Optimum CUSUM Tests

for

Detecting Changes in Continuous Time Processes

George V. Moustakides
INRIA, Rennes, France

Outline

- The change detection problem
- Overview of existing results
- Lorden’s criterion and the CUSUM test
- A modified Lorden criterion
- CUSUM tests for Ito processes
- Extensions

The change detection problem

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

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

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

Applications include: systems monitoring; quality control; financial decision making; remote sensing (radar, sonar, seismology); speech/image/video segmentation; …

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

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

- 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 $\mathcal{F}_t$ (the time we stop and declare an alarm).

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.

For any stopping time T define the criterion:

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

Optimization problem: $\inf_T J(T)$

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

and stop: $T_S = \inf_t \{\ t: \pi_t \geqslant \nu\ \}$

- 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 (Shiryayev-Roberts-Pollak)

The change time τ is deterministic but unknown.

For any stopping time T define the criterion:

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

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

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

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

Compute the statistics: $S_n = (S_{n-1} + 1)\,\dfrac{f_0(\xi_n)}{f_\infty(\xi_n)}$.

and stop (Yakir 1997): $T_{SRP} = \inf_n \{\ n:\ S_n \geqslant \nu\ \}$

Lorden's criterion and the CUSUM test

Alternative min-max approach (Lorden 1971):

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

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

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

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

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.

Define the CUSUM process y_t as follows:

$$y_t = u_t - m_t$$

where

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

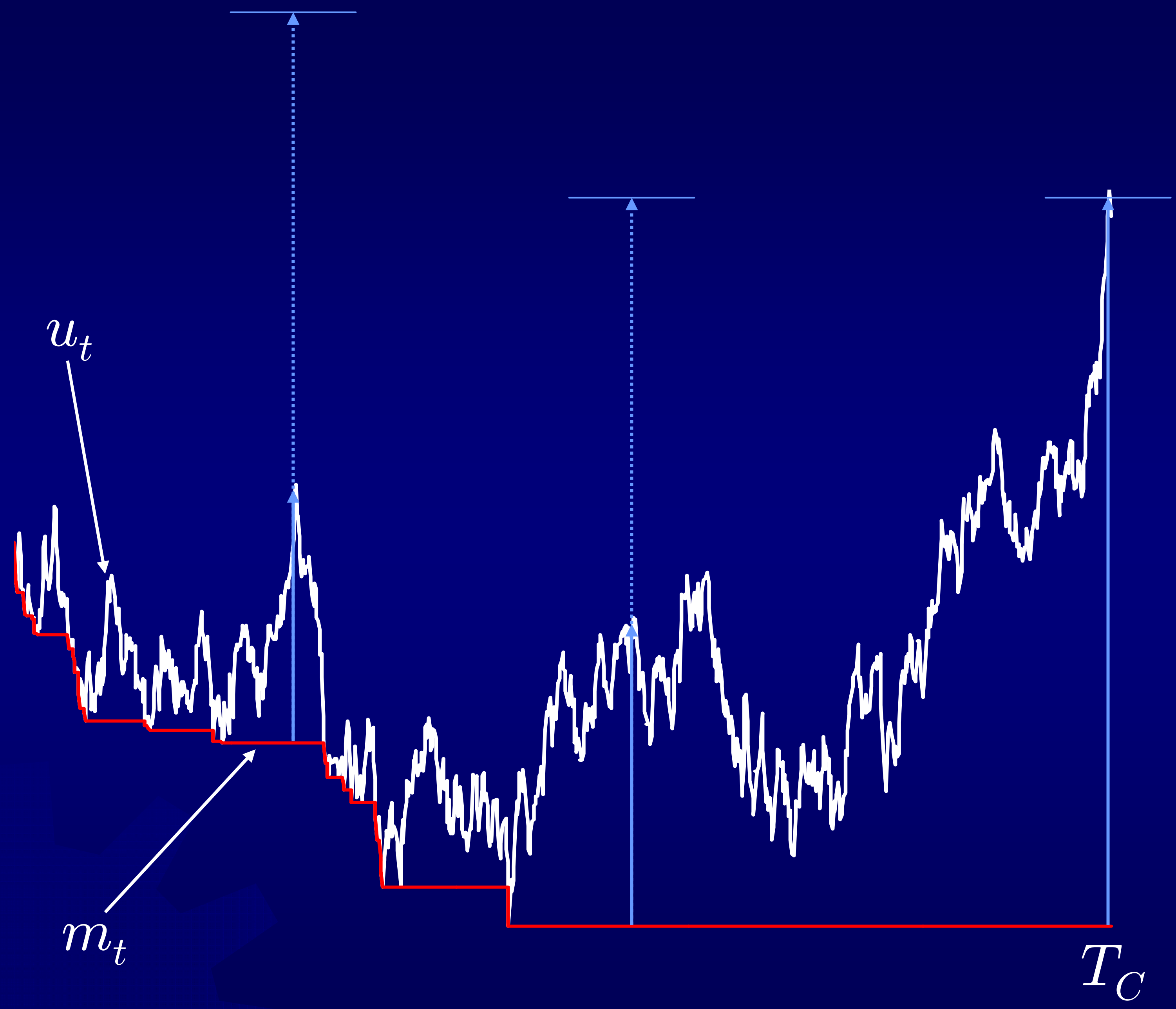
$$m_t = \inf_{0\leqslant s\leqslant t} u_s \,.$$

The CUSUM stopping time (Page 1954):

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

- Discrete time: when $\{\xi_n\}$ 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).


u_t
m_t
T_C

A modified Lorden criterion

We intend to extend the optimality of CUSUM to detection of changes in **Ito processes** by modifying Lorden's criterion using the **Kullback-Leibler Divergence (KLD)**.

Similar extension exists for the Sequential Probability Ratio Test (SPRT), applied in hypotheses testing, since 1978 (Liptser and Shiryayev)

The observation process $\{\xi_t\}$ satisfies the following sde:

$$d\xi_t = \begin{cases} dw_t & 0 \leqslant t \leqslant \tau \\ \alpha_t\, dt + dw_t & \tau < t \end{cases}$$

$\{w_t\}$ standard Brownian Motion
$\{\alpha_t\}$ adapted to the history $\mathcal{F}_t = \sigma\{\xi_s : 0 \leqslant s \leqslant t\}$

If $\alpha_t = \alpha(\xi_t)$, then ξ_t is a diffusion process for $t > \tau$.

To $\{\xi_t\}$ we correspond the following process $\{u_t\}$

$$du_t = \alpha_t \, d\xi_t - 0.5\alpha_t^2 \, dt \, .$$

We would like: $u_t = \log\Big(\frac{d\mathbb{P}_0}{d\mathbb{P}_\infty}(\mathcal{F}_t)\Big)$.

We need the following conditions:

1. $\mathbb{P}_0\Big[\int_0^t \alpha_s^2 \, ds < \infty\Big] = \mathbb{P}_\infty\Big[\int_0^t \alpha_s^2 \, ds < \infty\Big] = 1$

2. A “Novikov” condition

3. $\mathbb{P}_0\Big[\int_0^\infty \alpha_s^2 \, ds = \infty\Big] = \mathbb{P}_\infty\Big[\int_0^\infty \alpha_s^2 \, ds = \infty\Big] = 1$

From Condition 1&2 we have validity of Girsanov's theorem:

$$\frac{d\mathbb{P}_0}{d\mathbb{P}_\infty}(\mathcal{F}_t) = e^{u_t} \qquad \frac{d\mathbb{P}_\tau}{d\mathbb{P}_\infty}(\mathcal{F}_t) = e^{u_t - u_\tau}$$

The Kullback-Leibler Divergence can then be written as:

$$\mathbb{E}_\tau\Big[\log\Big(\frac{d\mathbb{P}_\tau}{d\mathbb{P}_\infty}(\mathcal{F}_t)\Big)\,\Big|\,\mathcal{F}_\tau\Big]$$

$$= \mathbb{E}_\tau\Big[\int_\tau^t \alpha_s\,dw_s + 0.5\int_\tau^t \alpha_s^2\,ds\,\Big|\,\mathcal{F}_\tau\Big]$$

$$= \mathbb{E}_\tau\Big[0.5\int_\tau^t \alpha_s^2\,ds\,\Big|\,\mathcal{F}_\tau\Big], \quad 0 \leqslant \tau \leqslant t$$

The original Lorden criterion

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

using the Kullback-Leibler Divergence can be modified as

$$J(T) = \sup_{\tau} \operatorname{essup} \mathbb{E}_{\tau}\left[\mathbb{1}_{\{T > \tau\}} 0.5 \int_{\tau}^{T} \alpha_t^2 \, dt \,\middle|\, \mathcal{F}_{\tau}\ \right]$$

The two criteria are equivalent in the case

$$\alpha_t^2 = constant$$

i.e. Brownian motion with constant drift.

Similarly

$$\mathbb{E}_\infty\Big[\log\Big(\frac{d\mathbb{P}_\infty}{d\mathbb{P}_0}(\mathcal{F}_t)\Big)\Big]$$

$$= \mathbb{E}_\infty\Big[- \int_0^t \alpha_s \, dw_s \; + 0.5 \int_0^t \alpha_s^2 \, ds \Big]$$

$$= \mathbb{E}_\infty\Big[0.5 \int_0^t \alpha_s^2 \, ds \Big]$$

This suggest replacing the constraint $\mathbb{E}_\infty[\, T \,] \geqslant \gamma$ with

$$\mathbb{E}_\infty\Big[0.5 \int_0^T \alpha_t^2 \, dt \Big] \geqslant \gamma$$

Summarizing:

$$J(T) = \sup_{\tau} \operatorname{essup} \mathbb{E}_{\tau}\Big[\mathbb{1}_{\{T > \tau\}} 0.5 \int_{\tau}^{T} \alpha_t^2 \, dt \;\Big|\; \mathcal{F}_{\tau}\Big]$$

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

subject to: $\mathbb{E}_{\infty}\Big[0.5 \int_0^T \alpha_t^2 \, dt\Big] \geqslant \gamma$

CUSUM tests for Ito processes

The CUSUM statistics y_t for Ito processes takes the form

$$du_t = \alpha_t \, d\xi_t - 0.5\alpha_t^2 \, dt$$

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

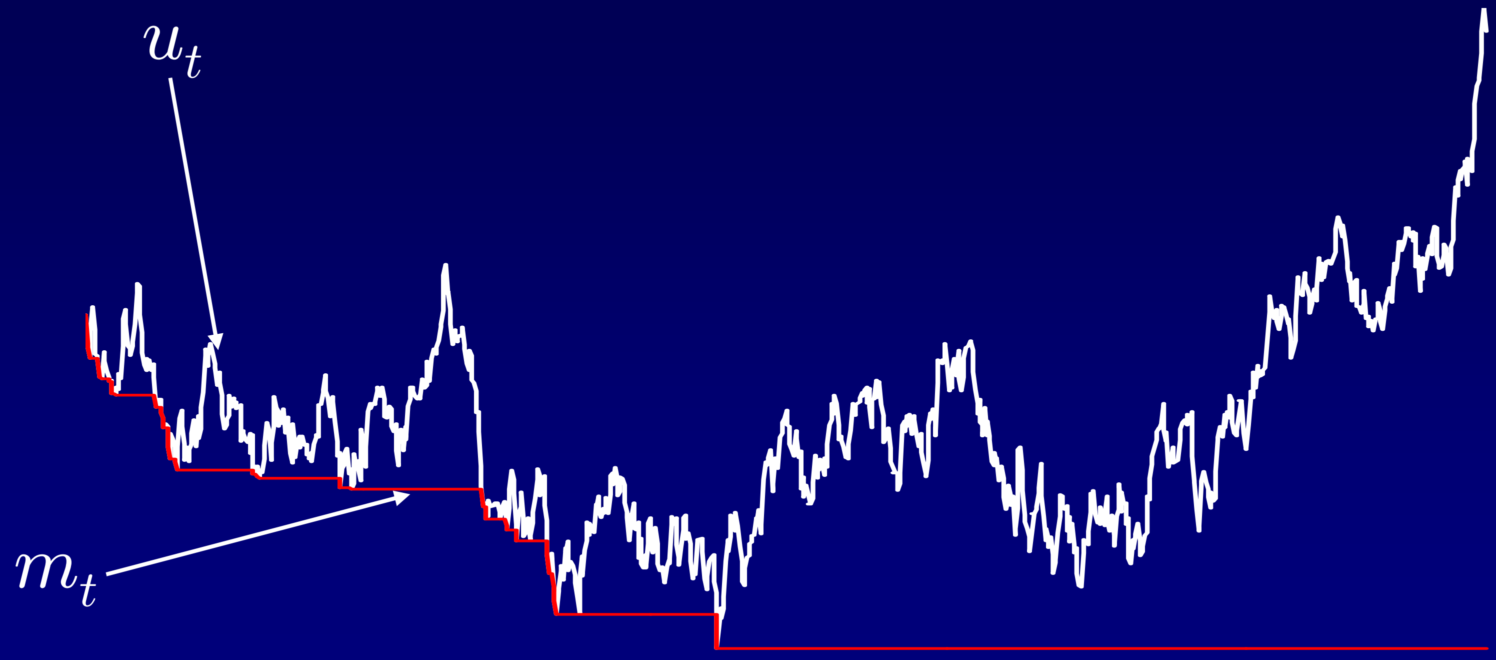
$$y_t = u_t - m_t$$

and the optimum CUSUM test is

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

where ν such that: $\mathbb{E}_\infty\left[0.5\int_0^{T_C} \alpha_t^2 \, dt\,\right] = \gamma$

Since y_t has continuous paths, when the CUSUM test stops we have: $y_{T_C} = \nu$.



- Since $u_t \geqslant m_t$ we conclude $y_t = u_t - m_t \geqslant 0$
- m_t is nonincreasing and $dm_t \neq 0$ only when $u_t = m_t$ or $y_t = u_t - m_t = 0$
- If $f(y)$ continuous with $f(0) = 0,$ then $f(y_t)dm_t = 0$

If $f(y)$ is a twice continuously differentiable function with $f'(0)=0$, using standard Ito calculus, we can write

$$df(y_t) = f'(y_t)(du_t - dm_t) + 0.5\alpha_t^2 f''(y_t)dt$$
$$= f'(y_t)du_t + 0.5\alpha_t^2 f''(y_t)dt$$

Theorem 1: T_C is a.s. finite, furthermore

$$\mathbb{E}_\tau\Big[\mathbb{1}_{\{T_C>\tau\}} 0.5\int_\tau^{T_C}\alpha_t^2\,dt \;\Big|\; \mathcal{F}_\tau\Big] = [g(\nu) - g(y_\tau)]\;\mathbb{1}_{\{T_C>\tau\}}$$

$$\mathbb{E}_\infty\Big[\mathbb{1}_{\{T_C>\tau\}} 0.5\int_\tau^{T_C}\alpha_t^2 dt \;\Big|\; \mathcal{F}_\tau\Big] = [h(\nu) - h(y_\tau)]\;\mathbb{1}_{\{T_C>\tau\}}$$

$$g(y) = y + e^{-y} - 1 \qquad\qquad h(y) = e^{y} - y - 1$$

The functions $g(y)$, $h(y)$ are increasing, strictly convex, with $g(0) = h(0) = 0$. We can therefore conclude

$$J(T_C) = \sup_\tau \text{ essup } \mathbb{E}_\tau\Big[\mathbb{1}_{\{T_C > \tau\}} 0.5\int_\tau^{T_C} \alpha_t^2 dt \,\Big|\, \mathcal{F}_\tau\Big]$$

$$= \sup_\tau \text{ essup}[g(\nu) - g(y_\tau)]\, \mathbb{1}_{\{T_C > \tau\}}$$

$$= g(\nu) - g(0) = g(\nu) = \nu + e^{-\nu} - 1$$

Similarly

$$\mathbb{E}_\infty\Big[0.5 \int_0^{T_C} \alpha_t^2 \, dt\Big] = h(\nu) - h(0) = h(\nu) = \gamma$$

$$\boxed{e^\nu - \nu - 1 = \gamma}$$

For any stopping time T, using again standard Ito calculus, we have the following corollary of Theorem 1

Corollary:

$$\mathbb{E}_\tau\left[\mathbb{1}_{\{T>\tau\}} 0.5\int_\tau^T \alpha_t^2\,dt\,\Big|\,\mathcal{F}_\tau\right] = \mathbb{E}_\tau[g(y_T)\text{-}g(y_\tau)|\mathcal{F}_\tau]\ \mathbb{1}_{\{T>\tau\}}$$

$$\mathbb{E}_\infty\left[\mathbb{1}_{\{T>\tau\}} 0.5\int_\tau^T \alpha_t^2 dt\,\Big|\,\mathcal{F}_\tau\right] = \mathbb{E}_\infty[h(y_T)\text{-}h(y_\tau)|\mathcal{F}_\tau]\ \mathbb{1}_{\{T>\tau\}}$$

Remark 1: The false alarm constraint can be written as

$$\begin{aligned}\mathbb{E}_\infty\left[0.5\int_0^T \alpha_t^2\,dt\right] &= \mathbb{E}_\infty[h(y_T)\text{-}h(0)]\\ &= \mathbb{E}_\infty[h(y_T)] \geqslant \gamma\end{aligned}$$

Remark 2: The modified performance measure $J(T)$ can be suitably lower bounded as follows

$$J(T) = \sup_{\tau} \operatorname{essup} \mathbb{E}_{\tau}\Big[\mathbb{1}_{\{T > \tau\}} 0.5 \int_{\tau}^{T} \alpha_t^2 \, dt \,\Big|\, \mathcal{F}_{\tau}\Big]$$

$$\geqslant \frac{\mathbb{E}_{\infty}[e^{y_T} g(y_T)]}{\mathbb{E}_{\infty}[e^{y_T}]}$$

In the case of CUSUM the lower bound coincides with the corresponding performance measure $J(T_C)$

Remark 3: We can limit ourselves to stopping times that satisfy the false alarm constraint **with equality**, i.e

$$\mathbb{E}_{\infty}[h(y_T)] = \gamma = h(\nu)$$

Theorem 2: Any stopping T that satisfies the false alarm constraint with equality has a performance measure $J(T)$ that is no less than $J(T_C)= g(\nu)$.

Proof: Let T satisfy the false alarm constraint with equality, i.e.

$$\mathbb{E}_\infty[h(y_T)] = \gamma = h(\nu)$$

we then like to show that: $J(T) \geqslant g(\nu)$.

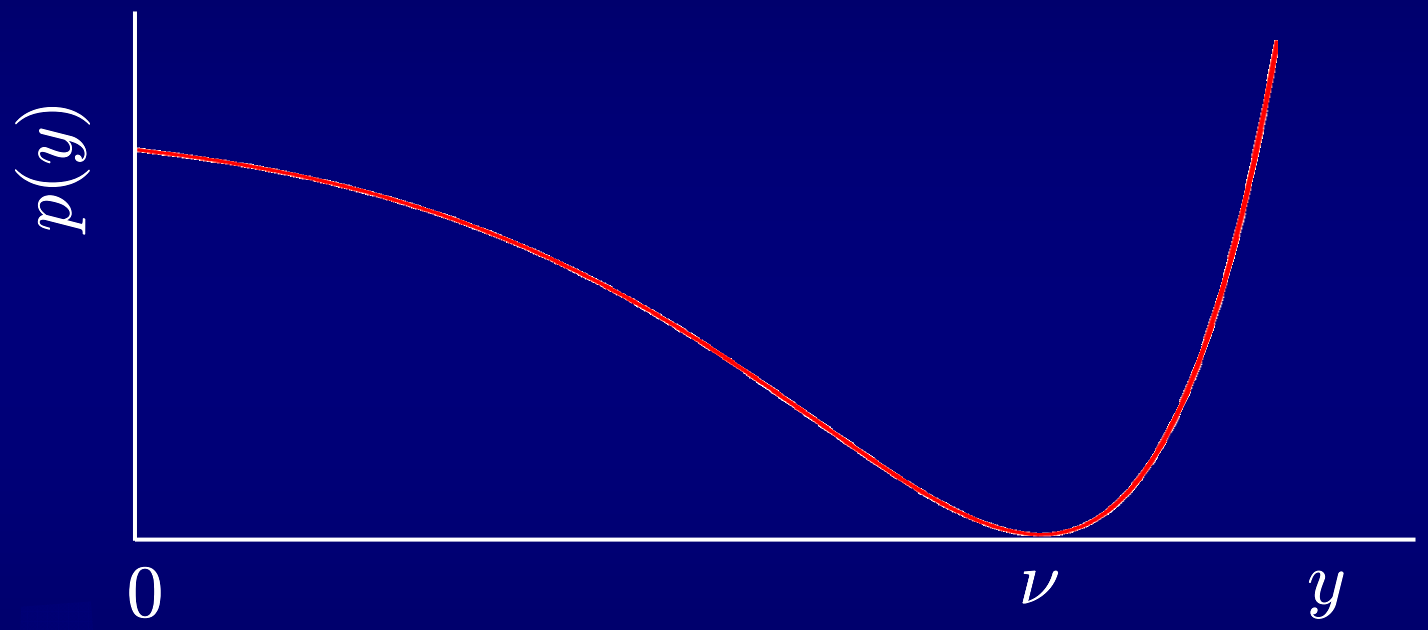
Since $J(T) \geqslant \dfrac{\mathbb{E}_\infty[e^{y_T} g(y_T)]}{\mathbb{E}_\infty[e^{y_T}]}$ it is sufficient to show

$$\mathbb{E}_\infty[e^{y_T}\{g(y_T) - g(\nu)\} + h(\nu) - h(y_T)] \geqslant 0$$

If we define the function

$$p(y) = e^y\{g(y) - g(\nu)\} + h(\nu) - h(y)$$

then the previous inequality becomes: $\mathbb{E}_\infty[p(y_T)] \geqslant 0$



We observe that $p(y) \geqslant 0$

therefore we also have $\mathbb{E}_\infty[p(y_T)] \geqslant 0$

with equality iff $y_T = \nu$, i.e. the CUSUM test.

Extensions

Can our result be extended to the discrete time case?

$$\xi_n = \begin{cases} w_n & 0 \leqslant n \leqslant \tau \\ \alpha_{n\text{-}1} + w_n & \tau < n \end{cases}$$

$\{w_n\}$ an i.i.d. Gaussian process

$\{\alpha_n\}$ adapted to the history $\mathcal{F}_n = \sigma\{\xi_k : 0 \leqslant k \leqslant n\}$

Not Straightforward !

$$\mathbb{E}\left[\mathbb{1}_{\{T > \tau\}}\, 0.5 \sum_{k=\tau}^{T} \alpha_k^2 \,\middle|\, \mathcal{F}_\tau\right] \stackrel{?}{=} \mathbb{E}[g(y_T)\text{-}g(y_\tau)|\mathcal{F}_\tau]\ \mathbb{1}_{\{T > \tau\}}$$

Similar problem exists for SPRT.

Straightforward extension for scalar processes

$$d\xi_t = \begin{cases} \alpha_t \, dt + \sigma_t dw_t & 0 \leqslant t \leqslant \tau \\ \beta_t \, dt + \sigma_t dw_t & \tau < t \end{cases}$$

or vector processes

$$d\Xi_t = \begin{cases} A_t \, dt + \Sigma_t dW_t & 0 \leqslant t \leqslant \tau \\ B_t \, dt + \Sigma_t dW_t & \tau < t \end{cases}$$

EnD