Change-point Problems
IMS Lecture Notes – Monograph Series (Volume 23, 1994)

BAYES PROBLEMS IN CHANGE-POINT MODELS FOR THE WIENER PROCESS

BY M. BEIBEL
University of Freiburg

We consider the continuous time change-point model of Shiryayev (1973) and study its structure. We also derive an asymptotic expansion of the Bayes risk when the costs of observation become small. This approach is then extended to related other problems.

1. Introduction and General Setup. A problem of sequential detection in continuous time is considered. Let B denote a standard Brownian motion. Let θ be a fixed and known real number. We study the process

$$W_t = B_t + \theta(t-\tau)^+ \quad t \in [0,\infty) ,$$

where τ is an unknown change-point in a Bayesian framework. This means that τ is a random variable with a known law. Following Shiryayev (1973) we assume that τ is independent of B and has a prior distribution with $P(\{\tau = 0\}) = p$ and $P(\{\tau > t\}) = (1-p)e^{-\lambda t}$, where $p \in [0,1)$ and $\lambda > 0$.

The statistician observes W and seeks for a stopping time T depending on W that detects τ as soon as possible. Here a stopping time means that

$$\{T \leq t\} \in \mathcal{F}_t^W = \sigma(W_s; s \leq t) .$$

The quality of a stopping time T is measured by the following risk function $R(T)$:

$$R(T) = P(T < \tau) + CE(T-\tau)^+ ,$$

where $C > 0$. We seek for a Bayes solution T^*, that is a stopping time T^* with

$$R(T^*) = \inf_T R(T) .$$

The infimum on the right hand side is taken over all stopping times T.

AMS 1991 Subject Classification: Primary 62L15; Secondary 62C10.

Key words and phrases: Change-point, sequential detection, Bayes problems, Brownian motion.

Let π_t denote a continuous version of the posterior probability (which exists in the situation we consider) that the change has already occurred given the observations up to time t

$$\pi_t := P(\tau \leq t|\mathcal{F}_t^W) .$$

For all finite stopping times T holds

$$R(T) = E\left\{1 - \pi_T + C\int_0^T \pi_s ds\right\} .$$

It turns out (see Shiryayev (1973)) that in the above problem T^* is of the form

$$T^* := \inf\{t > 0|\ \pi_t \geq p^*\} ,$$

where $p^* \in (0,1)$ is properly chosen. The key step to prove this without using results of the general theory of optimal stopping is to find a function g for which $g(x) + 1 - x$ assumes its minimum over $[0,1)$ uniquely at p^* such that for a sufficiently large class of stopping times T holds

$$R(T) = E\{g(\pi_T) + (1 - \pi_T)\} .$$

Since π_t is continuous in t one can stop exactly in the minimum. This method is well adapted to optimality proofs in sequential testing even in the case of composite hypotheses (see Lerche (1985)). A similar approach can be applied to the continuous time version of the Bayes problems of Ritov (1990) (see Beibel (1992)).

2. Shiryayev's Problem – Exact Solution. In Shiryayev (1973) Theorem 1 below is proved using the general theory of optimal stopping of Markov processes. Shiryayev considered a certain generalized Stefan problem. The solution of this free-boundary problem is time independent. Therefore it is possible to carry out our approach. Our arguments are modifications of Shiryayev's original reasoning.

π_t satisfies the equation (see Shiryayev (1973))

$$d\pi_t = \lambda(1 - \pi_t)dt + \theta\pi_t(1 - \pi_t)d\bar{W}_t ,$$

where $\pi_0 = p$ and $\bar{W}$ is a standard Brownian motion.

Let $g(\cdot)$ denote the function $-\int_p^{\cdot} \psi^*(u)du$, where ψ^* is defined as in Shiryayev (1973, p.163 equation (4.172)). Then g solves the differential equation

$$\frac{\theta^2}{2}x^2(1-x)^2g''(x) + \lambda(1-x)g'(x) = Cx \quad \text{for } x \in (0,1) .$$

THEOREM 1. (Shiryayev) *Let p^* denote the unique solution in $(0,1)$ of $g'(x) = 1$. Then*

$$T^* := \inf\{t > 0 |\ \pi_t \geq p^*\}$$

is a Bayes solution.

PROOF. Ito's formula implies that for all bounded stopping times T holds

$$R(T) = E\{g(\pi_T) + 1 - \pi_T\}\ .$$

A straightforward calculation shows that g is convex and that $g(x) + 1 - x$ assumes its minimum over $[0,1)$ at a unique point $p^* \in (0,1)$.

If $\pi_0 = p < p^*$ we immediately obtain the optimality of T^*. An easy limit argument yields that for all finite stopping times T holds $R(T) \geq g(p^*)+1-p^*$. π_t converges almost sure to one as t tends to infinity and thus $T^* < \infty$ and $\pi_{T^*} = p^*$. Another limit argument yields now $R(T^*) = g(p^*) + 1 - p^*$.

In the case $p > p^*$ one starts beyond the minimum and the best strategy is to stop immediately.

3. Shiryayev's Problem for Small Costs. Theorem 1 unfortunately gives no explicit expression for the optimal threshold p^*. The implicit equation $g'(p^*) = 1$ can only be solved numerically. Hence we will now study the behaviour of $R(T^*)$ and p^* in Shiryayev's problem for fixed λ and θ as the costs tend to zero. We now write c instead of C. We like to indicate the dependence of $R(\cdot)$, T^* and p^* on c and write $R_c(\cdot)$, T_c^* and p_c^*.

Using elementary calculus to study the behaviour of $g(x)$ for $x \to 1$ one can show that

$$\frac{1}{1-p_c^*} = \frac{\tilde{I}}{c} + \left(1 - \frac{2\lambda}{\theta^2}\right) + o(1) \ \text{ as } \ c \to 0\ ,$$

where $\tilde{I} = \lambda + \frac{\theta^2}{2}$. A similar expansion can be obtained for $R_c(T_c^*)$. The analytical arguments provide no probabilistic interpretation of the terms in the asymptotic expansions which they yield. We now give a stochastic argument which perhaps reveals more what happens for $c \to 0$.

The stochastic differential equation for π_t implies that

$$-d\log(1-\pi_t) = \lambda dt + \frac{\theta^2}{2}\pi_t^2 dt + \theta\pi_t d\bar{W}_t\ .$$

Then for all stopping times T with $ET < \infty$ it holds

$$R_c(T) = E\{\tilde{g}(\pi_T) + 1 - \pi_T\} + \tilde{c}EV_T - \tilde{g}(p)\ ,$$

where $\tilde{c} = \frac{c}{\tilde{I}}$, $\tilde{g}(x) = \tilde{c}\log\frac{1}{1-x} - \tilde{c}x$ and $V_t = \frac{\theta^2}{2}\int_0^t (\pi_s - \pi_s^2)ds$.
Obviously $V_t \uparrow V_\infty$ as $t \to \infty$ for some random variable V_∞ with $EV_\infty < \infty$. The function $\tilde{g}(x) + 1 - x$ is convex on $[0,1)$ and assumes its unique minimum over $[0,1)$ at $x = (\tilde{c}+1)^{-1}$. Let

$$T_c := \inf\left\{t > 0 \,\middle|\, \pi_t \geq \left(\tfrac{c}{\tilde{I}} + 1\right)^{-1}\right\} .$$

For $c \to 0$ it holds $T_c^* \to \infty$ P-a.s.. This yields $E(V_{T_c} - V_{T_c^*}) \to 0$ as $c \to 0$. Together with some additional arguments we obtain from that

$$0 \leq \frac{1}{c}\Big(R_c(T_c) - R_c(T_c^*)\Big) \leq \frac{1}{\tilde{I}}\Big\{EV_{T_c} - EV_{T_c^*}\Big\} \to 0 \text{ as } c \to 0 .$$

Thus T_c is asymptotically optimal and it holds

Proposition 1.

$$\begin{aligned} R_c(T_c^*) &= R_c(T_c) + o(c) \\ &= \tilde{c}\log\frac{1}{\tilde{c}} + \tilde{c}\left\{EV_\infty - \log\frac{1}{1-p} + p\right\} + o(c) \text{ as } c \to 0 . \end{aligned}$$

The asymptotic expression in Proposition 1 for the minimal Bayes risk $R_c(T_c^*)$ has an important consequence. If we rewrite $R_c(\cdot)$ as

$$R_c(T) = P(T < \tau) + \tilde{c}\tilde{I}E(T-\tau)^+$$

we can see, that taking the costs for the delay proportional to $\tilde{I}$, standardises the problems with different values of θ such that they are asymptotically for small costs in first order of equal difficulty. This unsual cost structure $(\lambda + \frac{\theta^2}{2})(T-\tau)^+$ for the delay term raises the question, wether it is possible to modify the problem such that $\frac{\theta^2}{2}(T-\tau)^+$ is the correct standardisation. This might be more natural, because $\frac{\theta^2}{2}$ is the Kullback-Leibler-Information of a normal distribution with mean θ and variance 1 relative to a normal distribution with mean 0 and variance 1. Let P_∞ denote the measure corresponding to no change at all. Let

$$L_c(T) = P_\infty(T < \infty) + c\frac{\theta^2}{2}E(T-\tau)^+ .$$

A continuous version of $\frac{dP}{dP_\infty}\Big|_{\mathcal{F}_t^W}$ is given by ψ_t with

$$\psi_t := (1-p)e^{-\lambda t}(1-\pi_t)^{-1} .$$

ψ_t satifies the equation

$$d\log(\psi_t) = \frac{\theta^2}{2}\pi_t^2 dt + \theta\pi_t d\bar{W}_t \ ,$$

where $\psi_0 = 1$ and thus for all stopping times T with $ET < \infty$ it holds

$$L_c(T) = E\Big\{\psi_T^{-1} + c\log(\psi_T) + cV_T\Big\} \ .$$

Let $S_c := \inf\Big\{t\Big|\psi_t \geq c^{-1}\Big\}$ for $c > 0$. Let S_c^* denote a Bayes solution corresponding to $L_c(\cdot)$. The existence of S_c^* can be shown using results of the general theory of optimal stopping of Markov processes. Similar arguments as before yield for $c \to 0$ the following result.

PROPOSITION 2.

$$\begin{aligned} L_c(S_c^*) &= L_c(S_c) + o(c) \\ &= c + c\log\frac{1}{c} + cEV_\infty + o(c) \quad \text{as} \quad c \to 0 \ . \end{aligned}$$

It is interesting to compare this statement with the exact results in the problem of optimal power one tests for simple alternatives (see Lerche (1985)). One can write the risk function considered there in a way similar to $L_c(\cdot)$. Then the minimal risk is equal to $c + c\log(1/c)$ and for a large class of stopping times T the risk is the expectation of $1/l_T + c\log l_T$, where l_t is the likelihood ratio at time t of Brownian motion with drift θ to Brownian motion with drift 0.

4. The Case of Unknown θ for Small Costs. We finish with a short remark about the case of small costs and unknown θ for the loss structures $R(\cdot)$ and $L(\cdot)$. We therefore like to indicate the dependence of the distribution of W on θ and write in the sequel $P^{(\theta)}$ and $E^{(\theta)}$ instead of P and E. The arguments of Section 3 on asymptotic optimality can be extended to handle the case when θ is unknown and has a prior distribution. We take the costs proportional to $\lambda + \frac{\theta^2}{2}$ or $\frac{\theta^2}{2}$ for $R(\cdot)$ or $L(\cdot)$ and consider

$$\bar{R}_c(T) = P_\infty(T < \tau) + c\int_{-\infty}^{+\infty}\Big(\lambda + \frac{1}{2}\theta^2\Big)E^{(\theta)}(T-\tau)^+ G(d\theta)$$

or

$$\bar{L}_c(T) = P_\infty(T < \infty) + c\int_{-\infty}^{+\infty}\frac{1}{2}\theta^2 E^{(\theta)}(T-\tau)^+ G(d\theta) \ ,$$

where G denotes a normal distribution function. We obtain similar results as Lerche (1985). An additional term of order $\frac{c}{2}\log\log\frac{1}{c}$ plus a term c times certain constants appear in the asymptotic expressions for the minimal Bayes risks as $c \to 0$. These terms result from not knowing θ and when the change occurs.

REFERENCES

BEIBEL, M. (1992). Bayes and minimax properties of the cusum procedure in a change-point model for the Wiener process. To appear in *Ann. Statist.*

LERCHE, H. R. (1985). On the optimality of sequential tests with parabolic boundaries. In *Proc. of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer.* (L. LeCam and R. Olshen, eds.) **2**, 579–597. Wadsworth, Monterey, CA.

RITOV, Y. (1990). Decision theoretic optimality of the cusum procedure. *Ann. Statist.*, **18**, 1464–1469

SHIRYAYEV, A. N. (1973). *Statistical sequential analysis.* Translations of Math. Monographs, vol. **8**, American Math. Society

UNIVERSITÄT FREIBURG
INSTITUT FÜR MATHEMATISCHE STOCHASTIK
HEBELSTR. 27
79104 FREIBURG
GERMANY