Quickest Change Detection Using Mismatched CUSUM

Sean Meyn, University of Florida

The field of quickest change detection (QCD) concerns design and analysis of algorithms to estimate in real time the time at which an important event takes place, and identify properties of the post-change behavior. The classical model is considered, with change time denoted τ_{a}:

$$Y_k = X_k^0 \mathbf{1}_{k<\tau_{\mathsf{a}}} + X_k^1 \mathbf{1}_{k\ge\tau_{\mathsf{a}}}\,, \qquad k \ge 0\,,$$

in which $\{X_k^0, X_k^1\}$ are stationary processes, evolving on a set Y. The goal is to devise a stopping time τ_{s}, adapted to the observations, and minimizing a mean weighted ℓ_1 loss: for fixed $\kappa > 0$,

$$J(\tau_{\mathsf{s}}) = \mathsf{E}\big[(\tau_{\mathsf{s}} - \tau_{\mathsf{a}})_+ + \kappa(\tau_{\mathsf{s}} - \tau_{\mathsf{a}})_-\big]$$

An approximately optimal solution is obtained based on the CUSUM statistic, defined as the one-dimensional reflected process,

$$\mathcal{X}_{n+1} = \max\{0, \mathcal{X}_n + F(Y_{n+1})\}$$

which defines the stopping rule $\tau_{\mathsf{s}} = \min\{n \ge 0 : \mathcal{X}_n \ge \mathrm{H}\}$. Subject to conditions on $\{X_k^0, X_k^1\}$ and compatible conditions on the function F, the average cost $J(\tau_{\mathsf{s}})$ admits an approximation based on two non-negative scalars:

$$\varrho_{\mathsf{a}} = -\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\{\tau_{\mathsf{a}} \ge n\} \quad \textit{and } \theta_+^* > 0 \textit{ the solution to } \ \Lambda_0(\theta_+^*) = \varrho_{\mathsf{a}}\,,$$

with Λ_0 the log cumulative generating function for the partial sums of $\{F(X_k^0)\}$.

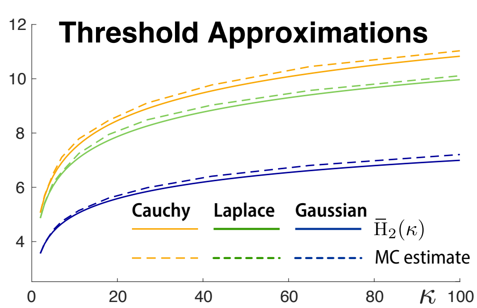
In the talk we will explain the main conclusions:

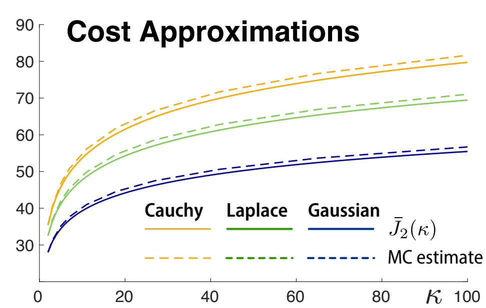
1. Large-κ approximations for the optimal threshold and corresponding average cost:

$$\bar{\mathrm{H}}_\infty^*(\kappa) = \frac{1}{\theta_+^*} \log(\kappa m_1 \theta_+^*)\,, \quad \bar{J}_\infty^*(\kappa) = \frac{1}{m_1\theta_+^*}\Big[1 + \log(\kappa m_1 \theta_+^*)\Big]\,, \quad \textit{with } m_1 = \mathsf{E}[F(X_k^1)] > 0.$$

2. For the special case $\varrho_{\mathsf{a}} = 0$, the approximation $\bar{J}_\infty^*(\kappa)$ is minimized when $F = L$, the log likelihood of the marginals of $\{X_k^0, X_k^1\}$.

3. The potential for the creation of architectures for reinforcement learning based on this theory.





Based on the recent research with Austin Cooper@UF, *Reinforcement Learning Design for Quickest Change Detection*, https://arxiv.org/abs/2403.14109