Change-point models and performance measures for sequential change detection

George V. Moustakides

Department of Electrical and Computer Engineering,
University of Patras, 26500 Rion, Greece
moustaki@upatras.gr

Abstract. For the problem of sequential change detection we propose a novel modelling of the change-point mechanism. In particular we regard the time of change as a stopping time controlled by Nature. Nature, in order to decide when to impose the change, accesses sequentially information which can be different from the information provided to the Statistician to detect the change. Using as performance measure the classical conditional detection delay, we recover most well known criteria of the literature by considering different dependency classes between the information accessed by Nature and the information accessed by the Statistician. According to our approach, the Shiryaev and Pollak measure correspond to informations that are completely independent while in Lorden's measure the two informations must coincide. By considering alternative models between these two extreme scenarios, we obtain a number of completely new criteria.

Keywords. Quickest detection, Sequential change detection.

1 Change-point probability measures

Suppose that we have a process $\{x_t\}_{t=-\infty}^{\infty}$ and two probability measures $\mathbb{P}_0, \mathbb{P}_\infty$ that both describe its statistics. Let us for simplicity assume that both measures accept a pdf which with a slight abuse of notation we are going to denote as $f_i(\ldots, x_1, x_0, x_{-1}, \ldots),\ i = 0, \infty$. Assume now that there is a *deterministic* time τ such that the data $\{x_t\}_{t=-\infty}^{\tau}$ follow the (nominal) probability measure $\mathbb{P}_\infty$ while after τ we have that $\{x_t\}_{t=\tau+1}^{\infty}$ are according to the (alternative) measure $\mathbb{P}_0$. In other words, there is a change in statistics at time τ. This change induces a new probability measure $\mathbb{P}_\tau$ and our goal in this introductory section is to find suitable model to describe it.

As we did in the case of the two initial measures, we are going to assume that there is a pdf for the desired $\mathbb{P}_\tau$ which we denote as $f_\tau(\ldots, x_{-1}, x_0, x_1, \ldots)$. Using the Bayes rule we can decompose the unknown density as follows

$$f_\tau(\ldots, x_{\tau+2}, x_{\tau+1}, x_\tau, x_{\tau-1}, \ldots) = f_\tau(x_\tau, x_{\tau-1}, \ldots) f_\tau(x_{\tau+1}, x_{\tau+2}, \ldots | x_\tau, x_{\tau-1}, \ldots). \tag{1}$$

It is clear that the first part describes the pre-change data while the second the post-change. Undoubtedly the first part mast be selected as

$$f_\tau(x_\tau, x_{\tau-1}, \ldots) = f_\infty(x_\tau, x_{\tau-1}, \ldots) \tag{2}$$

expressing the fact that the pre-change data follow the nominal regime. Of course the key point is how we select the second component. It turns out that there are two possibilities.

1.1 Pdf model1

We can clearly have a similar decomposition as in (1) for the alternative model, that is,

$$f_0(\ldots, x_{\tau+2}, x_{\tau+1}, x_\tau, x_{\tau-1}, \ldots) = f_0(x_\tau, x_{\tau-1}, \ldots) f_0(x_{\tau+1}, x_{\tau+2}, \ldots | x_\tau, x_{\tau-1}, \ldots).$$

and select the second component of this decomposition as the second part in (1), that is,

$$f_\tau(x_{\tau+1}, x_{\tau+2}, \ldots | x_\tau, x_{\tau-1}, \ldots) = f_0(x_{\tau+1}, x_{\tau+2}, \ldots | x_\tau, x_{\tau-1}, \ldots) \tag{3}$$

resulting in the following model for the probability measure imposed by the change

$$f_\tau(\ldots, x_{\tau+2}, x_{\tau+1}, x_\tau, x_{\tau-1}, \ldots) = f_\infty(x_\tau, x_{\tau-1}, \ldots) f_0(x_{\tau+1}, x_{\tau+2}, \ldots | x_\tau, x_{\tau-1}, \ldots). \tag{4}$$

This change-point model clearly suggests a continuity in the data flow with the pre-change data *influencing* the post-change (because of the conditional probability measure). As the next example reveals this type of change is perhaps the most frequent encountered in practice.

Example 1. A very common mathematical model used to describe physical systems is the following linear state-space difference equation

$$X_t = \mathbf{A}X_{t-1} + V_t \tag{5}$$

where $\{X_t\}$ denotes the state of the system; $\mathbf{A}$ is the system matrix and $\{V_t\}$ denotes the system excitation (input). This equation is usually accompanied by

$$y_t = \mathbf{C}X_t + W_t \tag{6}$$

where $\{y_t\}$ denotes the observation sequence (output); $\{W_t\}$ an additive measurement noise (usually taken to be i.i.d. and independent from $\{V_t\}$) and $\mathbf{C}$ a matrix that combines state elements to produce the observations. Equ.(5) is used to model physical systems while (6) measurement devices attached to the physical system. Regarding failures, we distinguish

System failures: corresponding to a change in the system matrix from $\mathbf{A}$ to $\mathbf{A}'$.

Device failures: corresponding to a change in the measurement device matrix from $\mathbf{C}$ to $\mathbf{C}'$.

In both failures the data before the change influence the data after the change. Indeed in the case of a system failure we have

$$X_t = \mathbf{A}'X_{t-1} + V_t, \text{ for } t > \tau.$$

It is clear that at $t = \tau + 1$ the state $X_{\tau+1}$ depends on X_τ with the latter being under the nominal statistics and serving as initial condition for the state evolution for $t > \tau$. The same is true if we have a device failure since the observations at time $t > \tau$ clearly depend on the observations before the change.

1.2 Pdf model 2

A different way to decompose the alternative measure is

$$f_0(\ldots, x_{\tau+2}, x_{\tau+1}, x_\tau, x_{\tau-1}, \ldots) = f_0(x_\tau, x_{\tau-1}, \ldots | x_{\tau+1}, x_{\tau+2}, \ldots) f_0(x_{\tau+1}, x_{\tau+2}, \ldots)$$

and select again the second component to describe the data after the change, yielding the following density

$$f_\tau(\ldots, x_{\tau+2}, x_{\tau+1}, x_\tau, x_{\tau-1}, \ldots) = f_\infty(x_\tau, x_{\tau-1}, \ldots) f_0(x_{\tau+1}, x_{\tau+2}, \ldots). \tag{7}$$

This model suggests that the data before the change are *independent* from the data after the change. Of course the samples themselves before and after the change can be dependent. This is like having two independent processes evolving in parallel (with the samples of each process being dependent) and the process $\{x_t\}$ originally follows the first process and at some point in time it switches to the second. This type of change-point occurs in Example1 when there is a device failure with $\mathbf{C}' = 0$. The device after the change is simply measuring white noise (a relatively common failure in practice) which is independent from the state and therefore from the previous observations. There is also the following characteristic application that falls under the second change-time model.

Example 2. Consider the problem of object segmentation in images where we need to specify the border of an object that lies over a random background. This problem can be viewed as a change detection problem with the change following (7). Indeed, if we process the image by lines then the sequence of pixel intensities constitutes the available data sequence. The transition from background to object is abrupt and there is clearly no information continuity as we switch from one regime to the other.

The two pdf models (4) and (7) are statistically very dissimilar and coincide when the process $\{x_t\}$ has independent components. Let us now see how we can apply the corresponding probability measures to compute the expectation of a random variable. Suppose $\mathcal{X}(\ldots, x_{\tau-1}, x_\tau, x_{\tau+1}, x_{\tau+2}, \ldots)$ is a function of the process $\{x_t\}$ and we are interested in computing its average. Using (4) we conclude

$$\begin{aligned}\mathbb{E}_\tau[\mathcal{X}] &= \int \left[\int \mathcal{X} f_0(x_{\tau+1}, x_{\tau+2}, \ldots | x_\tau, x_{\tau-1}, \ldots) dx_{\tau+1} dx_{\tau+2} \cdots \right] f_\infty(x_\tau, x_{\tau-1}, \ldots) dx_\tau dx_{\tau-1} \cdots \\ &= \mathbb{E}_\infty \left[\mathbb{E}_0[\mathcal{X} | x_\tau, x_{\tau-1}, \ldots] \right], \end{aligned} \tag{8}$$

In a similar way we can show that the second model yields

$$\mathbb{E}_\tau[\mathcal{X}] = \mathbb{E}_\infty\left[\mathbb{E}_0^{\tau+1,\infty}[\mathcal{X}|x_\tau, x_{\tau-1}, \ldots]\right], \tag{9}$$

where with $\mathbb{E}_0^{k,\infty}[\cdot]$ we denote expectation with respect to the pdf $f_0(x_k, x_{k+1}, \ldots)$. The expectation in (8) is the form more frequently encountered in the literature suggesting that the adopted change-point model is the one proposed in (4). Due to its practical importance, from now on we limit ourselves to the first model.

2 Change-point mechanism and performance measure

As it is the case in many problems in Statistics, we can view the sequential change detection problem as a conflict between Nature and the Statistician. Nature, at some unknown time τ, imposes a change in the statistics of an observed sequence, while the Statistician uses the observed sequence to decide whether the change took place or not.

As we pointed out, the Statistician has access only to the observed sequence to reach a decision. Nature, on the other hand, can access the same, additional or even completely different information to reach the decision as to when the change must be imposed. In order to formulate this difference in available information, we are going to assume that Nature obtains *sequentially* a processes $\{z_t\}_{t=-\infty}^{\infty}$ and uses this time increasing information to make a decision about the change. In parallel, the Statistician acquires, sequentially again, an alternative process $\{\xi_t\}_{t>0}^{\infty}$ for $t > 0$, in order to decide whether a change took place or not. Process $\{\xi_t\}_{t>0}$ is the *observation process* while $\{z_t\}_{t=-\infty}^{\infty}$ is *Nature's information process*. The two processes can be dependent and, we can even assume that $\{\xi_t\}$ is part of $\{z_t\}$, in order to model the case where Nature uses more information to decide the time of change, than the information provided to the Statistician do detect the change. A very interesting point that must be noted is the fact that the process $\{z_t\}_{t=-\infty}^{\infty}$ can be acquired by Nature *well before* $\{\xi_t\}_{t>0}^{\infty}$ is provided to the Statistician. This allows for modeling the case where the change takes place before the Statistician obtains any observations.

By selecting $x_t = (z_t, \xi_t)$ we can use the results of our previous section. Therefore let us assume that there are two probability measures $\mathbb{P}_i,\ i = \infty, 0$ that describe the joint statistics of the two processes. If τ is some deterministic time of change then the induced measure $\mathbb{P}_\tau$ follows (4). Consider also the two filtrations $\{\mathcal{F}_t^z\}_{t=-\infty}^{\infty}$, $\{\mathcal{F}_t^\xi\}_{t=0}^{\infty}$ with $\mathcal{F}_t^z = \sigma\{z_s; -\infty < s \leq t\}$; $\mathcal{F}_t^\xi = \sigma\{\xi_s; 0 < s \leq t\}$, with $\mathcal{F}_0^\xi$ being the trivial σ-algebra. In fact we will extend our last definition and assume that for any $t \leq 0$, $\mathcal{F}_t^\xi$ is the trivial σ-algebra. If $\mathcal{X}$ is now a random variable which is measurable with respect to $\mathcal{F}_\infty^z \cup \mathcal{F}_\infty^\xi$ then, following (8), we can write

$$\mathbb{E}_\tau[\mathcal{X}] = \mathbb{E}_\infty\left[\mathbb{E}_0[\mathcal{X}|\mathcal{F}_\tau^z \cup \mathcal{F}_\tau^\xi]\right]. \tag{10}$$

So far we have only discussed types of possible changes without making reference to the *change-point mechanism* which is responsible for generating the change (i.e. the rule used by Nature to decide for the change). In order to find a suitable model, let us refer to Example1. Usual source of system failure constitutes the entrance of the state X_t into some extremal set $\mathcal{A}$. For example in vibration monitoring of structures, whenever the amplitude of the vibrations becomes exceedingly large, this causes a failure resulting in a system matrix change from $\mathbf{A}$ to $\mathbf{A}'$. Similarly for device failures, when the value $\mathbf{C}X_t$ enters some extremal set $\mathcal{B}$ this causes a change from $\mathbf{C}$ to $\mathbf{C}'$. Note that the change-point mechanism that decides upon the change *uses currently available information to make its decision about imposing the change or not*. This suggests the following general model for the change-point mechanism.

Definition 1: A change-point mechanism employed by Nature is a *stopping time τ adapted to the filtration $\{\mathcal{F}_t^z\}$*.

According to this definition, the time of change τ is regarded as a stopping time (the last time we use the nominal statistics) which is adapted to the filtration $\{\mathcal{F}_t^z\}$ generated by the information provided to

Nature. For the Statistician that attempts to make a sequential detection of the change, we adopt the usual definition

Definition 2: Any sequential detection structure can me modeled as a stopping time T adapted to the filtration $\{\mathcal{F}_t^\xi\}$.

Next, our goal is to define a suitable performance measure the pays attention to the average detection delay, that is, the expectation of $T-\tau$. Note that this quantity, for every fixed τ is a random variable which is $\mathcal{F}_\infty^\xi$-measurable. Consequently let us first derive a convenient formula for the expectation of a randomly stopped sequence of $\mathcal{F}_\infty^\xi$-measurable random variables.

2.1 Expectation of a randomly stopped sequence

Consider a process $\{\mathcal{X}_t\}_{t=0}^\infty$ where for all $t\geq 0$ the random variables $\mathcal{X}_t$ are nonnegative and $\mathcal{F}_\infty^\xi$-measurable. We are interested in computing the expectation of the randomly stopped random variable $\mathcal{X}_\tau$ where τ is an $\{\mathcal{F}_t^z\}$-adapted stopping time. We define

$$\mathcal{X}_\tau \mathbb{1}_{\{\tau<\infty\}} = \mathcal{X}_0 \mathbb{1}_{\{\tau\leq 0\}} + \sum_{t=1}^\infty \mathcal{X}_t \mathbb{1}_{\{\tau=t\}}.$$

Taking expectation, applying (10) and using the fact that if τ is adapted to $\{\mathcal{F}_t^z\}$ it is certainly adapted to $\{\mathcal{F}_t^z\cup\mathcal{F}_t^\xi\}$, we obtain

$$\mathbb{E}_\tau[\mathcal{X}_\tau \mathbb{1}_{\{\tau<\infty\}}] = \mathbb{E}_\infty[\mathbb{E}_0[\mathcal{X}_0|\mathcal{F}_0^z\cup\mathcal{F}_0^\xi]\mathbb{1}_{\{\tau\leq 0\}}] + \sum_{t=1}^\infty \mathbb{E}_\infty[\mathbb{E}_0[\mathcal{X}_t|\mathcal{F}_t^z\cup\mathcal{F}_t^\xi]\mathbb{1}_{\{\tau=t\}}] \tag{11}$$

$$= \mathbb{E}_\infty[\mathbb{E}_0[\mathcal{X}_0|\mathcal{F}_0^z]\mathbb{1}_{\{\tau\leq 0\}}] + \sum_{t=1}^\infty \mathbb{E}_\infty[\mathbb{E}_\infty[\mathbb{E}_0[\mathcal{X}_t|\mathcal{F}_t^z\cup\mathcal{F}_t^\xi]|\mathcal{F}_t^z]\mathbb{1}_{\{\tau=t\}}], \tag{12}$$

where for the last equality we used the fact that $\mathcal{F}_0^\xi$ is the trivial σ-algebra and that τ is $\{\mathcal{F}_t^z\}$-adapted. Equ. (12) is very important for the development of our general performance measure. Note that by substituting $X_t=1$ we assess that $\mathbb{P}_\tau[\tau<\infty]=\mathbb{P}_\infty[\tau<\infty]$.

2.2 A general performance criterion

Let us now use the previous formulas to provide suitable expressions for the average detection delay. We recall that τ is the stopping time employed by Nature to impose the change whereas T is the stopping time employed by the Statistician to detect it. As our performance measure we propose

$$\mathcal{J}(T) = \mathbb{E}_\tau\left[T-\tau|T>\tau;\tau<\infty\right] = \frac{\mathbb{E}_\tau[(T-\tau)^+\mathbb{1}_{\{\tau<\infty\}}]}{\mathbb{E}_\tau[\mathbb{1}_{\{T>\tau\}}\mathbb{1}_{\{\tau<\infty\}}]}, \tag{13}$$

where we have assumed that $\mathbb{P}_\infty[\tau<\infty]>0$.

We note that the two sequences of random variables $\{(T-t)^+\}$ and $\{\mathbb{1}_{\{T>t\}}\}$ are both nonnegative and $\{\mathcal{F}_\infty^\xi\}$-measurable, consequently we can use (12) to evaluate the numerator and the denominator in (13). This yields

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

Using the fact that, since $\{T>t\}$ is $\mathcal{F}_t^\xi$-measurable it is also $\mathcal{F}_t^z\cup\mathcal{F}_t^\xi$-measurable and that $T>0$ a.s., we can simplify the denominator of the previous ratio to

$$\mathcal{J}(T) = \frac{\mathbb{E}_\infty[\mathbb{E}_0[T|\mathcal{F}_0^z]\mathbb{1}_{\{\tau\leq 0\}}] + \sum_{t=1}^\infty \mathbb{E}_\infty[\mathbb{E}_\infty[\mathbb{E}_0[(T-t)^+|\mathcal{F}_t^z\cup\mathcal{F}_t^\xi]|\mathcal{F}_t^z]\mathbb{1}_{\{\tau=t\}}]}{\mathbb{E}_\infty[\mathbb{1}_{\{\tau\leq 0\}}] + \sum_{t=1}^\infty \mathbb{E}_\infty[\mathbb{P}_\infty[T>t|\mathcal{F}_t^z]\mathbb{1}_{\{\tau=t\}}]}. \tag{14}$$

If we do not know exactly the change-point mechanism and we define an uncertainty class $\mathcal{T}$ of possible stopping times for τ, then it makes sense to extend the performance measure by considering the worst possible candidate from the class $\mathcal{T}$. More precisely

$$\mathcal{J}(T) = \sup_{\tau \in \mathcal{T}} \frac{\mathbb{E}_\infty[\mathbb{E}_0[T|\mathcal{F}_0^z]\mathbb{1}_{\{\tau \le 0\}}] + \sum_{t=1}^{\infty} \mathbb{E}_\infty[\mathbb{E}_\infty[\mathbb{E}_0[(T-t)^+|\mathcal{F}_t^z \cup \mathcal{F}_t^\xi]|\mathcal{F}_t^z]\mathbb{1}_{\{\tau=t\}}]}{\mathbb{E}_\infty[\mathbb{1}_{\{\tau \le 0\}}] + \sum_{t=1}^{\infty} \mathbb{E}_\infty[\mathbb{P}_\infty[T > t|\mathcal{F}_t^z]\mathbb{1}_{\{\tau=t\}}]}. \quad (15)$$

If we now consider $\mathcal{T}$ to be the class of all $\{\mathcal{F}_t^z\}$-adapted stopping times, in other words, we assume no other prior information about τ, then the next lemma provides the worst case performance by solving the previous maximization problem.

Lemma 1. *The criterion obtained by considering the worst case scenario for the change-time τ, where τ is any $\{\mathcal{F}_t^z\}$-adapted stopping time is*

$$\mathcal{J}_{\mathrm{M}}(T) = \sup_{\tau} \mathbb{E}_\tau[T-\tau|T>\tau; \tau<\infty] = \sup_{t \ge 0} \operatorname{ess\,sup} \mathbb{E}_t[T-t|T>t; \mathcal{F}_t^z]. \quad (16)$$

Proof. It is easy to see from (14) that

$$\begin{aligned}
\mathcal{J}(T) &\le \max\left\{ \frac{\mathbb{E}_\infty[\mathbb{E}_0[T|\mathcal{F}_0^z]\mathbb{1}_{\{\tau \le 0\}}]}{\mathbb{E}_\infty[\mathbb{1}_{\{\tau \le 0\}}]}, \sup_{t \ge 1} \frac{\mathbb{E}_\infty[\mathbb{E}_\infty[\mathbb{E}_0[(T-t)^+|\mathcal{F}_t^z \cup \mathcal{F}_t^\xi]|\mathcal{F}_t^z]\mathbb{1}_{\{\tau=t\}}]}{\mathbb{E}_\infty[\mathbb{P}_\infty[T>t|\mathcal{F}_t^z]\mathbb{1}_{\{\tau=t\}}]} \right\} \\
&\le \sup_{t \ge 0} \operatorname{ess\,sup} \frac{\mathbb{E}_\infty[\mathbb{E}_0[(T-t)^+|\mathcal{F}_t^z \cup \mathcal{F}_t^\xi]|\mathcal{F}_t^z]}{\mathbb{P}_\infty[T>t|\mathcal{F}_t^z]} \\
&= \sup_{t \ge 0} \operatorname{ess\,sup} \frac{\mathbb{E}_t[(T-t)^+|\mathcal{F}_t^z]}{\mathbb{P}_\infty[T>t|\mathcal{F}_t^z]} = \sup_{t \ge 0} \operatorname{ess\,sup} \mathbb{E}_t[T-t|T>t; \mathcal{F}_t^z].
\end{aligned}$$

The first and second inequality are rather straightforward. The first equality is an immediate outcome of the model we use for the change-point pdf and finally the last equality is again obvious. We can also see that it is always possible to select a stopping time τ that can put all its stopping probability on the event that attains the upper bound. Consequently there always exists an $\{\mathcal{F}_t^z\}$-adapted stopping time τ that attains the supremum. This concludes the proof. □

3 Special cases

The criterion introduced in (16) is novel and, in a sense, combines the two well known criteria of Lorden (1971) and Pollak (1985). What is also interesting and must be emphasized, is that this criterion can be used by the Statistician to evaluate the detection rule. Although the criterion involves conditioning with respect to Nature's information, for the evaluation we only need complete knowledge of the pre- and post-change joint statistics. Let us now focus on several interesting special cases.

Nature and Statistician access independent information: As a first example consider the case where the two sequences $\{z_t\}$ and $\{\xi_t\}$ are independent under both measures $\mathbb{P}_i, i = 0, \infty$. In other words Nature and the Statistician access information which is independent from each other before and after the change. It is then easy to see that (14) takes the special form

$$\mathcal{J}_{\mathrm{S}}(T) = \frac{\mathbb{E}_0[T]\pi_0 + \sum_{t=1}^{\infty} \mathbb{E}_t[(T-t)^+]\pi_t}{\pi_0 + \sum_{t=1}^{\infty} \mathbb{P}_\infty[T>t]\pi_t}.$$

where $\pi_0 = \mathbb{P}_\infty[\tau \le 0]$ and $\pi_t = \mathbb{P}_\infty[\tau = t],\ t \ge 1$. In other words we recover Shiryaev's (1978) Bayesian performance measure with π_t expressing the prior probability of having a change at time t.

By remaining within the same special class of independent processes, let us additionally assume that the sequence of probabilities $\{\pi_t\}_{t \ge 0}$ is unknown. In this case we would be interested in defining a performance measure by considering the worst possible scenario for this sequence, namely

$$\mathcal{J}_{\mathrm{P}}(T) = \sup_{\{\pi_i\}} \frac{\mathbb{E}_0[T]\pi_0 + \sum_{t=1}^{\infty} \mathbb{E}_t[(T-t)^+]\pi_t}{\pi_0 + \sum_{t=1}^{\infty} \mathbb{P}_\infty[T>t]\pi_t} = \sup_{t \ge 0} \frac{\mathbb{E}_t[(T-t)^+]}{\mathbb{P}_\infty[T>t]} = \sup_{t \ge 0} \mathbb{E}_t[T-t|T>t],$$

which yields Pollaks's (1985) performance measure. The same result can be obtained directly from (16) where, if we assume that T is independent from $\{z_t\}$, the criterion becomes independent from $\mathcal{F}_t^z$ and we recover Pollak's measure.

Nature and Statistician share the same information: Suppose now that Nature and the Statistician access exactly the same information. Then $\mathcal{F}_t^z = \mathcal{F}_t^\xi$ and since $\{T > t\}$ is $\mathcal{F}_t^\xi$-measurable, we can write $\mathbb{E}_t\left[T - t|T > t; \mathcal{F}_t^z\right] = \mathbb{E}_t[(T-t)^+|\mathcal{F}_t^\xi]$, suggesting that the worst-case performance criterion in (16) takes the special form of Lorden's (1971) measure

$$\mathcal{J}_{\mathrm{L}}(T) = \sup_{t\geq 0} \operatorname{ess\,sup} \mathbb{E}_t[(T-t)^+|\mathcal{F}_t^\xi] = \sup_{t\geq 0} \operatorname{ess\,sup} \mathbb{E}_0[(T-t)^+|\mathcal{F}_t^\xi].$$

Nature accesses more information than Statistician: Here we analyze the special case $\mathcal{F}_t^\xi \subseteq \mathcal{F}_t^z,\ t \geq 0$, in other words, Nature has access to more information than the Statistician, which is probably the most practically interesting scenario. Similarly to the previous case we have $\mathbb{E}_t\left[T - t|T > t; \mathcal{F}_t^z\right] = \mathbb{E}_t[(T-t)^+|\mathcal{F}_t^z]$, because $\{T > t\}$ is now $\mathcal{F}_t^z$-measurable as well, consequently

$$\mathcal{J}_{\mathrm{L}}(T) = \sup_{t\geq 0} \operatorname{ess\,sup} \mathbb{E}_t[(T-t)^+|\mathcal{F}_t^z] = \sup_{t\geq 0} \operatorname{ess\,sup} \mathbb{E}_0[(T-t)^+|\mathcal{F}_t^z].$$

One might argue that this is the same as Lorden's measure. However this is not exactly correct because conditioning is with respect to the σ-algebra $\mathcal{F}_t^z$ which is larger than $\mathcal{F}_t^\xi$. It is of course a Lorden-like criterion.

Nature accesses less information than Statistician: This is the opposite of the previous case, that is, $\mathcal{F}_t^\xi \supseteq \mathcal{F}_t^z,\ t \geq 0$. In other words, Nature accesses less information than the Statistician. Observing that $\mathbb{E}_i[X_t|\mathcal{F}_t^z \cup \mathcal{F}_t^\xi] = \mathbb{E}_i[X_t|\mathcal{F}_t^\xi]$ and using elements from the proof of Lemma1, the general criterion can be treated as follows

$$\mathcal{J}_{\mathrm{M}}(T) = \sup_{t\geq 0} \operatorname{ess\,sup} \frac{\mathbb{E}_\infty[\mathbb{E}_0[(T-t)^+|\mathcal{F}_t^z \cup \mathcal{F}_t^\xi]|\mathcal{F}_t^z]}{\mathbb{P}_\infty\left[T > t|\mathcal{F}_t^z\right]} = \sup_{t\geq 0} \operatorname{ess\,sup} \frac{\mathbb{E}_\infty[\mathbb{E}_0[(T-t)^+|\mathcal{F}_t^\xi]|\mathcal{F}_t^z]}{\mathbb{P}_\infty\left[T > t|\mathcal{F}_t^z\right]}.$$

Other possibilities: It is also possible to consider various combinations of the previous special cases. For example we can have independent informations before the change and Nature accessing more (or less) information than the Statistician after the change, or the opposite.

4 Discussion

The approach we followed here differs from the one adopted in Moustakides (2008). This is the reason why we obtain a number of interesting alternative measures. The main conclusion is that Shiryaev's and Pollak's measure rely on the assumption that Nature and the Statistician use completely independent information. Lorden on the other hand requires the two informations to coincide. Of course there is a variety of interesting application where the two information processes lie between these two extremes, that is, Nature and the Statistician use partially overlapping or simply dependent information.

The ultimate goal for introducing these alternative criteria is to show that the sequential change detection problem accepts a rich variety of mathematical formulations, that can describe it in more detail by taking into account the corresponding data scenarios. It is basically anticipated that with at least one of these new setups we might have a better chance in solving the change detection problem for dependent data. As we recall, with the existing min-max formulations and criteria no (nonasymptotically) optimum solution is currently available.

References

Lorden, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Stat.* **42**, 1897–1908.

Moustakides, G. V. (2008). Sequential change detection revisited. *Ann. Statist.* **36**, 787–807.

Pollak, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.* **13**, 206–227.

Shiryaev, A. N. (1978). *Optimal Stopping Rules.* Springer, New York.