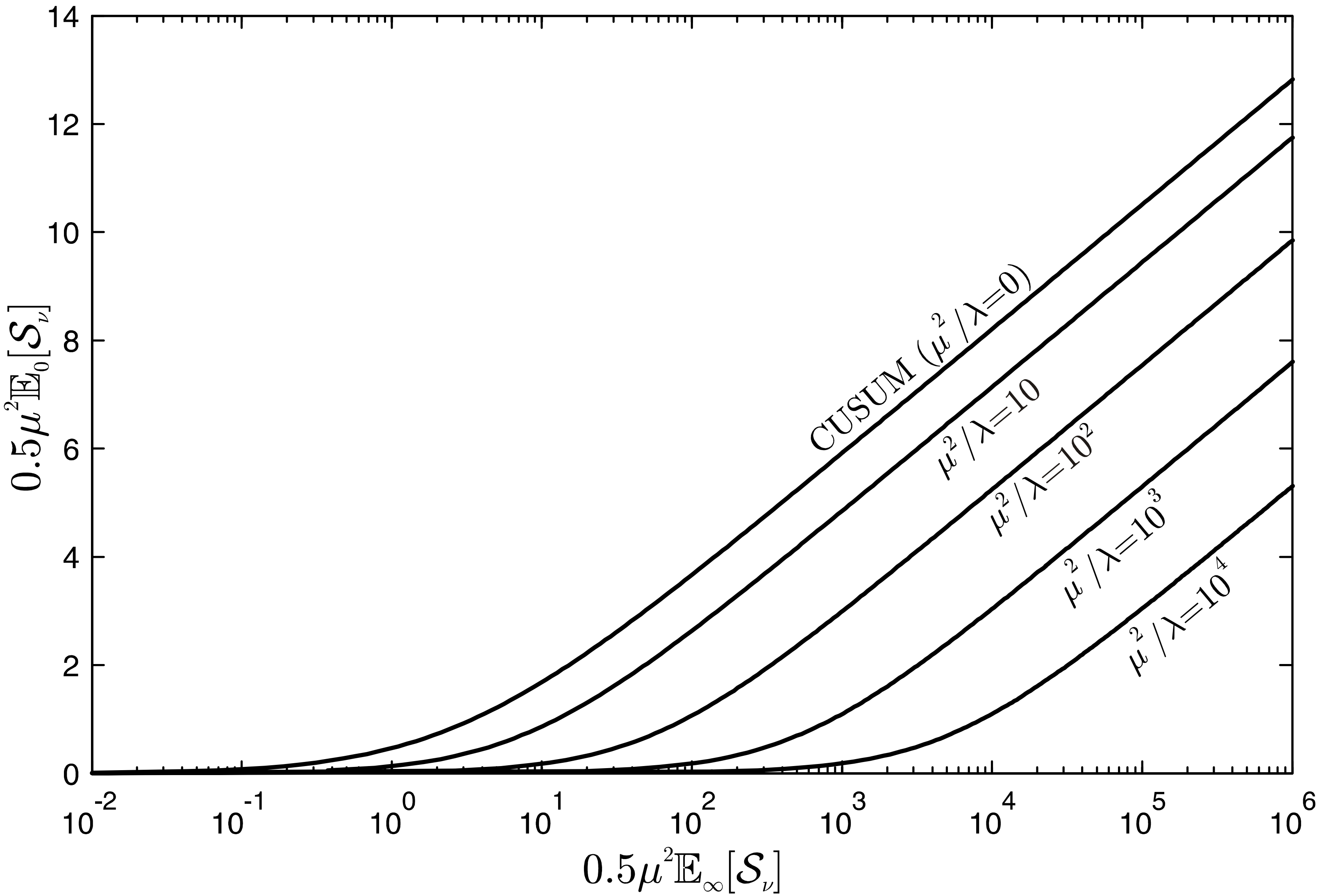


τ_1
τ_2
τ_3
τ_4
τ_5
u_t
m_t
ν
$\mathcal{S}$

Optimality when:

$\{\xi_t\}$ is BM with constant drift before and after the change and $\{\tau_n\}$ is Poisson.

$\{\xi_t\}$ is i.i.d. before and after the change and $\{\tau_n\}$ has differences $(\tau_n - \tau_{n-1})$ exponentially distributed.


14
12
10
8
6
4
2
0
$0.5\mu^2\mathbb{E}_0[\mathcal{S}_\nu]$
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
10^{6}
$0.5\mu^2\mathbb{E}_\infty[\mathcal{S}_\nu]$
CUSUM (μ^2/λ=0)
μ^2/λ=10
μ^2/λ=10^2
μ^2/λ=10^3
μ^2/λ=10^4

EnD

Thank you for your attention!!