Bayesian sequential hypothesis testing in discrete time

Yuqiong Wang | *yuqiong.wang@math.uu.se*
Joint work with Erik Ekström

Department of Mathematics
Uppsala University

June 30, 2021

Outline

Motivation

Testing unknown parameter with a Bernoulli prior: (Shiryaev, 1978)

- Observe a sequence of i.i.d. random variables distributed with density $p_\theta(x)$
- $\theta \in \{0, 1\}$ (Bernoulli prior)
- Can formulate stopping problem in terms of Π, the posterior probability process (Markovian)
- $V(\pi) = \inf_\tau \mathbb{E}_\pi[\Pi_\tau \wedge (1 - \Pi_\tau) + c\tau]$
- $\implies V(\pi)$ is concave
- $\implies$ There exists constant stopping boundaries

Motivation

Testing the unknown drift of a Brownian motion: (Ekström and Vaicenavicius, 2015)

- $dX_t = Bdt + dW_t$, B is a r.v.
- B has a general prior μ
- $V(0,\pi) = \inf_\tau \mathbb{E}_\pi[\Pi_\tau \wedge (1-\Pi_\tau) + c\tau]$
- $\implies V(\pi)$ is concave
- $\implies$ Volatility of Π is non-increasing in time
- $\implies$ There exists monotone stopping boundaries

Motivation

Questions

- Can we test other distributions in discrete time? e.g. unknown variance of a Gaussian?
- Does the problem exhibit similar structural properties?

Problem Setting

- The tester observes $X_1, X_2, \dots$ sequentially with cost c at each step
- X_k's are drawn from a one-parameter exponential family depending on r.v. Θ
- Conditioning on $\Theta = u$, X_k's are independent, and

$$\mathbb{P}(X_k \in A | \Theta = u) = \int_A p_u(x) \nu(dx)$$

where

$$p_u(x) := \exp\{ux - B(u)\},$$

and ν is a σ-finite measure on $\mathbb{R}$.

Bayesian set-up

- μ: prior of the unknown parameter Θ, denote the support of μ by S
- Want to test:

$$\begin{aligned} H_0: &\quad \Theta \leq \theta_0, \\ H_1: &\quad \Theta > \theta_0, \end{aligned}$$

- Let $d = i$ represents H_i is accepted,
- Define the cost function

$$V := \inf_{\tau \in \mathscr{T}} \inf_{d \in \mathscr{D}^\tau} \{\mathbb{P}(d = 1, \Theta \leq \theta_0) + \mathbb{P}(d = 0, \Theta > \theta_0) + c\mathbb{E}[\tau]\}$$

Reformulation in the π coordinate

- Define the posterior probability process Π

$$\Pi_n := \mathbb{P}(\Theta > \theta_0 | \mathscr{F}_n^X),$$

with $\Pi_0 = \pi$,

- Given $\tau \in \mathscr{T}$,

$$d = \begin{cases} 0 & \text{if } \Pi_\tau \leq 1/2 \\ 1 & \text{if } \Pi_\tau > 1/2, \end{cases}$$

- Consequently,

$$V = \inf_{\tau \in \mathscr{T}} \mathbb{E}\left[\Pi_\tau \wedge (1 - \Pi_\tau) + c\tau\right],$$

Properties of the Π process

- At time n, given $X_1 = x_1, \ldots, X_n = x_n$, by independence,

$$\mathbb{P}(\Theta > \theta_0 | X_1 = x_1, \ldots, X_n = x_n) = \frac{\int_{S^+} \prod_{i=1}^n p_u(x_i) \mu(du)}{\int_S \prod_{i=1}^n p_u(x_i) \mu(du)} = \frac{\int_{S^+} \exp\{u \sum_{i=1}^n x_i - nB(u)\} \mu(du)}{\int_S \exp\{u \sum_{i=1}^n x_i - nB(u)\} \mu(du)}.$$

- Denoting $Y_n := \sum_{i=1}^n X_i$:

$$\Pi_n = q(n, Y_n).$$

where

$$q(n, y) := \frac{\int_{S^+} e^{uy - nB(u)} \mu(du)}{\int_S e^{uy - nB(u)} \mu(du)}.$$

Parameterization of the posterior distribution

Denote by

$$\mu_{n,y}(du) := \frac{e^{uy-nB(u)}\mu(du)}{\int_S e^{uy-nB(u)}\mu(du)}$$

the posterior distribution of Θ at time n conditional on $Y_n = y$.

Lemma

The function $y \mapsto q(n,y) : \mathbb{R} \to (0,1)$ is an increasing bijection for each fixed n.

Remark

- Π *is a Markov process*
- *y can take any value in $\mathbb{R}$*
- *At time n, knowing y gives all the information of the posterior.*
- *Refer the set $\{(n, y(n,\pi)), n \geq 0\}$ as the $\pi-$level curve.*

Properties of the value function

The optimal stopping problem can be written as

$$V(n,\pi) = \inf_{\tau \in \mathscr{T}} \mathbb{E}_{n,\pi}[\Pi_{\tau+n} \wedge (1-\Pi_{\tau+n}) + c\tau]. \quad (1)$$

Lemma

The value function $V(n,\pi)$ satisfies

$$V(n-1,\pi) = \min\{\pi \wedge (1-\pi), c + \mathbb{E}_{n-1,\pi}[V(n,\Pi_n)]\}.$$

Lemma

Let $f : [0,1] \to [0,\infty)$ be a concave function. Then $\pi \mapsto \mathbb{E}_{n,\pi}[f(\Pi_{n+1})]$ is concave on $(0,1)$.

Proof. By implicit differentiation.

Concavity and its consequence

Main Theorem (1)

The function $\pi \mapsto V(n,\pi)$ is concave for each fixed $n \geq 0$.

Remark. V can be extended for every $\pi \in [0,1]$ and the concavity is preserved.

Concavity and its consequence

Introduce now

- The continuation region $\mathscr{C}$:

$$\mathscr{C} := \{(n,\pi) \in \mathbb{N}_0 \times [0,1] : V(n,\pi) < \pi \wedge (1-\pi)\},$$

- The stopping region $\mathscr{D}$ by

$$\mathscr{D} := \{(n,\pi) \in \mathbb{N}_0 \times [0,1] : V(n,\pi) = \pi \wedge (1-\pi)\}.$$

The stopping time

$$\tau^* := \inf\{k \geq 0 : (n+k, \Pi_{n+k}) \in \mathscr{D}\}$$

is an optimal strategy.

Remark

- *The continuation region is of the form* $(b_1(n), b_2(n))$
- *The concavity result is connected with time-monotonicity of the value function*

The posterior distribution gradually squeezes in

Main Theorem (2)

If $a < \theta_0 < b$*, then*

$$n \mapsto \mathbb{P}_{n,\pi}(\Theta \le a) \quad \& \quad n \mapsto \mathbb{P}_{n,\pi}(\Theta > b)$$

are decreasing.

Remark

As a consequence, let $0 < \pi_1 < \pi_2 < 1$*. Then*

$$n \mapsto y(n, \pi_2) - y(n, \pi_1)$$

is non-decreasing.

Conditions for monotonicity in time

Assumption

For any $\pi \in (0,1)$ and $n \geq m \geq 0$, the random variable $\Pi_{m+1}|\{\Pi_m = \pi\}$ dominates $\Pi_{n+1}|\{\Pi_n = \pi\}$ in convex order.

Theorem

Assume the above holds. Then $V(n,\pi)$ is non-decreasing in n, and the boundaries b_1 and b_2 are thus non-decreasing and non-increasing, respectively.

But is the assumption correct?

One sufficient condition

If there exists a π_0, such that Π_{n+1} is more concentrated around it than Π_{m+1}, then we are done.

Main Theorem (3)

- *Observations are continuously distributed with density $h(x)p_u(x)$ s.t. $I := \{h > 0\}$ is an interval.*
- *$h(x)p_u(x)$ is increasing in x on I, and S^+ is a singleton*

then $V(n,\pi)$ is non-decreasing in n.

Remark. The following cases also go through

- The symmetric case when $h(x)p_u(x)$ is decreasing in x on I, and S^- is a singleton.
- When $I := \{h > 0\}$ is a union of disjoint intervals.

Examples of Main theorem 3

Example 1. (Exponential observations)

$$h(x)p_u(x) = \begin{cases} \exp\{ux + \log u\} & x < 0 \\ 0 & x \geq 0 \end{cases}$$

Example 2. (Gaussian observations with unknown variance)

$$h(x)p_u(x) = \begin{cases} \frac{2}{\sqrt{-\pi x}} \exp\left\{ux + \frac{1}{2}\log u\right\} & x < 0 \\ 0 & x \geq 0 \end{cases}$$

Does it hold for other distributions?

For arbitrary priors:

- **Example 3. (Gaussian observations with unknown mean)**

$$h(x)p_u(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2}} \exp\{ux - B(u)\}.$$

Time-monotonicity follows from the continuous case.

- **Example 4. (Bernoulli observations)**

$$h(x)p_{\hat{u}}(x) = \exp\{\hat{u}x - \log(1+\hat{u})\},$$

Convex order of Π_n can be shown.

- **Example 5. (Binomial observations)**
  Can be regarded as modification of the Bernoulli case.

Further discussion

Conjecture

The function $V(n, \pi)$ is non-decreasing in n.

Questions and suggestions

Thank you!