Sequential Procedure for Testing Hypothesis about Mean of Latent Gaussian Process

Julia Bondarenko

Helmut-Schmidt University Hamburg
(University of the Federal Armed Forces Hamburg)
Holstenhofweg 85, 22043 Hamburg, Germany
bonda@hsu-hh.de

Abstract

Sequential Probability Ratio Test (SPRT) has been widely used to detect process anomalies. The purpose of this paper is to present a practical procedure on the basis of SPRT, which recognizes if the mean of a latent Gaussian process $Y_t$ significantly deviates from presumed value, via an observable signal resulting from $Y_t$. An equation for estimating the unknown nuisance parameter of the process is obtained, values of likelihood ratio statistic and thresholds are determined. The empirical analysis of the procedure performance is conducted.

Keywords: Sequential Probability Ratio Test; Maximum Likelihood Estimate; Nuisance Parameter; Average Sample Number

1 Introduction

A. Wald [11] has developed his famous sequential probability ratio test (SPRT) - a sequential extension of the Neyman-Pearson Lemma, which implies that likelihood ratio test gives the best result in fixed size samples. Let $Y_1, Y_2, \ldots$ be an infinite sequence of the random variables with values in some data space $Y$, and let $\mu$ be a sigma-finite measure on $Y$. Then $(Y_1, Y_2, \ldots, Y_n) \in Y^n$ and $\mu_n = \mu^n$ (product measure) is sigma-finite on $Y^n$. Let $f(Y_1, Y_2, \ldots, Y_n)$, $n = 1, 2, \ldots$, denote the density function of $(Y_1, Y_2, \ldots, Y_n)$ with respect to $\mu_n$ on $Y^n$. In the case of parameterized family of probability density functions, we can also write $f(Y_1, Y_2, \ldots, Y_n; \theta)$. This includes the continuous case (where $\mu_n$ is the Lebesque measure and $f(Y_1, Y_2, \ldots, Y_n; \theta)$ is a Lebesque density, or joint probability density function - PDF), the discrete case (where $\mu_n$ is the counting measure and $f(Y_1, Y_2, \ldots, Y_n; \theta)$ is a probability function - PF), as well as more complex models.

Let us assume first that we want to test two simple hypotheses, $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$, $\theta_0 \neq \theta_1$. The corresponding densities are $f(Y_1, Y_2, \ldots, Y_n; \theta_0) =$
We consider the first $n-1$ random variables (observations) $Y_1, Y_2, \ldots, Y_{n-1}$. When the new observation $Y_n$ arrives, the SPRT decides whether to reject the alternative hypothesis, reject the null hypothesis, or continue sampling. The test involves the likelihood ratio

$$
\Lambda_n = \frac{f_1(Y_1, Y_2, \ldots, Y_n)}{f_0(Y_1, Y_2, \ldots, Y_n)} = \frac{L(Y_1, Y_2, \ldots, Y_n; H_1)}{L(Y_1, Y_2, \ldots, Y_n; H_0)}
$$

where $n = 1, 2, \ldots, L(Y_1, Y_2, \ldots, Y_n; H_j)$ is the likelihood value under hypothesis $j$, $j = \{0; 1\}$.

The decision rule defined by SPRT splits the sample space $Y^n$ into three subsets called acceptance, rejection and continuation/indifference regions. In fixed size samples, accordingly to the Neyman-Pearson Lemma, the likelihood ratio statistic is compared to certain threshold value, which determines accuracy. By analogy, we select two thresholds $A$ and $B$ (which are in general depending on $n$), $0 < B < A < \infty$, and draw a comparison between them and our likelihood ratio at each consecutive observation. If a threshold is reached or crossed, a decision regarding the null or alternative hypothesis is reached as well, that is: if $\Lambda_n \leq B$ - terminate and reject $H_1$ (assuming that either $H_0$ or $H_1$ is true, it can be also interpreted as ”terminate and accept $H_0$” - acceptance region); if $\Lambda_n \geq A$ - terminate and reject $H_0$ (rejection region); if $B < \Lambda_n < A$ - continue sampling since our hope for a reduced sample size $i$ is not justified and we need more observations for decision (continuation/indifference region).

Constants $A$ and $B$ are related to the error rates $\alpha$ (probability of eventually rejecting $H_0$ when it is true, $\alpha = P[H_1| H_0] = P_0[H_1]$ - error of Type I, or the false alarm probability) and $\beta$ (probability of eventually accepting $H_0$ when it is false, $\beta = P[H_0| H_1] = P_1[H_0]$ - error of Type II, or missed alarm probability). The probabilities $\alpha$ and $\beta$ reflect the ”costs” of making these two types of errors. Note, that the Neyman-Pearson Lemma provides the test with maximal power $1 - \beta$, where the above-mentioned threshold value is chosen by setting the false alarm probability $\alpha$.

Provided that the procedure terminates with probability 1, Wald has also shown, that $\alpha$ and $\beta$ and the bounds $A$ and $B$ satisfy the following inequalities ([11], p. 43; [8]; [2]): $B \geq \frac{\beta}{1-\alpha}$, $A \leq \frac{1-\beta}{\alpha}$, see Appendix 2 for the proof. In general, these inequalities need not to hold with equalities because by termination $\Lambda_n$ might fall not exactly in $B$ or $A$. But usually, under large $n$, the value of $\Lambda_n$ is only slightly smaller than $B$ or larger than $A$. Therefore, Wald’s approximation replaces the inequalities by equalities, for obtaining in practice the thresholds for desired $\alpha = \alpha'$ and $\beta = \beta'$, see, for example, [8], p. 11: $B' = \frac{\beta'}{1-\alpha'}$, $A' = \frac{1-\beta'}{\alpha'}$. Note, that a special case of SPRT is the open-ended test, when the test statistic $\Lambda_n$ excludes the lower bound $B$.

Going back to the above-mentioned Neyman-Pearson Lemma, let’s remind that to provide the most powerful test for a fixed samples one needs to fix a
false alarm probability (test size) $\alpha$, that is, the constraints here are a sample size and $\alpha$. The Wald-Wolfowitz Theorem ([12]) cited below states, that the SPRT in testing simple $H_0$ vs simple $H_1$ hypotheses based on the i.i.d. observations, which operates with infinite sample, can present even better result. Really, additionally to the desired error probabilities $\alpha$ and $\beta$, playing a role of constraints in this case, it gives also the smaller average sample number (ASN) $E[n]$ under both $H_0$ and $H_1$. We remind briefly, that ASN is the average number of observations required for decision making on step $n$ under assumption that $H_0$ or $H_1$ is true. Hence, by implementing the SPRT to fixed sample the sampling procedure may be terminated even before the entire sample is considered.

**Theorem 1 (Wald and Wolfowitz).** Among all tests (fixed sample or sequential) for which $P[H_1 | H_0] \leq \alpha$ and $P[H_0 | H_1] \leq \beta$ and for which $E[n | H_j] < \infty$, $j = \{0; 1\}$, $H_j$ are simple, the SPRT with error probabilities $\alpha$ and $\beta$ minimizes both $E[n | H_0]$ and $E[n | H_1]$, under i.i.d.-assumption.

For the case of non-i.i.d. observations the optimality of the SPRT in the sense of the Wald-Wolfowitz Theorem, is in general not ensured, although the asymptotic optimality under small error probabilities and some mild constraint (stability of the process $\{\log \Lambda_n\}$) is fulfilled.

Note, that by the Wald and Wolfowitz Theorem, the optimality of SPRT implies that the expected sample size needed is smallest only when $\theta = \theta_0$ or $\theta = \theta_1$. But for values $\theta$ between $\theta_0$ and $\theta_1$, the SPRT may perform quite unsatisfactory: the expected sample sizes tend to be large and the decision about hypothesis acceptance is delayed, hence, optimality doesn’t hold here and SPRT may require a sample size even larger than a non-sequential test. It gave rise to consider the following optimization problem: one tests $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$ such that the expected sample size for $\theta = \theta'$, $\theta_0 < \theta' < \theta_1$, is minimized subject to error probabilities at two other points, $\theta_0$ and $\theta_1$, that is, $P[H_1 | H_0] \leq \alpha$ and $P[H_0 | H_1] \leq \beta$. This problem was solved by Kiefer and Weiss [4], and also extended in later works.

In most realistic situations the hypotheses to be tested are composite, and the simple hypotheses are only a particular case of them. The composite hypotheses can be written as

$$H_0 : \theta < \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0 \, (\theta > \theta_1), \, \theta_0 < \theta_1,$$

with Type I and Type II error probabilities not exceeding $\alpha$ and $\beta$ and indifference zone (i.e., zone separating zero and alternative hypotheses) $(\theta_0, \theta_1)$. For fixed sample size, the composite hypotheses may be converted to simple ones under monotone likelihood ratio property. Unfortunately, it is not true in
the sequential case ([1], p. 54). There are numerous approaches to the testing of such kind of hypotheses, but an obvious optimality criterion doesn’t exist here. For example, Wald’s proposal was to transform the problem to testing between two simple hypotheses $K_0 : \theta = \theta_0$ vs $K_1 : \theta = \theta_1$ by means the method of ”weight functions” - some sort of prior distribution ([11]). The sequential testing solution for composite hypotheses based upon Bayesian strategy is considered also in [6]. A detailed review of methods and approaches for sequential testing of composite hypotheses is given in ([5], [7]).

In this work, we consider a sequential test of composite hypotheses about parameter of (weakly) stationary gaussian latent process. The outline of the paper is as follows. In Section 2, we introduce our testing problem for parameter of interest, present a nuisance parameter eliminating procedure and derive the test thresholds. Section 3 shows the empirical results of the problem solution and concludes with a few remarks.

2 Sequential Testing Procedure

2.1 Problem Formulation

Let $Y_1, Y_2, ..$ be some underlying unobservable (hidden) time series (realization of $Y_t$) supposed to be independent normally distributed random variables, $Y_t \sim N(\mu, \sigma^2)$, for each $t = 1, 2, ...$, where $\mu > 0$ and $\sigma^2$ are unknown. Assume that if $Y_t > \theta$, where $\theta = const$, then the signal is observed. Without losing generality, we may take $\theta = 0$. That is, we shall concerned then with the nonnegative random variables

$$X_t = \max(0, Y_t), \quad i = 1, 2, ... .$$

(3)

The purpose is to detect as quickly as possible if the value of parameter of interest $\mu$ differs from a known value $\mu_0$, in the presence of nuisance parameter $\sigma^2$. We can formulate the following, analogous to (2), hypotheses:

$$H_0 : \mu < \mu_0 \ vs \ H_1 : \mu > \mu_0 ,$$

(4)

with false and missed alarm probabilities $\alpha$ and $\beta$, correspondingly. Similar problem statement can arise in medical, socio-economic and physical fields, for example, with application to financial decision-making (in particular, option pricing problem), accidents monitoring, disease control, seismic signal processing, etc.

In order to construct a likelihood ratio (1) corresponding to the testing problem (4), we need, in the first instance, to derive the distribution of the random variable $X$, under all values $\mu$. The distribution of $X$ can be easily obtained as follows. Let’s replace the zero value in (3) for every $t$ with a random variable $Z$, normally distributed with parameters 0 and $\tau^2$, where
\( \tau \to 0, Z \sim N(0, \tau^2) \), and independent of \( Y \). Then the distribution and density functions of the random variable \( X_t = \max(Z_t, Y_t) \) are, correspondingly, 

\[
F_X(x; \mu, \sigma^2, \tau^2) = P(Z < x, Y < x) = F_Z(x; 0, \tau^2) F_Y(x; \mu, \sigma^2), \quad f_X(x; \mu, \sigma^2, \tau^2) = F_Z(x; 0, \tau^2) f_Y(x; \mu, \sigma^2) + F_Y(x; \mu, \sigma^2) f_Z(x; 0, \tau^2).
\]

**Notation 2** Another possibility would be to assign a uniform distribution to the variable \( Z \), i.e., \( Z \sim U(-\tau, \tau) \), where \( \tau \to 0 \).

### 2.2 Nuisance Parameter Estimation

Elimination of the nuisance parameter \( \sigma^2 \) still remains an open question. The algorithm that produces approximate maximum likelihood estimator for parameter \( \sigma^2 \) is similar to the Stein two-stage procedure [9]. A training set of data (first \( M \) observations \( X_1, ..., X_M \)) has been used to construct the estimate \( \hat{\sigma}^2 \).

**Notation 3** Alternatively, one can consider "purely sequential" procedure for the problem solution, based on the Adaptive Sequential Probability Ratio Test (ASPRT) (see [10]). ASPRT employs estimates of the unknown nuisance parameters (parameter \( \sigma^2 \) in our case), which involve only the data up to the previous time point.

First, we specify the maximum (log)likelihood equations:

\[
\begin{align*}
\frac{\partial l}{\partial \mu}(X_1, ..., X_M | \mu, \sigma^2, \tau^2) &= 0, \\
\frac{\partial l}{\partial \sigma}(X_1, ..., X_M | \mu, \sigma^2, \tau^2) &= 0,
\end{align*}
\]

where

\[
l(X_1, ..., X_M | \mu, \sigma^2, \tau^2) = \ln L(X_1, ..., X_M | \mu, \sigma^2, \tau^2) = \\
\sum_{i=1}^{M} \ln (F_1(X_i) f_2(X_i) + F_2(X_i) f_1(X_i)),
\]

\[
F_1(x) = F_Z(x; 0, \tau^2), \quad f_1(x) = f_Z(x; 0, \tau^2), \quad F_2(x) = F_Y(x; \mu, \sigma^2), \quad f_2(x) = f_Y(x; \mu, \sigma^2).
\]

Let \( \eta \) be a parameter, \( \eta = \{\mu, \sigma\} \), then we denote

\[
S = \frac{\partial l}{\partial \eta}(X_1, ..., X_M | \mu, \sigma^2, \tau^2) = \sum_{i=1}^{M} \frac{F_1(X_i) \frac{\partial f_2(X_i)}{\partial \sigma} + \frac{\partial F_2(X_i)}{\partial \mu} f_1(X_i)}{F_1(X_i) f_2(X_i) + F_2(X_i) f_1(X_i)}.
\]

Passing to the limit under \( \tau \to 0 \), one can distinguish two cases: 1) \( X_i \leq 3\tau \) and 2) \( X_i > 3\tau \). The sum \( S \) is partitioned into two smaller sums: the first one, \( S_1 \), involves only \( K \) summands with \( X_i \leq 3\tau \), the second one, \( S_2 \), \( (M - K) \) summands with \( X_i > 3\tau \), \( S = S_1 + S_2 \). In first case we can see that \( f_1(X_i) \to \infty \) and \( F_1(X_i) \to \frac{1}{2} \), and therefore
the system of equations (6) is transformed in the following one:

\[ S_1 = \sum_{i=1}^{K} \frac{f_i(X_i)}{F_i(X_i)} \frac{\partial^2 f_i(X_i)}{\partial \eta^2} f_i(X_i) = \sum_{i=1}^{K} \frac{f_i(X_i)}{F_i(X_i)} \frac{\partial f_i(X_i)}{\partial \eta} \frac{\partial f_i(X_i)}{\partial \eta} f_i(X_i) \approx \sum_{i=1}^{K} \frac{f_i(X_i)}{F_i(X_i)} \frac{\partial f_i(X_i)}{\partial \eta} (0) \]

\[ S_2 = \sum_{i=1}^{M-K} \frac{f_i(X_i)}{F_i(X_i)} \frac{\partial^2 f_i(X_i)}{\partial \eta^2} f_i(X_i) = \sum_{i=1}^{M-K} \frac{f_i(X_i)}{F_i(X_i)} \frac{\partial f_i(X_i)}{\partial \eta} (0) \]

In respect that

\[ \frac{\partial f_2}{\partial \mu} (X_i) = \frac{X_i - \mu}{\sigma} f_2 (X_i) , \]

\[ \frac{\partial f_2}{\partial \sigma} (X_i) = \left( -\frac{1}{\sigma} + \frac{(X_i - \mu)^2}{\sigma^2} \right) f_2 (X_i) , \]

\[ \frac{\partial^2 f_2}{\partial \sigma^2} (0) = -\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{0} (s - \mu) \exp \left( \frac{(s - \mu)^2}{2\sigma^2} \right) ds , \]

\[ \frac{\partial^2 f_2}{\partial \mu} (0) = -\frac{1}{\sigma} f_2 (0) + \frac{1}{\sqrt{2\pi\sigma^4}} \int_{-\infty}^{0} (s - \mu)^2 \exp \left( \frac{(s - \mu)^2}{2\sigma^2} \right) ds , \]

the system of equations (6) is transformed in the following one:

\[
\begin{cases}
K \exp \left( \frac{-\mu^2}{2\sigma^2} \right) = \frac{1}{\sigma} \sum_{X_i>0} X_i - (M-K) \frac{\mu}{\sigma} , \\
K \mu \exp \left( \frac{-\mu^2}{2\sigma^2} \right) = M - K - \frac{1}{\sigma^2} \sum_{X_i>0} X_i^2 + \frac{2}{\sigma^2} \mu \sum_{X_i>0} X_i - \frac{M-K}{\sigma^2} \mu^2 ,
\end{cases}
\]

where \( \Phi(x) = \frac{1}{\sqrt{2\pi}} \exp \left( \frac{-x^2}{2} \right) \) is a standard normal distribution function. Multiplying the first equation in (7) by \( \frac{\mu}{\sigma} \) and subtracting it from the second one, we obtain that \( \mu = \frac{\sum_{X_i>0} X_i^2 - (M-K) \sigma^2}{\sum_{X_i>0} X_i} \). We substitute then \( \mu \) into first equation and resolve it with respect to \( \sigma \).

**Notation 4** Method of moments is another way to the solution of parameter estimation problem.

### 2.3 Sequential Likelihood Ratio Test Statistic and Boundaries

For practical purposes, we can replace the composite hypotheses by simple ones (see [11], p. 71). Let’s choose two constants \( \epsilon_1 \) and \( \epsilon_2 \), \( \epsilon_1 > 0 \), \( \epsilon_2 > 0 \), and consider two simple hypotheses
known variance. We can also write (9) in the following form
\[ l \left( X_{M+1}, \ldots, X_n \mid \mu_0 + \epsilon_2, \sigma^2 \right) \approx l \left( X_{M+1}, \ldots, X_n \mid \mu_0 - \epsilon_1, \sigma^2 \right) \]

as
\[ \Lambda_n = \frac{l \left( X_{M+1}, \ldots, X_n \mid \mu_0 + \epsilon_2, \sigma^2 \right)}{l \left( X_{M+1}, \ldots, X_n \mid \mu_0 - \epsilon_1, \sigma^2 \right)} \sum_{t=M+1}^{n} \ln \frac{F_1(X_t)}{F_1(X_t)} f_2(X_t; \mu_0 + \epsilon_2) + F_2(X_t; \mu_0 + \epsilon_2) f_1(X_t) \]
\[ f_2(X_t; \mu_0 - \epsilon_1) + F_2(X_t; \mu_0 - \epsilon_1) f_1(X_t) \],

where \( l(\cdot) \) is a log-likelihood function, \( F_1(x) \equiv F_Z(x; 0, \tau^2) \), \( f_1(x) \equiv f_Z(x; 0, \tau^2) \), \( F_2(x; s) \equiv F_Y(x; s, \sigma^2) \), \( f_2(x; s) \equiv f_Y(x; s, \sigma^2) \) and \( \sigma^2 \) is an estimate of unknown variance. We can also write (9) in the following form
\[ \Lambda_n = \Lambda_{n-1} + C_n \]

where \( C_n = \ln \frac{F_1(X_t) f_2(X_t; \mu_0 + \epsilon_2) + F_2(X_t; \mu_0 + \epsilon_2) f_1(X_t)}{F_1(X_t) f_2(X_t; \mu_0 - \epsilon_1) + F_2(X_t; \mu_0 - \epsilon_1) f_1(X_t)} \). The test involves comparing the value \( \Lambda_n \) to thresholds \( b = \log \left( \frac{\beta}{1-\alpha} \right) \) and \( a = \log \left( \frac{1-\beta}{\alpha} \right) \), where \( \alpha \) and \( \beta \) are the error probabilities of Type I and Type II, respectively. If \( \Lambda_n \leq b \), we stop the test and accept \( H_0 \). If \( \Lambda_n \geq a \), we stop and declare \( H_1 \). Otherwise, no decision is made and we keep on testing.

As before, we distinguish two cases: 1) \( X_n \leq 3\tau \) and 2) \( X_n > C \). In second case we have that \( f_1(X_n) \to 0 \) and \( F_1(X_n) \to 1 \), and therefore, from (10),
\[ \Lambda_n \approx \Lambda_{n-1} + \frac{1}{\sigma^2} \left( \epsilon_2 + \epsilon_1 \right) \left( X_n - \mu_0 \right) - \frac{1}{\sigma^2} \left( \epsilon_2 - \epsilon_1 \right) \]

that corresponds to the normal case. More detailed, \( C_n = \frac{1}{\sigma^2} \left( \epsilon_2 + \epsilon_1 \right) \left( X_n - \mu_0 \right) - \frac{1}{\sigma^2} \left( \epsilon_2 - \epsilon_1 \right) \).

The first case gives us \( f_1(X_n) \to 0 \) \( \to \infty \) and \( F_1(X_n) \to 1 \) under \( \tau \to 0 \).

One obtains here \( C_n = \ln \frac{f_1(0)[F_1(0)f_2(0; \mu_0 + \epsilon_2) + F_2(0; \mu_0 + \epsilon_2)]}{f_1(0)[F_1(0)f_2(0; \mu_0 - \epsilon_1) + F_2(0; \mu_0 - \epsilon_1)]} \approx \ln \frac{F_2(0; \mu_0 + \epsilon_2)}{F_2(0; \mu_0 - \epsilon_1)} \), thus,
Finally, we can write our likelihood ratio statistic in a slightly different form:

\[
\Lambda^*_n = \Lambda^*_{n-1} + \begin{cases} 
\frac{1}{\sigma^2} (\epsilon_1 + \epsilon_2) & \text{if } X_n > 3\tau, \\
0 & \text{if } X_n \leq 3\tau.
\end{cases}
\]  

(13)

with upper and lower bounds

\[
\begin{cases} 
U^*_n = U^*_{n-1} + \frac{1}{2\sigma^2} (\epsilon_2^2 - \epsilon_1^2) + \frac{\mu_0}{\sigma^2} (\epsilon_2 + \epsilon_1) & \text{and} \\
L^*_n = L^*_{n-1} + \frac{1}{2\sigma^2} (\epsilon_2^2 - \epsilon_1^2) + \frac{\mu_0}{\sigma^2} (\epsilon_2 + \epsilon_1) & \text{if } X_n > 3\tau,
\end{cases}
\]

(14)

\[
\begin{cases} 
U^*_n = U^*_{n-1} - \ln \frac{F_2(0; \mu_0 + \epsilon_2)}{F_2(0; \mu_0 - \epsilon_1)} & \text{and} \\
L^*_n = L^*_{n-1} - \ln \frac{F_2(0; \mu_0 + \epsilon_2)}{F_2(0; \mu_0 - \epsilon_1)} & \text{if } X_n \leq 3\tau,
\end{cases}
\]

(15)

where \(\Lambda^*_M = 0, U^*_M = a, L^*_M = b\).

3 Empirical Results and Future Research Direction

Here we illustrate a SPRT-procedure performance comparison for different values of the mean parameter \(\mu\), standard deviation parameter \(\sigma\) and indifference zone width. The boundaries defined by \(\epsilon_1\) and \(\epsilon_2\) in (8) are equal to \(\mu_0 - \epsilon\) and \(\mu_0 + \epsilon\), where \(\epsilon = \epsilon_1 = \epsilon_2\), \(\mu_0 = 0.5\). Error probabilities have been set to \(\alpha = \beta = 0.05\). We assume also that a set of training data of size \(M = 150\) for parameter \(\sigma^2\) estimation is available. Each set of parameters \((\mu, \sigma, \epsilon)\) implies simulation with 2000 replications. The reported values of power and ASN are presented in Figures 1-4.

In general, for fixed values \(\alpha \) and \(\beta\), an increase in \(\epsilon\)-value entails better results. The relatively small value of \(\epsilon\) produces high detection probability and short ASN only under small variance \(\sigma\). For large variance values, a certain hesitation between indifference and weak preference is illustrative. Note, that as the difference between \(\mu\) and \(\mu_0\) increases, performance of the procedure becomes not sensitive to the variance values: power converges to 1 and ASN go quickly down.

Notation 5 Average Sample Number can be also determined analytically as following (see, for instance, [3]):

\[
E(n|H_1) = \frac{E(\Lambda_n|H_1)}{E(C_t|H_1)},
\]

(16)

where \(C_t\) is the (log)likelihood ratio component corresponding to the \(t\)th observation. From the previous section,
Sequential procedure

Figure 1: Empirical Power and Average Sample Length, $\mu = 0.55$, $\mu_0 = 0.5$, $\alpha = \beta = 0.05$

Figure 2: Empirical Power and Average Sample Length, $\mu = 0.6$, $\mu_0 = 0.5$, $\alpha = \beta = 0.05$

Figure 3: Empirical Power and Average Sample Length, $\mu = 1.0$, $\mu_0 = 0.5$, $\alpha = \beta = 0.05$
Figure 4: Empirical Power and Average Sample Length, $\mu = 2.0$, $\mu_0 = 0.5$, $\alpha = \beta = 0.05$

$$C_t = \begin{cases} \frac{1}{\sigma^2} (\epsilon_2 + \epsilon_1) (X_t - \mu_0) - \frac{1}{2} \frac{(\epsilon_2^2 - \epsilon_1^2)}{\sigma^2}, & \text{if } X_t > 3\tau, \\ \ln \frac{F_2(0; \mu_0^0 + \epsilon_2)}{F_2(0; \mu_0 - \epsilon_1)}, & \text{if } X_t \leq 3\tau. \end{cases}$$

Computing an expectation of the random value $C_t$ under alternative hypothesis $H_1$ and keeping in mind that the numerator in (16) may be approximated as

$$E(\Lambda_n|H_1) = (1 - \beta) \ln(\frac{1-\beta}{\alpha}) + \beta \ln(\frac{\beta}{1-\alpha}),$$

we can obtain the theoretical results and compare them with empirical evidence. The ASN under null hypothesis $H_0$ are evaluated analogously.

Theoretical analysis of ASN proposed above is a subject tackled in future research. Moreover, the results shown here are also preliminary steps towards a complexification of the considered sequential problem: particularly, we are interested in solving change detection problems from option pricing when the underlying unobservable process $Y_t$ is a geometric Brownian motion process.

References


Received: April, 2010