Metrics and Optimum Tests
for Sequential Change-Detection

George V. Moustakides
University of Patras, Greece
& Rutgers University, USA
Outline

- Definition of the change detection problem, Applications, Questions with existing formulations
- Model for mechanism that imposes change
- Two generic setups for sequential detection of changes
  - Complete knowledge
  - Incomplete knowledge
Problem definition

Change-time

\{\xi_t\}

\[ P_0(E_0) \]

\[ P_1(E_1) \]

\( \tau \)

Detect change as soon as possible

Data become available sequentially: at each instant \( t \) obtain new sample \( \xi_t \).

Detector: Every instant \( t \) consult available data \( \xi_1, \ldots, \xi_t \) Use them to make Binary Decision.
Each instant $t$ decide between:
- A change took place before and including $t$.
- A change didn’t take place before and including $t$.

There will be a point in time (call it) $T$ where we stop.
$T$ is random time controlled by the observations.
Need a test to implement binary decision at each time $t$:

$$T = \inf \{ G_t(\xi_1, \ldots, \xi_t) \geq \nu \}$$

Stopping rule

Class of stopping times: Extremely rich

Design $T$ to minimize Detection Delay ($T - \tau$) while controlling False Alarms.
Quality monitoring of manufacturing process

Production line → Continuous Measurements → Quality Assessment

Medical Applications

Epidemic Detection
- Disease rate measurements → Increase in rate → Epidemic outbreak?

Early Detection of Epilepsy Episode
- EEG Wearables → Divergence from normal → Episode?
Financial Applications

- Structural Change-detection in Exchange Rates
- Portfolio Monitoring

Electronic Communications

Seismology

Speech & Image Processing (segmentation)

Vibration monitoring (Structural health monitoring)

Security monitoring (fraud detection)

Spectrum monitoring

Scene monitoring

Network monitoring (router failures, attack detection)

CUSUM: 3,000 hits in 2015. Google Scholar.

80% in Change Detection: 2300 articles
Performance Metrics

\[ J(T) = E_1[T - \tau \mid T > \tau] \]

Metric must measure only success
Failures will be dealt through False Alarm control

**Shiryaevo (1963):** $\tau$ is random with known prior:

\[ J_S(T) = E_1[T - \tau \mid T > \tau] \]

Too restrictive!

**Pollak (1985):** $\tau$ is deterministic and unknown:

\[ J_P(T) = \sup_{t \geq 0} E_1[T - t \mid T > t] \]

Leads to SR test

**Lorden (1971):** $\tau$ is deterministic and unknown:

\[ J_L(T) = \sup_{t \geq 0} \sup_{\xi_1, \ldots, \xi_t} E_1[T - t \mid T > t, \xi_1, \ldots, \xi_t] \]

Leads to CUSUM

Too pessimistic (?)
Model for change imposing mechanism

A random vector process $\{X_t\}$ evolves in time in $\mathbb{R}^K$

$A$ is a subset in $\mathbb{R}^K$

$\tau = \inf\{t > 0 : X_t \in A\}$

$\tau$ is a first entry time, depends on $\{X_t\}$.

We would like to detect it.

If $\{X_t\}$ observable and $A$ known, problem is trivial.

If $\{X_t\}$ (partially) hidden and/or $A$ unknown, problem is challenging.

Instead of \( \{X_t\} \) we observe process \( \{\xi_t\} \)

Entry to the set, generates 
change in statistical behavior

Sequential Change Detection Problem
First-entry: Model for change-imposing mechanism.

- Unifies all existing formulations
- Understanding of change-imposing mechanism can explain existing metrics
- May lead to more efficient detectors.

**Goal: detect occurrence of** $\tau$

$\tau$ is a stopping time adapted to the history generated by the (partially) hidden process $\{X_t\}$.

$T$ is a stopping time adapted to the history generated by the observation sequence $\{\xi_t\}$.
Power Grid:

$X_t$: Energy at major points in the grid.

$\xi_t = X_t + W_t$ noisy measurements.

$A$: If $X_t \in A$ then, after short time major blackout.

$A$ is known

Structure health monitoring:

\[ X_t: \text{ Vibrations at every point of the structure (state)} \]

\[ \xi_t = A X_t + W_t: \text{ Low dimensional noisy measurements} \]

\[ A: \text{ If } X_t \in A \text{ then cracks (change in the structure)} \]

\[ A \text{ known or unknown.} \]

Independent $\{X_t\}$ and $\{\xi_t\}$ ?:

$X_t$: Field coordinates of the ball

$\xi_t$: Noisy vibration measurements

$A$: Volume under the goal net.
Performance metrics - known entry set

Delayed Detection

\[ \mathcal{J}(T) = \mathbb{E}_1[T - \tau | T > \tau] \]

\[ \inf_T \mathcal{J}(T) = \inf_T \mathbb{E}_1[T - \tau | T > \tau] \]
subject to: \(P_0(T \leq \tau) \leq \alpha\)

Hard Limited Delay

\[ \mathcal{P}(T) = \mathbb{P}_1(T \leq \tau + M | T > \tau) \]

\[ \sup_T \mathcal{P}(T) = \sup_T \mathbb{P}_1(T \leq \tau + M | T > \tau) \]
subject to: \(P_0(T \leq \tau) \leq \alpha\)
Delayed detection - known entry set

\[
\inf_T \mathcal{J}(T) = \inf_T \mathbb{E}_1[T - \tau | T > \tau]
\]

subject to: \( P_0(T \leq \tau) \leq \alpha \)

**Optimal Stopping Theory**

Pair process \( \{(X_t, \xi_t)\} \) is i.i.d. before and after \( \tau \) with joint pdfs \( f_0, f_1 \).

Moustakides (2016): The optimum test

\[
\pi_t = P_0(X_t \in A | \xi_t)
\]

\[
T_o = \inf\{t > 0 : S_t \geq \nu\}
\]

\[
S_t = S_{t-1} \frac{f_1(\xi_t)}{(1 - \pi_t) f_0(\xi_t)} + \frac{\pi_t}{1 - \pi_t}
\]

If additionally \( \{X_t\} \) and \( \{\xi_t\} \) independent processes, then

\[
\pi_t = P_0(X_t \in A|\xi_t) = P_0(X_t \in A) = \pi
\]

\[
S_t = S_{t-1} \frac{f_1(\xi_t)}{(1-\pi)f_0(\xi_t)} + \frac{\pi}{1-\pi}
\]

\[
\tilde{S}_t = \frac{1-\pi}{\pi} S_t - 1, \quad \tilde{\nu} = \frac{1-\pi}{\pi} \nu - 1
\]

\[
T_0 = \inf\{t > 0 : \tilde{S}_t \geq \tilde{\nu}\}
\]

\[
\tilde{S}_t = (\tilde{S}_{t-1} + 1) \frac{f_1(\xi_t)}{(1-\pi)f_0(\xi_t)}
\]

Shiryayev test (1963)
Unknown entry set

- What if entry set $A$ is unknown?
- Can detect the first-entry to an unknown set? Equivalently: can detect the change-time $\tau$ that inflicts a change in the statistical behavior?

Focus on change of the statistics
Performance metrics - unknown entry set

Delayed Detection

\[ J(T) = \sup_{\tau} E_1[T - \tau | T > \tau] \]

\[ \inf_{T} J(T) = \inf_{T} \sup_{\tau} E_1[T - \tau | T > \tau] \]

subject to: \( E_0[T] \geq \gamma \)

Hard Limited Delay

\[ P(T) = \inf_{\tau} P_1(T \leq \tau + M | T > \tau) \]

\[ \sup_{T} P(T) = \sup_{T} \inf_{\tau} P_1(T \leq \tau + M | T > \tau) \]

subject to: \( E_0[T] \geq \gamma \)
Delayed detection - unknown entry set

\[ \mathcal{J}(T) = \sup_{\tau} \mathbb{E}_{1}[T - \tau | T > \tau] \]

We can show (Moustakides 2008):

\{X_t\} and \{\xi_t\} independent processes

\[ \mathcal{J}_P(T) = \sup_{t \geq 0} \mathbb{E}_{1}[T - t | T > t] \quad \text{Pollak (1985)} \]

\{X_t\} and \{\xi_t\} dependent processes

\[ \mathcal{J}_L(T) = \sup_{t \geq 0} \sup_{\xi_1, \ldots, \xi_t} \mathbb{E}_{1}[T - t | T > t, \xi_1, \ldots, \xi_t] \quad \text{Lorden (1971)} \]
\[ J_P(T) = \sup_{t \geq 0} E_1[T - t | T > t] \quad \text{Pollak (1985)} \]

\[
\inf_T J_P(T) \quad \text{subject to: } E_0[T] \geq \gamma
\]

Discrete time: i.i.d. data before and after the change with pdfs \( f_0, f_1 \).

**Shiryaev-Roberts-Pollak test**

Compute recursively the following statistic:

\[
S_t = (S_{t-1} + 1) \frac{f_1(\xi_t)}{f_0(\xi_t)} \quad T_P = \inf \{ t > 0 : S_t \geq \nu \}
\]

Pollak (1985): If \( S_0 \) specially designed, then

\[
[J_P(T_P) - \inf_T J_P(T)] \to 0; \quad \text{as } \gamma \to \infty
\]
Exact optimality?


Change in the drift of a BM: Polunchenko (2016)

Dependence? Multiple pre- and/or post-change possibilities? Time variation?
\[ \mathcal{J}_L(T) = \sup_{t \geq 0} \sup_{\xi_1, \ldots, \xi_t} \mathbb{E}_1[T - t | T > t, \xi_1, \ldots, \xi_t] \]

Lorden (1971)

\[ \inf_T \mathcal{J}_L(T) \text{ subject to: } \mathbb{E}_0[T] \geq \gamma \]

Discrete time: i.i.d. data before and after the change with pdfs \( f_0, f_1 \).

**CUSUM test**

\[ S_t = \max\{S_{t-1}, 0\} + \log \frac{f_1(\xi_t)}{f_0(\xi_t)} \]

\[ T_C = \inf\{t > 0 : S_t \geq \nu\} \]

\[ u_t = \sum_{s=1}^{t} \log \frac{f_1(\xi_s)}{f_0(\xi_s)} \]

\[ m_t = \min_{0 \leq s \leq t} u_s \]

\[ S_t = u_t - m_t \]


Discrete time: Moustakides-Veeravalli (2016) Non abrupt changes

Dependence? Multiple pre- and/or post-change possibilities?
Hard Limited Delay

$$P_1(T \leq \tau + M | T > \tau)$$

Only for $M = 1$:  $P_1(T = \tau + 1 | T > \tau)$

Corresponds to immediate detection with the first sample after the change

$$P_S(T) = P_1(T = \tau + 1 | T > \tau) \quad \text{Shiryaev like}$$

$$P_P(T) = \inf_{t \geq 0} P_1(T = t + 1 | T > t) \quad \text{Pollak like}$$

$$P_L(T) = \inf_{t \geq 0} \inf_{\xi_1, \ldots, \xi_t} P_1(T = t + 1 | T > t, \xi_1, \ldots, \xi_t) \quad \text{Lorden like}$$

$$\sup_T P_S(T) \quad \text{s.t.} \quad P_0(T \leq \tau) \leq \alpha$$

$$\sup_T P_P(L)(T) \quad \text{s.t.} \quad E_0[T] \geq \gamma$$

\[ T_{Sh} = \inf \left\{ t > 0 : \frac{f_1(\xi_t)}{f_0(\xi_t)} \geq \nu \right\} \]  

Shewhart test (1931)

Optimality: Bojdecki (1979): Shiryaev like
Pollak and Krieger (2013): Pollak like
Moustakides (2014): Lorden like

Moustakides (2014): Post change time variation
Dependent observations

\[ P_L(T) = \inf_{t \geq 0} \inf_{\xi_1, \ldots, \xi_t} P_1(T = t + 1 | T > t, \xi_1, \ldots, \xi_t) \]

\[ \sup_T P_L(T) \quad \text{subject to} \quad E_0[T] \geq \gamma \]

**Markovian** pre- and post-change observations \( \{\xi_t\} \)

\[ T_{Sh} = \inf \left\{ t > 0 : c(\xi_{t-1}) \frac{f_1(\xi_t | \xi_{t-1})}{f_0(\xi_t | \xi_{t-1})} \geq \nu(\xi_t) \right\} \]

Moustakides (2015): With properly designed functions \( c(\xi) \) and \( \nu(\xi) \) we solve the constrained optimization.

Simple solution for conditionally Gaussian pdfs.

\[ \mathcal{J}_L(T) = \sup_{t \geq 0} \sup_{\xi_1, \ldots, \xi_t} E_1[T - t | T > t, \xi_1, \ldots, \xi_t] \]

\[
\inf_T \mathcal{J}_L(T) \text{ subject to : } E_0[T] \geq \gamma
\]

For Markovian pre- and post-change \( \{\xi_t\} \): Solution ??

Acknowledgement

NSF: CIF-1513373 through Rutgers University
Collaboration program with UIUC

Thank you for your attention!