Optimal stopping in mathematical statistics with applications to finance

Peter Johnson

School of Mathematics, The University of Manchester

9 Feb 2016

Sequential testing.

Problem.

Say we observe diffusion process $X = (X_t)_{t \geq 0}$ such that X solves the following SDE

$$dX_t = \Big(\mu_0 + \theta\big(\mu_1 - \mu_0\big)\Big)dt + \sigma dB_t,$$

with $X_0 = 0$.

However, θ is an unobservable random variable that takes either value 1 or 0 with probability π or $(1 - \pi)$ respectively. It is also assumed to be independent of B_t.

Then through observation of X we wish to determine the value of θ as quickly and accurately as possible.

Hypothesis test.

This can be seen as a hypothesis testing problem on the value of θ.

$$
\begin{aligned}
H_0: &\quad \theta = 0, \quad X_t = \mu_0 t + \sigma B_t, \\
H_1: &\quad \theta = 1, \quad X_t = \mu_1 t + \sigma B_t
\end{aligned}
$$

To solve this problem we can formalise it as an optimal stopping problem.

Defining the probability measure.

Firstly we can define a measure P_π such that

$$P_\pi = \pi P^1 + (1-\pi)P^0$$

where under $P^1(\theta = 1) = 1$ and $P^0(\theta = 0) = 1$.

Formulation of problem.

In order to find the optimal time to come to a decision on the value of θ the problem can be formulated as an optimal stopping problem where we wish to minimise the following costs;

- Expected time to come to a decision

$$E_\pi[\tau]$$

- Expected costs of a wrong terminal decision

$$aP_\pi(d=0,\theta=1)+bP_\pi(d=1,\theta=0)$$

Hence this can be written as the following optimal stopping problem

$$V(\pi)=\inf_{(\tau,d)}\left(E_\pi(\tau)+aP_\pi(d=0,\theta=1)+bP_\pi(d=1,\theta=0)\right)$$

Traditional formulation.

This formulation may be rewritten in terms of the *a posteriori process*, defined as

$$\pi_t = P_\pi(\theta = 1 | F_t^X), \quad \pi_0 = \pi.$$

Becoming of the form

$$V(\pi) = \inf_\tau E_\pi \left[\tau + a\pi_\tau \wedge b(1 - \pi_\tau) \right]$$

where

- $d = 1$ if $\pi_\tau > b/(a+b)$
- $d = 0$ if $\pi_\tau < b/(a+b)$

Likelihood ratio

The a posteriori process can be defined via the likelihood ratio process L_t in the one-to-one function

$$\pi_t = \frac{\frac{\pi}{1-\pi} L_t}{1 + \frac{\pi}{1-\pi} L_t}$$

where L_t is defined as the Girsanov measure change from P^1 to P^0 given to be

$$L_t = \exp\left(\frac{\mu_1 - \mu_0}{\sigma^2} X_t - \frac{\mu_1^2 - \mu_0^2}{2\sigma^2} t \right)$$

SDE of a posteriori process.

The a posteriori process can be seen to solve the following SDE

$$d\pi_t = \frac{\mu_1 - \mu_0}{\sigma}\pi_t(1-\pi_t)d\bar{B}_t$$

where

$$\bar{B}_t = \frac{1}{\sigma}\left(X_t - \int_0^t (\mu_0 + \pi_s(\mu_1 - \mu_0))\, ds\right)$$

Reduction to ODE.

$$\frac{(\mu_1-\mu_0)^2}{2\sigma^2}\pi^2(1-\pi)^2\frac{d^2}{d\pi^2}V(\pi)=-1 \quad \textit{for} \quad \pi\in(A,B),$$

$$V(A)=aA$$

$$V(B)=b(1-A)$$

$$V'(A)=a \quad (smoothfit)$$

$$V'(B)=-b \quad (smoothfit)$$

$$V<a\pi\wedge b(1-\pi) \quad \textit{for} \quad \pi\in(A,B)$$

$$V=a\pi\wedge b(1-\pi) \quad \textit{for} \quad \pi\in[0,A)\cup(B,1]$$

Solution to ODE.

Letting

$$\psi(\pi) = (1-2\pi)log(\frac{\pi}{1-\pi})$$

then we have that

$$V(\pi) = \frac{2\sigma^2}{\mu^2}(\psi(\pi) - \psi(A)) + (a - \frac{2\sigma^2}{\mu^2}\psi'(A))(\pi - A) + aA$$

if $\pi \in (A, B)$ and

$$V(\pi) = a\pi \wedge b(1-\pi)$$

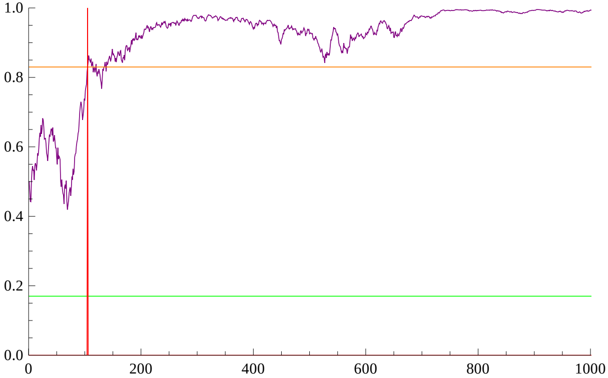
if $\pi \in [0, A) \cup (B, 1]$. While $A \in (0, c)$ and $B \in (c, 1)$ are given as the solution of the following transcendental equations:

$$V(B; A) = b(1-B)$$

$$V'(B; A) = -b$$

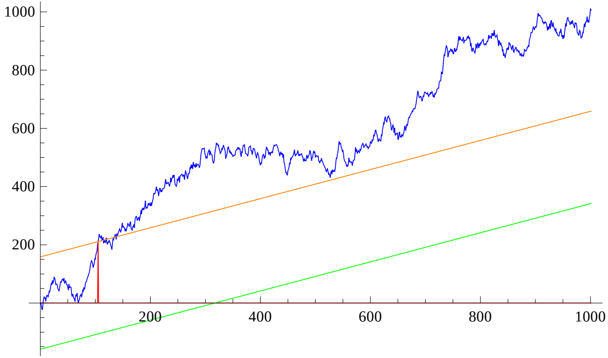
Pictures.

Out[3034]=



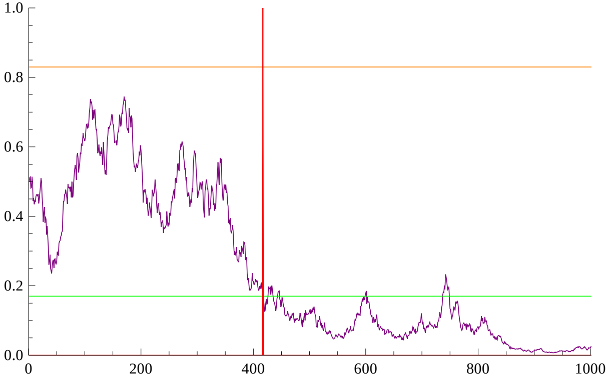
Pictures.

Out[3035]=



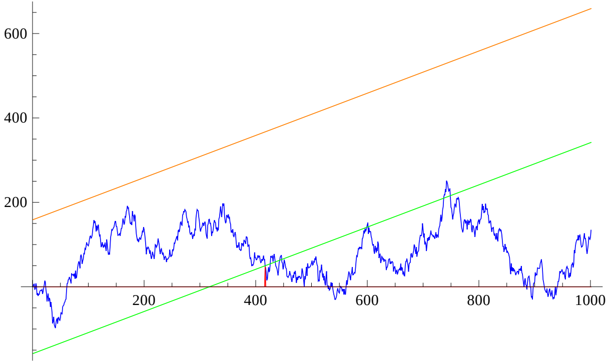
Pictures.

Out[3106]=



Pictures.

Out[3107]=



Bayesian quickest detection.

Problem.

Say we observe diffusion process $X = (X_t)_{t \geq 0}$ such that X solves the following SDE

$$dX_t = \Big(\mu_0 + I(t \geq \Theta)(\mu_1 - \mu_0)\Big) dt + \sigma dB_t,$$

with $X_0 = 0$.

However, Θ is an unobservable random variable which takes the value $\Theta = 0$ with probability π and is exponentially distributed with parameter λ given $\Theta > 0$ with probability $(1 - \pi)$.
It is also assumed to be independent of B_t.

Then through observation of X we wish to determine when the change point Θ has occurred, as quickly and accurately as possible .

Bayesian quickest detection.

So to reiterate before the change point the observed process behaves like

$$X_t = \mu_0 t + \sigma B_t,$$

and after the change point Θ it behaves like

$$X_t = \mu_1 t + \sigma B_t,$$

To solve this problem we can again formalise it as an optimal stopping problem.

Defining the probability measure.

Firstly we can define a measure P_π such that

$$P_\pi = \pi P^0 + (1-\pi)\int_0^\infty \lambda e^{-\lambda s} P^s ds$$

where under $P^s(\Theta = s) = 1$.

Formulation of the problem.

In order to find the optimal stopping time the problem can be formulated as an optimal stopping problem where we wish to minimise the following;

- Probability of false alarm

$$P_{\pi}(\tau < \Theta)$$

- The delay taken to come to a decision after Θ

$$E_{\pi}[(\tau - \Theta)^{+}]$$

Hence this can be written as the following optimal stopping problem

$$V(\pi) = \inf_{\tau} \left(P_{\pi}(\tau < \Theta) + cE_{\pi}[(\tau - \Theta)^{+}] \right)$$

Traditional formulation.

This formulation may be rewritten in terms of the *a posteriori process*, defined as

$$\pi_t = P_\pi(\Theta \leq t | F_t^X), \quad \pi_0 = \pi.$$

Becoming of the form

$$V(\pi) = \inf_\tau E_\pi \left[1 - \pi_\tau + c \int_0^\tau \pi_t dt \right]$$

Likelihood process.

The a posteriori process can be defined via the likelihood ratio process φ_t in the one-to-one function

$$\pi_t = \frac{\varphi_t}{1+\varphi_t}$$

where φ_t is defined by

$$\varphi_t = e^{\lambda t} L_t \left(\frac{\pi}{1-\pi} + \lambda \int_0^t \frac{e^{-\lambda s}}{L_s} ds \right)$$

and as before L_t is defined by the Girsanov measure change given to be

$$L_t = \exp\left(\frac{\mu_1 - \mu_0}{\sigma^2} X_t - \frac{\mu_1^2 - \mu_0^2}{2\sigma^2} t \right)$$

SDE of posteriori process.

The a posteriori process can be seen to solve the following SDE

$$d\pi_t = \lambda(1-\pi_t)dt + \frac{\mu_1 - \mu_0}{\sigma}\pi_t(1-\pi_t)d\bar{B}_t$$

where again we define

$$\bar{B}_t = \frac{1}{\sigma}\left(X_t - \int_0^t (\mu_0 + \pi_s(\mu_1 - \mu_0))\, ds\right)$$

Reduction to ODE.

$$\mathbb{L}_\pi V(\pi) = -c\pi \quad \textit{for} \quad \pi \in [0, A),$$

$$V(A) = 1 - A,$$

$$V'(A) = -1, \quad (\textit{smoothfit})$$

$$V(\pi) < 1 - \pi \quad \textit{for} \quad \pi \in [0, A),$$

$$V(\pi) = 1 - \pi \quad \textit{for} \quad \pi \in [A, 1],$$

where

$$\mathbb{L}_\pi = \lambda(1-\pi)\frac{d}{d\pi} + \frac{(\mu_1 - \mu_0)^2}{2\sigma^2}\pi^2(1-\pi)^2\frac{d^2}{d\pi^2}.$$

Solution to ODE.

Letting

$$\gamma = \frac{\mu^2}{2\sigma^2},$$

$$\psi(\pi) = -\frac{c}{\gamma} e^{-\frac{\lambda}{\gamma}\alpha(\pi)} \int_0^\pi \frac{e^{\frac{\lambda}{\gamma}\alpha(\rho)}}{\rho(1-\rho)^2} d\rho$$

such that A is the unique solution of

$$\psi(A) = -1.$$

Then the value function is explicitly given by

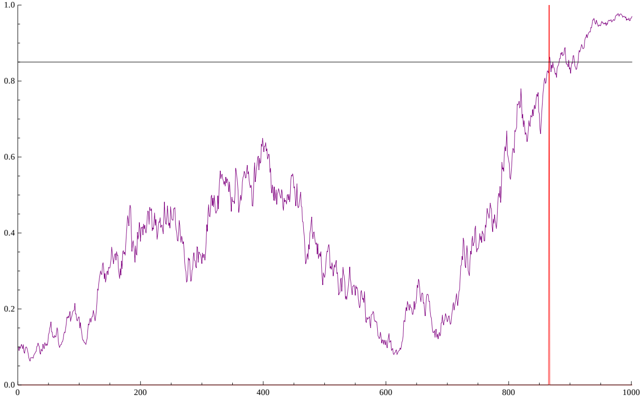
$$V(\pi) = \begin{cases} (1-A) + \int_A^\pi \psi(\rho) d\rho & \text{if } \pi \in [0, A) \\ 1-\pi & \text{if } \pi \in [A, 1] \end{cases}$$

and the optimal stopping time is defined to be

$$\tau = \inf\{t \geq 0 : \pi_t \geq A\}.$$

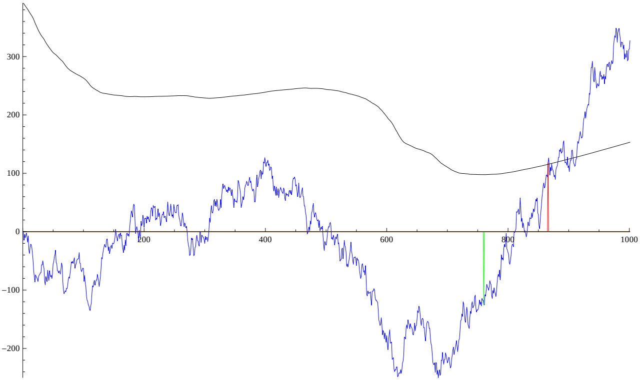
Pictures.

Out[2126]=



Pictures.

Out[2127]=



This type of problem can be tackled for more general diffusions, where the drift and diffusion coefficients may depend the current position of the observed process, I.e. for sequential testing we observe

$$dX_t = (\mu_0(X_t) + \theta(\mu_1(X_t) - \mu_0(X_t)))dt + \sigma(X_t)dB_t$$

or in a quickest detection setting we have

$$dX_t = (\mu_0(X_t)+I(\Theta \geq t)(\mu_1(X_t) - \mu_0(X_t)))dt + \sigma(X_t)dB_t$$

A very important quantity in both these sets of problems is known as the signal-to-noise ratio (SNR) defined by

$$\rho(x) = \frac{\mu_1(x) - \mu_0(x)}{\sigma(x)}$$

If this quantity is constant then the problems can be solved in a similar way to the canonical Brownian motion with drift case.

However, if the SNR is non-constant is then the problems become much more difficult to solve.

In this case the sufficient statistic is now the pair (X_t, π_t), which makes the problem two dimensional (much harder as now the boundaries will also be a function of the position of the observed process X_t).

Recent developments have seen the first examples of these problems being solved for a process with non-constant SNR, in particular a Bessel process with unknown/changing dimension (see references).

Applications.

Pairs trading strategy.

"The strategy monitors performance of two historically correlated securities. When the correlation between the two securities temporarily weakens, i.e. one stock moves up while the other moves down, the pairs trade would be to short the outperforming stock and to long the underperforming one, betting that the "spread" between the two would eventually converge. The divergence within a pair can be caused by temporary supply/demand changes, large buy/sell orders for one security, reaction for important news about one of the companies, and so on."

Pairs trading strategy.

The difference between the two stocks if often modelled using a Ornstein-Uhlenbeck process.

$$dX_t = \alpha(\beta - X_t)dt + \sigma dB_t.$$

However if the stocks become ‘uncoupled’ then this difference will lose its mean-reversion property so changing to a simple Brownian motion.

Quickest detection of this change-point would reduce losses incurred due to continuing a pairs trading strategy when the stocks are uncoupled.

Failure to detect such a uncoupling, caused in part by the 1998 Russian financial crisis, cost ‘Long-term capital management’ (whose directors included Scholes and Merton) around $150 million before its ultimate collapse.

Detection of arbitrage.

Stocks are often modelled using geometric Brownian motion solving

$$dX_t = \mu_0 X_t dt + \sigma X_t dB_t$$

which is then used to price derivatives under the risk-neutral measure such that $\mu_0 = r$.

A detection of a change in the drift rate away from the risk-free rate, $\mu_1 \neq r$, would represent an arbitrage opportunity using derivatives priced in this manner.

Updating model parameters.

Bessel processes have been used to model financial bubbles through considering an observed process of $1/X_t$ such that

$$dX_t = \frac{\delta_0 - 1}{2X_t} dt + dB_t$$

However if you can detect a change of dimension in the observed data you are modelling then this can be quickly integrated into your model for a more accurate description and response of the appearance of the financial bubble.

Other applications

- Seismic activity.
- Psychological testing.
- Quality control.
- Communications.
- Radar detection.
- Detection of instability in power systems.
- Sonar detection.
- Detecting breakages in atomic clocks in space.
- Detecting radioactive materials.

References

PESKIR, G. *and* SHIRYAEV, A. (2006). *Optimal Stopping and Free-Boundary Problems*. Birkhuser.

SHIRYAEV, A. N. (1978). *Optimal Stopping Rules*, Springer-Verlag.

GAPEEV, P. *and* SHIRYAEV, A. (2011). On the Sequential testing Problem for some Diffusion Processes, *Stochastics: An International Journal of Probability and Stochastic Processes*, Vol. 83, No. 4-6, (519–535)

GAPEEV, P. V. *and* SHIRYAEV, A. N. (2013). Bayesian quickest detection problems for some diffusion processes, *Adv. in Appl. Probab.*, Vol. 45, (164–185).

References

- JOHNSON, P. J. *and* PESKIR, G. (2014). Sequential testing problems for Bessel Processes, *Research Report*, No. 24, *Probab. Statist. Group Manchester* (to appear).
- JOHNSON, P. J. *and* PESKIR, G. (2014). Quickest detection problems for Bessel Processes, *Research Report*, No. 25, *Probab. Statist. Group Manchester* (to appear).
- SHIRYAEV, A. N. (2010). Quickest detection problems: Fifty years later, *Sequential Anal.*, Vol. 29, (345–385).
- BUONAGUIDI, B. *and* MULIERE, P. (2012). A note on some sequential problems for the equilibrium value of a Vasicek process, *Pioneer Journal of Theoretical and Applied Statistics*, Vol. 4, No. 2, (101–116).

Thank you for listening.