The Annals of Applied Probability
1998, Vol. 8, No. 3, 749–774

A NEW REPRESENTATION FOR A RENEWAL-THEORETIC CONSTANT APPEARING IN ASYMPTOTIC APPROXIMATIONS OF LARGE DEVIATIONS[1]

BY BENJAMIN YAKIR AND MOSHE POLLAK

Hebrew University

The probability that a stochastic process with negative drift exceed a value a often has a renewal-theoretic approximation as $a \to \infty$. Except for a process of iid random variables, this approximation involves a constant which is not amenable to analytic calculation. Naive simulation of this constant has the drawback of necessitating a choice of finite a, thereby hurting assessment of the precision of a Monte Carlo simulation estimate, as the effect of the discrepancy between a and ∞ is usually difficult to evaluate.

Here we suggest a new way of representing the constant. Our approach enables simulation of the constant with prescribed accuracy. We exemplify our approach by working out the details of a sequential power one hypothesis testing problem of whether a sequence of observations is iid standard normal against the alternative that the sequence is AR(1). Monte Carlo results are reported.

1. Introduction. In many contexts, the probability α that the maximal value of a stochastic process exceed a prespecified value is a quantity of considerable importance. In risk theory it shows up as the probability of ruin, in sequential analysis it appears in the form of error probabilities, in options pricing it is the probability that an option will be exercised, in branching processes it is the probability that the population size be large. Its value is usually hard to fix precisely, and approximations are often called for. When the stochastic process under study is the sequence of partial sums of iid observations, renewal theory supplies practical formulas which in turn provide useful approximations. [For an overview, see Siegmund (1985).] Renewal theory has been developed for other processes, too—such as when the underlying observations are generated by a Markov chain [Kesten (1974)] or by a time series [Lalley (1986)]. However, in these cases the renewal-theoretic results are not as useful as in the iid case, for, although they provide limiting expressions which (if evaluated) could be used as approximations, these expressions contain constants which, in contrast to the iid case, are not amenable to calculation.

In this article we develop a different renewal-theoretic approximation for the probability α. Our approximation, too, contains a constant which cannot

---

Received August 1996; revised September 1997.
[1] Supported in part by grants of the Israel Science Foundation.
AMS 1991 *subject classifications*. Primary 60K05; secondary 62L10.
Key words and phrases. Overshoot, sequential test, time series.

be calculated analytically. However, this constant can be evaluated by Monte Carlo. In principle, the constants appearing in the standard renewal-theoretic form can also be evaluated by Monte Carlo. However, the standard representation suffers from difficulties involved in measuring the precision of the Monte Carlo estimate, as renewal theory involves crossing a barrier which tends to infinity, and in a given simulation it is not easy to evaluate the effect of the discrepancy between infinity and the (necessarily finite) barrier used. In contrast, the constant appearing in our representation does not involve a barrier tending to infinity and can be evaluated by Monte Carlo to any degree of prescribed accuracy.

In the following, we regard the sequential analytic problem of a power one test of hypotheses. We chose this problem to exemplify our approach because it is relatively simple in structure and because it is a basic underlying building block for calculating the ARL to false alarm in changepoint problems. We describe our approach in Section 2. To illustrate the considerations involved, we first consider an iid case of a power one test of a shift of a normal mean. Then we apply our method to testing a null hypothesis that a sequence of observations is iid standard normal against an alternative that the sequence is AR(1). Monte Carlo results are reported.

Both of the examples worked out in this paper can be interpreted as the probability of a stochastic process crossing a straight line boundary. With appropriate modifications, the approach can be applied to more complex problems such as repeated significance testing. These modifications entail nonnegligible technical considerations, the spelling out of which would make an already long paper even lengthier and would not add enough insight to the basic understanding of the approach to justify their inclusion.

2. A rule of thumb. The changepoint problem deals with monitoring a sequence of observations for a change from one probability regime to another. With this as background, we envision the following.

Let $\underline{X} = X_1, X_2, \dots$ be a sequence of observations. Let $\mathrm{P}_0, \mathrm{P}_1$ be probability measures for $\underline{X}$ which have the same support for each finite sequence $X_1, \dots, X_n$. A (usually power one) test of $H_0\colon \underline{X} \sim \mathrm{P}_0$ vs. $H_1\colon \underline{X} \sim \mathrm{P}_1$ is to reject the null hypothesis H_0 if $\max_{1 \le n < \infty} L_n \ge A$, where $L_n = d\mathrm{P}_1(X_1, \dots, X_n)/d\mathrm{P}_0(X_1, \dots, X_n)$ and A is a prespecified constant. The level of significance of this test is $\alpha = \mathrm{P}_0(\max_{1 \le n < \infty} L_n \ge A)$. By a well-known martingale argument [Ville (1939), page 100], $\alpha \le 1/A$. Renewal theory is often called upon to obtain an approximation $\alpha \approx \mathrm{const}/A$ as $A \to \infty$.

Let $\mathrm{P}^{(k)}$ be a measure under which $X_1, \dots, X_k$ behave according to P_1 and $X_{k+1}, X_{k+2}, \dots$ according to P_0 [in the sense that the distribution of $X_{k+1}, X_{k+2}, \dots$ conditional on $X_1, \dots, X_k$ is the same as it would have been had $X_1, \dots, X_k$ been distributed according to P_0 (and had attained the same values)].

Such a situation is easy to describe if X_i are iid under both P_0 and P_1, in which case $\mathrm{P}^{(k)}$ is the measure under which the first k observations are distributed $d\mathrm{P}_1$ and the ensuing $d\mathrm{P}_0$. Other cases of interest include a transition from one Markov chain to another, one time series model to another, and so on.

Note that

$$L_n/L_k = \begin{cases} d\mathrm{P}^{(n)}(X_1,\ldots,X_k)/d\mathrm{P}_1(X_1,\ldots,X_k), & \text{if } n \le k, \\ d\mathrm{P}_1(X_1,\ldots,X_n)/d\mathrm{P}^{(k)}(X_1,\ldots,X_n), & \text{if } n \ge k. \end{cases}$$

Define

$$M_k = \max_{1 \le n < \infty} (L_n/L_k),$$

$$S_k = \sum_{n=1}^{\infty} (L_n/L_k).$$

In most cases of interest (including all of those mentioned above) if $\mathrm{P}^{(k)}$ is the true distribution of the sequence $\underline{X}$, then M_k will be attained at a time n close to k. In fact, the order of magnitude of L_n/L_k will be exponential in $-|n-k|$, so that S_k will be finite. Furthermore, if k is large, L_n/L_k will contribute little to both M_k and S_k when $|n-k|$ is large. Therefore, (M_k, S_k) under $\mathrm{P}^{(k)}$ will have a limit in distribution (M, S) as $k \to \infty$. Since $M_k < S_k$, it follows that $M \le S$.

In most cases of interest (including all of those mentioned above) the limit

$$I = \lim_{n\to\infty} \frac{1}{n} \int \log L_n \, d\mathrm{P}_1$$

exists. Define

$$\lambda = \frac{1}{I}\mathrm{E}\frac{M}{S}. \tag{2.1}$$

Arguments in subsequent sections will justify the following:

Rule of thumb. In most cases of interest,

$$A\mathrm{P}_0\Big(\max_{1 \le n < \infty} L_n \ge A\Big) \to \lambda \qquad \text{as } A \to \infty \tag{2.2}$$

so that for large A,

$$a \approx \frac{\lambda}{A}.$$

We suggest that the representation (2.1) is useful from a practical point of view, since usually one wants to be able to evaluate λ to any desired degree of accuracy. The limit I can usually be obtained analytically. Although $\mathrm{E}(M/S)$ is in general analytically intractable, it is readily amenable to simulation: $0 < M/S \le 1$, and one can usually give theoretical bounds on the exponential decay of L_n/L_k (as $|n-k|$ becomes large). In the non-iid case, this is an improvement over the renewal-theoretic results available at present, which at best yield the existence of a limit λ in (2.2), but do not provide means for its calculation. Its approximation by Monte Carlo as $A \to \infty$ [by the left side of (2.2) or even by importance sampling [cf. Siegmund (1976)] leaves open the question "When is A large enough?," so that one doesn't have a handle on the accuracy of the simulation.

Insight into the relation between $\mathrm{E}(M/S)$ and the renewal-theoretic constant can be gained by noticing the following heuristic argument, which is due to Siegmund. Given a large m, the expression

$$\mathrm{E}^{(k)}\left[\frac{\max_{1\le n\le m} L_n}{\sum_{n=1}^m L_n}\right]$$

is (almost) independent of k. (Some dependence is introduced by negligible boundary effects.) It follows that

$$\begin{aligned}\mathrm{E}(M/S) &\sim \frac{1}{m}\sum_{k=1}^m \mathrm{E}^{(k)}\left[\frac{\max_{1\le n\le m} L_n}{\sum_{n=1}^m L_n}\right]\\ &= \frac{1}{m}\mathrm{E}_0\left[\max_{1\le n\le m} L_n\right],\end{aligned}$$

which is known to be asymptotically (as $m \to \infty$) equivalent to the traditional representation of the constant as the Laplace transformation of the overshoot. [See Hogan and Siegmund (1986), Lemmas 3.3 and 3.4.]

At first glance, our "rule of thumb" admittedly looks mysterious. The intuition behind (2.1) is basically technical. The classical renewal-theoretic argument regarding the asymptotics of the significance level of a sequential test entails a change-of-measure component. This enables studying a tail probability (of the original measure) in terms of the behavior of a central part of a distribution (the transformed measure). Therefore, it is natural to look for a suitable measure to which the problem under study can be transformed.

In the classical iid setting, the behavior of Z_n^+, the ladder increment of the log-likelihood process, is independent of $l_{n-1}^+ = \sum_{i=1}^{n-1} Z_i^+$. Hence, the distribution of the overshoot can be represented as a convolution of the renewal measure and the distribution of a single ladder increment of the log-likelihood process

$$\int_0^a \mathrm{P}_1(Z^+ \ge a + y - x)\left[\sum_{n=1}^\infty f_{l_{n-1}^+}(x)\right] dx, \tag{2.3}$$

where $f_{l_{n-1}^+}(x)$ is the density of l_{n-1}^+ and $y > 0$. This fails when the dependence structure is more complicated. Therefore, it is reasonable to look for an alternative to the renewal measure $[\sum_{n=1}^\infty f_{l_{n-1}^+}(x)]\,dx$, one which will enable a separation between local behavior and long-term characteristics of the process.

When studying the changepoint problem, a measure which shows up quite naturally is $\sum_{n=1}^\infty \mathrm{P}^{(n)}$ [cf. Yakir (1995)], whose likelihood ratio with respect to $\mathrm{P}^{(k)}$ is $\sum_{n=1}^\infty L_n/L_k$. Therefore, it is natural to attempt to use this measure to separate local behavior and long-term characteristics of the process. So

$$\begin{aligned}
P_0\Big(\max_{1\le n<\infty} L_n \ge A\Big) &= \sum_{k=1}^{\infty} E^{(k)}\left[\mathbb{1}\Big(\max_{1\le n<\infty} L_n \ge A\Big)\Big/\sum_{n=1}^{\infty} L_n\right] \\
&= \sum_{k=1}^{\infty} E^{(k)}\left[\frac{\max_{1\le n<\infty} L_n/L_k}{\sum_{n=1}^{\infty} L_n/L_k} \times \frac{\mathbb{1}(\max_{1\le n<\infty} L_n \ge A)}{\max_{1\le n<\infty} L_n}\right] \\
&\sim E\frac{M}{S} \times \int \frac{\mathbb{1}(\max_{1\le n<\infty} L_n \ge A)}{\max_{1\le n<\infty} L_n}\left[\sum_{k=1}^{\infty} dP^{(k)}\right],
\end{aligned}$$

which is reminiscent of (2.3) in terms of separating local and long-term behavior. The content of this paper is to make this argument rigorous.

The basic ingredients of the proof are asymptotic independence between large blocks of observations, local central limit theorems regarding log-likelihood ratios and large deviations arguments. Although the examples we work out entail normal observations, the arguments should hold in general. Nonetheless, a full proof of (2.2) seems to require a case-by-case treatment. In Section 3, we give a proof for an iid case, which can be taken as a blueprint for the basic ideas. In Section 4, we deal with a more complicated case, which we believe exemplifies the problems arising in the general case. We conjecture that our rule of thumb is valid in most cases which possess the aforementioned basic ingredients. Clearly, the rule won't work always: if the dependence is too strong—such as when all observations are identically the same—the result is wrong. In intermediate cases, such as interchangeable sequences, appropriate modifications to our rule should hold.

3. An iid case. In this section, we exemplify our approach by considering a power one test of a shift of a normal mean. The considerations involved are prototypical to more complicated problems.

Let $X_1, X_2, \ldots$ be a sequence of observations. Let P_0 be a measure according to which these observations are iid $N(0,1)$ and let P_1 be a measure according to which they are iid $N(\mu, 1)$. Let

$$\begin{aligned}
l_n &= \log\left[\frac{dP_1(X_1,\ldots,X_n)}{dP_0(X_1,\ldots,X_n)}\right] \\
&= \sum_{i=1}^{n} \big(\mu X_i - \mu^2/2\big) \\
&= \sum_{i=1}^{n} Z_i,
\end{aligned}$$

so that l_n is the log-likelihood ratio statistic based on n observations and Z_i is the log-likelihood ratio statistic based on the ith observation. Let $I = E_1 Z_i = \mu^2/2$.

The null hypothesis P_0 is rejected in favor of P_1 if

$$\max_{1 \le n < \infty} l_n \ge a. \tag{3.1}$$

where $a = \log A$. The significance level of this test is given by

$$P_0\Big(\max_{1 \le n < \infty} l_n \ge a\Big),$$

which we want to approximate by calculating the limit

$$\lim_{a \to \infty} e^a P_0\Big(\max_{1 \le n < \infty} l_n \ge a\Big).$$

Using the notation of Section 2, we formulated the theorem.

THEOREM 3.1.

$$\lim_{a \to \infty} e^a P_0\Big(\max_{1 \le n < \infty} l_n \ge a\Big) = \lambda = \frac{1}{I} E \frac{M}{S}. \tag{3.2}$$

PROOF. The proof will require the following lemmas.

LEMMA 3.1. *Given* $\varepsilon > 0$, *there exists a finite constant* $c > 0$ *such that*

$$\lim_{a \to \infty} e^a P_0\Big(\max_{n < 2a/\mu^2 - c\sqrt{a}} l_n \ge a\Big) \le \varepsilon/2, \tag{3.3}$$

$$\lim_{a \to \infty} e^a P_0\Big(\max_{n \le 2a/\mu^2 + c\sqrt{a}} l_n < a, \max_{1 \le n < \infty} l_n \ge a\Big) \le \varepsilon/2. \tag{3.4}$$

PROOF. Let $\lfloor x \rfloor$ denote the integer part of x. In order to show (3.3), consider the stopping time of the power one SPRT:

$$N = \min\{n \colon l_n \ge a\}.$$

Using the usual technique of turning P_0-calculations into E_1,

$$\begin{aligned} e^a P_0\Big(\max_{n < 2a/\mu^2 - c\sqrt{a}} l_n \ge a\Big) &= E_1\Big[\exp(-(l_N - a))\mathbb{1}\Big(N < \frac{2a}{\mu^2} - c\sqrt{a}\Big)\Big] \\ &\le P_1\Big(\max_{n < 2a/\mu^2 - c\sqrt{a}} l_n \ge a\Big) \\ &\le \frac{\mathrm{Var}_1\big(l_{\lfloor 2a/\mu^2 - c\sqrt{a} \rfloor}\big)}{ac^2\mu^4/4}, \end{aligned}$$

where the last inequality follows by applying Doob's inequality to the P_1-martingale $l_n - n\mu^2/2$. Finally,

$$\frac{\mathrm{Var}_1\big(l_{\lfloor 2a/\mu^2 - c\sqrt{a} \rfloor}\big)}{ac^2\mu^4/4} = \frac{\lfloor 2a - \mu^2 c\sqrt{a} \rfloor}{ac^2\mu^4/4},$$

which can be made as small as desired, uniformly in a, by choosing a large enough c.

The claim (3.4) can be shown in a similar way. Indeed,

$$\begin{aligned}
e^a \mathrm{P}_0\Big(\max_{n \le (2a/\mu^2)+c\sqrt{a}} l_n < a, \max_{1 \le n < \infty} l_n \ge a\Big) &= e^a \mathrm{P}_0\left(\frac{2a}{\mu^2} + c\sqrt{a} < N < \infty\right) \\
&\le \mathrm{P}_1\left(\frac{2a}{\mu^2} + c\sqrt{a} < N < \infty\right) \\
&\le \mathrm{P}_1\big(l_{\lfloor 2a/\mu^2 + c\sqrt{a}\rfloor} < a\big) \\
&= \Phi\left(-\frac{c\mu^2\sqrt{a}/2}{\sqrt{\lfloor 2a + c\mu^2\sqrt{a}\rfloor}}\right),
\end{aligned}$$

and the proof of the lemma follows. □

Denote by $J(k,t)$ the set of integers $\{i\colon |i-k| \le t\}$. Define the set of indices $J = J(2a/\mu^2, c\sqrt{a})$. For any $k \in J$ consider the changepoint measure $\mathrm{P}^{(k)}$ given by

$$\begin{aligned}
&X_1, X_2, \ldots, X_k \sim \text{iid}, && N(\mu, 1),\\
&X_{k+1}, X_{k+2} \ldots \sim \text{iid}, && N(0,1).
\end{aligned}$$

In the next lemma the measure P_0 is transformed to the measure $\sum_{k \in J} \mathrm{P}^{(k)}$.

LEMMA 3.2.

$$\mathrm{P}_0\Big(\max_{n \in J} l_n \ge a\Big) = \sum_{k \in J} \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}.}\right]$$

PROOF. The log-likelihood ratio of $\mathrm{P}^{(n)}$ to P_0, based on the complete sequence of observations, is l_n. Hence the likelihood ratio of P_0 to $\sum_{n \in J} \mathrm{P}^{(n)}$ is $1/\sum_{n \in J} \exp\{l_n\}$. □

Next we turn to the investigation of the term

$$\mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}}\right]$$

as a function of k and a. Given k, it will be shown that this term can be approximated by a similar term for which the set of indices J is replaced by the set $J(k,t)$, $t = ((32/\mu^2) \vee 1)\log a$.

LEMMA 3.3. *Let $\varepsilon > 0$ be given. Then for all $2a/\mu^2 - c\sqrt{a} + t \le k \le 2a/\mu^2 + c\sqrt{a} - t$ it is true that*

$$
\begin{aligned}
&e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}}\right] \\
&\qquad \le e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \ge a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] + \frac{\varepsilon}{a} \quad \text{and} \\
&e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}}\right] \\
&\qquad \ge \frac{1}{1+\varepsilon} e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \ge a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] - \frac{\varepsilon}{a},
\end{aligned}
\tag{3.5}
$$

provided that a is large enough.

PROOF. On the one hand,

$$
\begin{aligned}
e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}}\right] &\le e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}}\right] \\
&\quad + \mathrm{P}^{(k)}\Big(\max_{n \in J} l_n > \max_{n \in J(k,t)} l_n\Big) \\
&\le e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \ge a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] \\
&\quad + \mathrm{P}^{(k)}\Big(\max_{n > k+t} l_n - l_k > 0\Big) \\
&\quad + \mathrm{P}^{(k)}\Big(\max_{1 \le n < k-t} l_n - l_k > 0\Big).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}}\right] &\ge e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \ge a)}{\sum_{n \in J} \exp\{l_n\}}\right] \\
&\ge \frac{1}{1+\varepsilon} e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \ge a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] \\
&\quad - \mathrm{P}^{(k)}\left(\sum_{n=k+t+1}^{\infty} \exp\{l_n - l_k\} \ge \varepsilon/2\right) \\
&\quad - \mathrm{P}^{(k)}\left(\sum_{n=1}^{k-t-1} \exp\{l_n - l_k\} \ge \varepsilon/2\right).
\end{aligned}
$$

Notice that

$$l_n - l_k = \begin{cases} \sum_{i=k+1}^{n} Z_i, & n > k, \\ -\sum_{i=n+1}^{k} Z_i, & n < k, \end{cases}$$

and for large enough values of $n - k$,

$$\mathrm{P}^{(k)}\left(\sum_{i=k+1}^{n} Z_i > -\frac{\mu^2}{4}(n-k) \right) = 1 - \Phi\left(\mu\sqrt{n-k}/4 \right) \tag{3.6}$$
$$\leq \frac{4\exp\{-(\mu^2/32)(n-k)\}}{\mu\sqrt{2\pi(n-k)}}.$$

The sum of these probabilities over $n, n > k + t$, is $o(1/a)$. This observation leads to the conclusion that for large enough a,

$$\mathrm{P}^{(k)}\left(\sum_{n=k+t+1}^{\infty} \exp\{l_n - l_k\} \geq \frac{\varepsilon}{2} \right) < \frac{\varepsilon}{2a}.$$

Similar derivations give bounds to the other probability terms under consideration and the proof of the lemma follows. □

REMARK. The lemma, with appropriate changes in the definition of $J(k, t)$, is valid also for k such that $c\sqrt{a} - t < |k - 2a/\mu^2| \leq c\sqrt{a}$.

One can rewrite the term

$$e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \geq a)}{\sum_{n \in J(k,t)} \exp\{l_n\}} \right]$$

in the form

$$\mathrm{E}^{(k)}\Bigg[\frac{\max_{n \in J(k,t)} \exp\{l_n\}}{\sum_{n \in J(k,t)} \exp\{l_n\}} \exp\left\{ -\left(l_{k-t} + \max_{n \in J(k,t)} l_n - l_{k-t} - a \right) \right\}$$
$$\times \mathbb{1}\left(l_{k-t} + \max_{n \in J(k,t)} l_n - l_{k-t} \geq a \right) \Bigg].$$

It can be seen that the term, in this form, is an expectation of the product of two random variables. The first random variable,

$$\frac{\max_{n \in J(k,t)} \exp\{l_n - l_k\}}{\sum_{n \in J(k,t)} \exp\{l_n - l_k\}} = \frac{M_k^*}{S_k^*}$$

is positive and bounded by one. Its distribution, under $\mathrm{P}^{(k)}$, is independent of k and of l_{k-t}. The second random variable is an exponent, over a set, of a sum of two independent variables l_{k-t}, which has a normal distribution, and $\max_{n \in J(k,t)}(l_n - l_{k-t})$, which is nonnegative.

LEMMA 3.4. *Let $\varepsilon > 0$ be given. Then, for large enough a,*

$$e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \geq a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] \leq e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(a + t \geq \max_{n \in J(k,t)} l_n \geq a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] + \frac{\varepsilon}{\sqrt{a}}.$$

PROOF. From the alternative form in which the term was rewritten it can be concluded that

$$e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n > a + t)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] \leq e^{-t}. \qquad \square$$

The process variable $\{l_n - l_{k-t}\colon n \in J(k,t)\}$ has a positive drift up to time k and a negative drift thereafter. In the next lemma we show that its maximum is of a controllable order.

LEMMA 3.5. *Let $\varepsilon > 0$ be given. Then*

$$\mathrm{P}^{(k)}\left(\max_{n \in J(k,t)} l_n - l_{k-t} > \varepsilon\sqrt{a}\right) \leq \varepsilon/\sqrt{a}\,.$$

PROOF. The process $\exp\{l_n - l_k\}$ is a $\mathrm{P}^{(k)}$-martingale of mean one. Hence,

$$\mathrm{P}^{(k)}\left(\max_{n \in J(k,t)} l_n - l_k > \varepsilon\sqrt{a}\,/2\right) \leq \exp\left\{-\varepsilon\sqrt{a}\,/2\right\}.$$

The random variable $l_k - l_{k-t}$ has a normal distribution and

$$\mathrm{P}^{(k)}\left(l_k - l_{k-t} > \varepsilon\sqrt{a}\,/2\right) = 1 - \Phi\left(\frac{\varepsilon\sqrt{a}\, - t\mu^2}{2\mu\sqrt{t}}\right),$$

which converges to zero at a rate faster than $1/a$. $\square$

Lemmas 3.4 and 3.5 can be summarized by saying that the term

$$e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \geq a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right]$$

can be approximated, up to an $o(1/\sqrt{a})$ term, by

$$\begin{aligned}
&\mathrm{E}^{(k)}\left[\frac{\max_{n\in J(k,t)}\exp\{l_n\}}{\sum_{n\in J(k,t)}\exp\{l_n\}}\exp\left\{-\left(l_{k-t}+\max_{n\in J(k,t)} l_n - l_{k-t}-a\right)\right\}\right.\\
&\qquad\times\mathbb{1}\left(a+t\ge l_{k-t}+\max_{n\in J(k,t)} l_n - l_{k-t}\ge a;\right.\\
&\qquad\qquad\left.\left.\max_{n\in J(k,t)} l_n - l_{k-t}\le\varepsilon\sqrt{a}\right)\right].
\end{aligned}\tag{3.7}$$

This expectation will be approximated by conditioning on the values of $\max_{n\in J(k,t)}\exp\{l_n-l_{k-t}\}$ and $\sum_{n\in J(k,t)}\exp\{l_n-l_{k-t}\}$ and then integrating over the values of the independent random variable l_{k-t}. This random variable has a normal distribution. The approximation will result from the following lemma.

LEMMA 3.6. *Assume that*

$$X_n\sim N(n\tau,n\sigma^2),\qquad n=1,2,\dots.$$

Let $n=n(a)$ be a sequence of integers where $(n-a/\tau)/\sqrt{a}\to y$ as $a\to\infty$ and let $m=m(a)$ and $t=t(a)$ be two sequences of real numbers that are $o(\sqrt{a})$ with $t(a)\to\infty$ as $a\to\infty$. Then

$$\begin{aligned}
&\lim_{a\to\infty}\sqrt{a}\,\mathrm{E}\exp(-(X_n+m-a))\mathbb{1}(a+t\ge X_n+m\ge a)\\
&\qquad=\frac{\sqrt{\tau}}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{\tau^3}{2\sigma^2}y^2\right\}
\end{aligned}$$

uniformly for y in a compact set.

PROOF. The density of X_n at x is given by

$$\begin{aligned}
f(x)&=\frac{1}{\sqrt{2\pi}\,\sigma\sqrt{n}}\exp\left\{-\frac{(x-n\tau)^2}{2n\sigma^2}\right\}\\
&=\frac{1}{\sqrt{n}}\,\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{a\tau^2}{2n\sigma^2}\left(\frac{x-a}{\tau\sqrt{a}}-\frac{n-a/\tau}{\sqrt{a}}\right)^2\right\}.
\end{aligned}$$

However, $a/n\to\tau$, hence

$$\lim_{a\to\infty}\sqrt{a}\,f(x)=\frac{\sqrt{\tau}}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{\tau^3}{2\sigma^2}y^2\right\}$$

uniformly in x, for x for in the range $[a-m,a-m+t]$.

The integral

$$\int_{a-m}^{a-m+t} \exp(-(x+m-1))\,dx$$

converges to 1. The proof of the lemma follows from standard convergence arguments. □

From Lemma 3.6 we conclude that the $\mathrm{P}^{(k)}$-conditional expectation, given M_k^* and S_k^*, of the ingrand in (3.7) can be approximated by the term

$$\frac{M_k}{S_k}\frac{\sqrt{\tau}}{\sqrt{2\pi a}\,\sigma}\exp\left\{-\frac{\tau^3}{2\sigma^2}\frac{(k-a/\tau)^2}{a}\right\}.$$

The unconditional expectation becomes

$$\mathrm{E}^{(k)}\left[\frac{M_k}{S_k}\right]\frac{\sqrt{\tau}}{\sqrt{2\pi a}\,\sigma}\exp\left\{-\frac{\tau^3}{2\sigma^2}\frac{(k-a/\tau)^2}{a}\right\} < \frac{\mathrm{const}_1}{\sqrt{a}}\exp\{-\mathrm{const}_2\cdot c^2\},$$

where $c\sqrt{a}$ is the radius of the interval J, the interval of indices centered about a/τ. Notice, moreover that $\mathrm{E}^{(k)}[M_k/S_k]$ converges, as k increases, to a constant we denote by $\mathrm{E}[M/S]$.

Next we return to the summation, over k in the set J, of the approximated terms. Summation is approximate to integration against the counting measure. When we transform the variable of integration k to the variable $y=(k-a/\tau)/\sqrt{a}$ we get that

$$\begin{aligned}
e^a\mathrm{P}_0\Big(\max_{n\in J} l_n \ge a\Big) &= e^a\sum_{k\in J}\mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J} l_n\ge a)}{\sum_{n\in J}\exp\{l_n\}}\right]\\
&= (1+O(\varepsilon))\sum_{k\in J}\mathrm{E}\left[\frac{M}{S}\right]\frac{\sqrt{\tau}}{\sqrt{2\pi a}\,\sigma}\exp\left\{-\frac{\tau^3}{2\sigma^2}\frac{(k-a/\tau)^2}{a}\right\}\\
&= (1+O(\varepsilon))\mathrm{E}\left[\frac{M}{S}\right]\int_{-c}^{c}\frac{\sqrt{\tau}}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{\tau^3}{2\sigma^2}y^2\right\}dy\\
&= (1+O(\varepsilon))\mathrm{E}\left[\frac{M}{S}\right]\frac{1}{\tau}\left(2\Phi\left(\frac{c\tau^{3/2}}{\sigma}\right)-1\right).
\end{aligned}$$

Finally, letting $\varepsilon\to 0$, in which case $c\to\infty$, we can conclude that the limit λ of e^a times the tail probability can be represented as

$$\lambda = \mathrm{E}\left[\frac{M}{S}\right]\frac{1}{\tau} = \mathrm{E}\left[\frac{M}{S}\right]\frac{2}{\mu^2}. \qquad \square$$

4. A more complicated example: testing H_0: $\underline{X}\sim$ iid $N(0,1)$ vs. H_1: $\underline{X}\sim$ AR(1). In this section we show our rule of thumb to hold for testing H_0: $\underline{X}\sim$ iid $N(0,1)$ vs. H_1: $\underline{X}$ is an autoregressive sequence with $p=1$ and with a known autoregression parameter θ. Apart from this problem being of interest in its own right, the proof of the rule of thumb in this case is an example for the type of considerations, beyond the blueprint proof of Section 3, which are required for more complicated cases.

Let $X_0, X_1, X_2, \dots$ be a sequence of observations. Let P_0 be the measure according to which these observations are iid $N(0,1)$, and let P_1 be a measure according to which $X_0 \sim N(0,1)$ and $X_i = \theta X_{i-1} + \epsilon_i$ for $i \geq 1$, where ϵ_i are iid $N(0,1)$ and are independent of X_0, and $|\theta| < 1$. Define

$$l_n = \log \frac{d\mathrm{P}_1(X_0,\dots,X_n)}{d\mathrm{P}_0(X_0,\dots,X_n)} = \sum_{i=1}^{n} \left(\theta X_i X_{i-1} - \theta^2 X_{i-1}^2/2\right) = \sum_{i=1}^{n} Z_i,$$

$$I = \frac{1}{2}\frac{\theta^2}{1-\theta^2}.$$

Let M, S be as in Section 2.

THEOREM 4.1.

$$\lim_{a\to\infty} e^a \mathrm{P}_0\Big(\max_{1\leq n<\infty} l_n \geq a\Big) = \frac{1}{I}\mathrm{E}\frac{M}{S}.$$

The proof of Theorem 4.1 follows the lines of the proof of Theorem 3.1. The differences which have to be taken into consideration are the nonnormality of Z_i and the dependence between l_k and $l_n - l_k$. A conditioning argument will take care of the problems caused by this dependence, and the asymptotic normal limit of the density of $\sum_{i=1}^n Z_i$ (standardized) as well as large deviations arguments will be substituted for the normality of Z_i. We will sketch the proof in this section. Some of the formal details are relegated to the Appendix.

LEMMA 4.1. *Given $\varepsilon > 0$, there exists a finite constant c_ϵ such that if $c > c_\varepsilon$*

$$\lim_{a\to\infty} e^a \mathrm{P}_0\Big(\max_{n<a/I-c\sqrt{a}} l_n \geq a\Big) \leq \frac{\varepsilon}{2}, \tag{4.1}$$

$$\lim_{a\to\infty} e^a \mathrm{P}_0\Big(\max_{n<a/I+c\sqrt{a}} l_n < a,\ \max_{1\leq n<\infty} l_n \geq a\Big) \leq \frac{\varepsilon}{2}. \tag{4.2}$$

SKETCH OF PROOF. This is along the same lines of the proof of Lemma 4.1, except that the P_1-compensator of $\sum_{i=1}^n Z_i$ is $(\theta^2/2)\sum_{i=1}^n X_{i-1}^2$. The variances of $\sum_{i=1}^n (Z_i - (\theta^2/2)X_{i-1}^2)$ and $\sum_{i=1}^n X_{i-1}^2$ are $O(n)$. $\square$

In the following three lemmas the moment generating function of the log-likelihood statistic is investigated, both under P_0 and under P_1. The basic ingredients of the proof—asymptotic independence between large blocks of observations, local central limit theorems regarding log-likelihood ratios and large deviations arguments—are later shown to hold, using properties of the moment generating function. The problems of dependence of $l_k, l_n - l_k$ and the nonnormality of Z_i can thus be overcome. The proofs of these lemmas is by somewhat lengthy calculations.

LEMMA 4.2.

$$\mathrm{E}_1\left(\exp\left\{\gamma \sum_{i=1}^{n} Z_i\right\} \,\middle|\, X_0, X_{n+1}\right)$$
$$= \frac{\left(\prod_{i=1}^{n} \delta_i\right)^{\gamma/2}}{\left(\prod_{i=1}^{n}(1+\gamma(\delta_i - 1)/\delta_i)\right)^{1/2}} \exp\left\{(1/2)\gamma \underline{\mu}'\left[\sum_{j=0}^{\infty}\left(\frac{\gamma}{\gamma+1}\right)^j (\Sigma_n)^j\right]\underline{\mu}\right\}$$

where

$$\delta_j = 1 + \theta^2 - 2\theta\cos\left(\frac{\pi j}{n+1}\right), \qquad j = 1, \dots, n,$$
$$\Sigma_n = (\sigma_{ij})_{n \times n}^{-1},$$
$$\underline{\mu}' = (\mu_1, \dots, \mu_n),$$

with

$$\sigma_{ij} = \begin{cases} 1+\theta^2, & |i-j| = 0, \\ -\theta, & |i-j| = 1, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\mu_i = \frac{1-\theta^2}{1-\theta^{2n+2}}\left(\theta^i - \theta^{2n+2-i}, \theta^{n+1-i} - \theta^{n+1-i}\right)\begin{pmatrix} X_0 \\ X_{n+1}\end{pmatrix}.$$

LEMMA 4.3. *Let $\underline{\mu}, \Sigma_n$ be as in Lemma 4.2. For small enough γ there exists a constant $\omega > 0$ such that $\underline{\mu}'[\sum_{j=0}^{\infty}(\gamma/(\gamma+1))^j(\Sigma_n)^j]\underline{\mu} \le \omega(X_0^2 + X_{n+1}^2)$.*

LEMMA 4.4. *Let $X_0 \sim N(0, \tau^2)$ where $0 < \tau < 1/(1-\theta^2)$, $X_i \sim N(0,1)$ for $i = 1, 2, \dots$ all be independent. Let s be such that*

$$1 + \theta^2 s + \sqrt{(1+\theta^2 s) - 4\theta^2 s^2} > 0.$$

Then

$$\mathrm{E}\exp\left\{s\sum_{i=1}^{n}\left(\theta X_{i-1}X_i - \theta^2 X_{i-1}^2/2\right)\right\}$$
$$\le O\left(\left(\frac{2}{1+\theta^2 s + \sqrt{(1+\theta^2 s)^2 - 4\theta^2 s^2}}\right)^{n/2}\right).$$

LEMMA 4.5.

$$P_0\Big(\max_{n\in J} l_n \ge a\Big) = \sum_{k\in J} E^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J} l_n \ge a)}{\sum_{n\in J}\exp\{l_n\}}\right].$$

The proof is verbatim as with Lemma 3.2.

LEMMA 4.6. *Let $\varepsilon > 0$ be given and let c be as in Lemma* 4.1. *There exists $1 < \gamma < \infty$ such that if $t = \gamma \log a$ then for all $a/I - c\sqrt{a} + t \le k \le a/I + c\sqrt{a} - t$, it is true that*

$$E^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J} l_n \ge a)}{\sum_{n\in J}\exp\{l_n\}}\right] \le E^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J(k,t)} l_n \ge a)}{\sum_{n\in J(k,t)}\exp\{l_n\}}\right] + \frac{\varepsilon}{a}$$

and

$$E^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J} l_n \ge a)}{\sum_{n\in J}\exp\{l_n\}}\right] \ge \frac{1}{1+\varepsilon}E^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J(k,t)} l_n \ge a)}{\sum_{n\in J(k,t)}\exp\{l_n\}}\right] - \frac{\varepsilon}{a}$$

provided that a is large enough.

SKETCH OF PROOF. This is along the lines of the proof of Lemma 3.3, verbatim until (3.6) and (3.6) replaced by a large deviation argument which utilizes Lemmas 4.2 and 4.4.

REMARK. Lemma 4.6, with appropriate changes in the definition of $J(k,t)$, is also valid for k such that $c\sqrt{a} - t < |k - a/I| \le c\sqrt{a}$.

One can rewrite the term

$$e^a E^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J(k,t)} l_n \ge a)}{\sum_{n\in J(k,t)}\exp\{l_n\}}\right]$$

in the following form:

$$E^{(k)}\Bigg[\frac{\max_{n\in J(k,t)}\exp\{l_n\}}{\sum_{n\in J(k,t)}\exp\{l_n\}}\exp\Big(-\Big(l_{k-t} + \max_{n\in J(k,t)}(l_n - l_{k-t}) - a\Big)\Big) \times \mathbb{1}\Big(l_{k-t} + \max_{n\in J(k,t)}(l_n - l_{k-t}) \ge a\Big)\Bigg].$$

This is an expectation of the product of two random variables. The first is

$$\frac{\max_{n\in J(k,t)}\exp\{l_n\}}{\sum_{n\in J(k,t)}\exp\{l_n\}} = \frac{M_k^*}{S_k^*},$$

which is positive and bounded by 1. Its conditional distribution under $P^{(k)}$, given $X_0 = 0$ and X_{k-t}, is independent of l_{k-t} and does not depend on k. The second random variable is an exponent (on a set) of a sum of two variables, l_{k-t} and $\max_{n \in J(k,t)}(l_n - l_{k-t})$, which are conditionally (on X_{k-t}) independent.

In the next lemma we assume that $t = \gamma \log a$, with γ as in Lemma 4.6.

LEMMA 4.7. *Let $\epsilon > 0$ be given. Then, for large enough a,*

$$e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \geq a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] \leq e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(a + t \geq \max_{n \in J(k,t)} l_n \geq a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right] + \frac{\varepsilon}{\sqrt{a}}.$$

The proof is verbatim as with Lemma 3.4.

The process variable $\{l_n - l_k : n \in J(k,t)\}$ has a positive drift up to time k and a negative drift thereafter. In the next lemma we show that its maximum is of a controllable order.

LEMMA 4.8. *Let $\epsilon > 0$ be given. Then*

$$\mathrm{P}^{(k)}\left(\max_{n \in J(k,t)} (l_n - l_{k-t}) > \varepsilon\sqrt{a}\right) \leq \frac{\varepsilon}{\sqrt{a}}.$$

SKETCH OF PROOF. This is along the lines of the proof of Lemma 3.5, using a large deviation argument (instead of the normality of $l_k - l_{k-t}$) which utilizes Lemma 4.2 and 4.3.

Lemmas 4.7 and 4.8 can be summarized by saying that the term

$$e^a \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n \in J(k,t)} l_n \geq a)}{\sum_{n \in J(k,t)} \exp\{l_n\}}\right]$$

can be approximated, up to a $o(1/\sqrt{a})$ term, by

$$\mathrm{E}^{(k)}\left[\frac{\max_{n \in J(k,t)} \exp\{l_n\}}{\sum_{n \in J(k,t)} \exp\{l_n\}} \exp\left(-\left(l_{k-t} + \max_{n \in J(k,t)} (l_n - l_{k-t}) - a\right)\right)\right.$$
$$\left.\times \mathbb{1}\left(a + t \geq l_{k-t} + \max_{n \in J(k,t)} (l_n - l_{k-t}) \geq a;\ \max_{n \in J(k,t)} (l_n - l_{k-t}) \leq \varepsilon\sqrt{a}\right)\right].$$

This expectation will be approximated by conditioning on the value of X_{k-t}, $M_k^* = \max_{n \in J(k,t)} \exp\{l_n - l_{k-t}\}$ and $S_k^* = \sum_{n \in J(k,t)} \exp\{l_n - l_{k-t}\}$ and then integrating over the (conditionally) independent random variable l_{k-t}. We will need the following two lemmas.

LEMMA 4.9. *Suppose $\{X_n\}$ is a sequence such that the density of $(X_n - n\tau)/(\sigma\sqrt{n})$ converges to the $N(0,1)$ density uniformly on compact sets. Let*

$m = m(a)$ and $t = t(a) \to_{a\to\infty} \infty$ be two sequences of real numbers that are $o(\sqrt{a})$. Then, for $n = y\sqrt{a} + a/\tau$,

$$\lim_{a\to\infty} \sqrt{a}\,\mathrm{E}\exp(-(X_n + m - a))\mathbb{1}(a + t \geq X_n + m \geq a)$$
$$= \frac{\sqrt{\tau}}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{\tau^3}{2\sigma^2}y^2\right).$$

The proof is analogous to that of Lemma 3.6.

LEMMA 4.10. *Conditional on X_{k-t}, the $\mathrm{P}^{(k)}$-conditional density of the random variable $[\sum_{i=1}^{k-t} Z_i - \tau(k-t)]/[\sigma\sqrt{k-t}]$, with $\tau = 1/I$ and $\sigma^2 = \theta^2/(1-\theta^2)$, converges to the $N(0,1)$ density as $a \to \infty$. The convergence is uniform on compact sets.*

SKETCH OF PROOF. The integral of the absolute value of the characteristic function of $[\sum_{i=1}^{k-t} Z_i - \tau(k-t)]/[\sigma\sqrt{k-t}]$ is finite, by virtue of Lemma 4.2 and Lemma 4.3. □

From Lemmas 4.9 and 4.10 we can conclude that the $\mathrm{P}^{(k)}$-conditional expectation (conditional on M_k^* and S_k^*) of

$$\frac{M_k^*}{S_k^*}\exp(-(l_{k-t} + M_k^* - a))\mathbb{1}\left(a + t \geq l_{k-t} + M_k^* \geq a;\ M_k^* \leq \varepsilon\sqrt{a}\right)$$

can be approximated by

$$\frac{M_k^*}{S_k^*}\frac{1}{\sqrt{a}}\frac{\sqrt{\tau}}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{\tau^3}{2\sigma^2}\frac{(k-a/\tau)^2}{a}\right)$$

with τ and σ as in Lemma 4.10. Notice that $\mathrm{E}^{(k)}(M_k^*/S_k^*)$ converges, as k increases, to a constant we denote by $\mathrm{E}(M/S)$.

Next we turn to the summation, over k in the set J, of the approximated terms. Summation is approximated by integration with respect to counting measure. When we transform the variable of integration to the variable $y = (k - a/\tau)/\sqrt{a}$ we get that

$$\begin{aligned}
e^a\mathrm{P}_0\left(\max_{n\in J} l_n \geq a\right) &= e^a \sum_{k\in J}\mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J} l_n \geq a)}{\sum_{n\in J}\exp\{l_n\}}\right] \\
&\approx \mathrm{E}\frac{M}{S}\sum_{k=a/\tau-c\sqrt{a}}^{a/\tau+c\sqrt{a}}\frac{1}{\sqrt{a}}\frac{\sqrt{\tau}}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{\tau^3}{2\sigma^2}\frac{(k-a/\tau)^2}{a}\right) \\
&\approx \mathrm{E}\frac{M}{S}\int_{-c}^{c}\frac{\sqrt{\tau}}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{1}{2}\frac{y^2}{\sigma^2/\tau^3}\right)dy \\
&\to \mathrm{E}\frac{M}{S}\frac{1}{\tau} \qquad \text{as } c \to \infty
\end{aligned}$$

and so

$$\lim_{a\to\infty} e^a \mathrm{P}_0\Big(\max_{1\le n<\infty} l_n \ge a\Big) = \mathrm{E}\frac{M}{S}\frac{1}{I}. \qquad \square$$

5. Monte Carlo. Using the notation of Section 4, we know from Theorem 4.1 that

$$\lim_{a\to\infty} e^a \mathrm{P}_0\Big(\max_{1\le n<\infty} l_n \ge a\Big) = \frac{1}{I}\mathrm{E}\frac{M}{S} \overset{\mathrm{def}}{=} \lambda.$$

Table 1 gives an idea of the rate of convergence, where the Monte Carlo was done by the importance sampling formula

$$e^a \mathrm{P}_0\Big(\max_{1\le n<\infty} l_n \ge a\Big) = \mathrm{E}_1 \exp\{a - l_N\},$$

where $N = \min\{n: l_n \ge a\}$ and E_1 denotes expectation with respect to P_1. Hence $\exp\{a - l_N\}$ was simulated under P_1, and Table 1 represents the means and standard errors for $a = j\log 10$, $j = 1,\dots,6$, based on 100,000 simulations, when $\theta = 0.7$. (The value $\theta = 0.7$ corresponds to $I = (1/2)\theta^2/(1-\theta^2) = 0.48 \approx 0.5$, which is the information in the iid hypothesis testing problem of H_0: $X_i \sim N(0,1)$ vs. H_1: $X_i \sim N(1,1)$ where, by the way, $\lambda = 0.5604$).

Clearly, although the convergence is not slow—by $a = 2\log 10$ it seems that it is within 10% of the limit—evaluating the limit by this method is problematic, as it is not clear how large a value of a one should choose. Evaluating λ directly (by simulating M/S) produces a much more reliable estimate: 0.4261 as the mean and 0.0010 as the standard error (see Table 2). Clearly, precision to two significant digits is not achieved if $e^a < 10^5$.

Using the methods of Lemma 4.4 to get bounds on the moment generating function of the log-likelihood statistic, one can achieve exponential bounds on likelihood ratios and thus bound the expression

$$\left|\mathrm{E}\frac{M}{S} - \mathrm{E}\frac{M^*}{S^*}\right|,$$

TABLE 1

Means and standard errors of $\exp\{a - l_N\}$ *based on* 100,000 *simulations made under* P_1 *for* $\theta = 0.7$ *and for various values of a*

e^a	10	10^2	10^3	10^4	10^5	10^6
Mean	0.4921	0.4605	0.4474	0.4384	0.4356	0.4327
StdErr	0.0009	0.0009	0.0009	0.0009	0.0009	0.0010

TABLE 2
Means and standard errors of $(1/I)(M^*/S^*)$ *and numbers of observations* $(2t+1)$ *on which* M^*, S^* *are based, for various values of* θ

$\boldsymbol{\theta}$	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Mean	0.9174	0.8342	0.7527	0.6711	0.5921	0.5120	0.4261	0.3296	0.2057
StdErr	0.0012	0.0012	0.0012	0.0013	0.0013	0.0012	0.0010	0.0008	0.0004
$2t+1$	40001	10001	4001	2001	2001	1001	1001	1001	1001

where M^* and S^* are computed from a process of $2t+1$ observations (t to each "side" and one in the middle). We chose t to bound this expression by 10^{-7}, and ran 100,000 replications of $(1/I)(M^*/S^*)$ for each of $\theta = 0.1, 0.2, \dots, 0.9$. The results are entered in Table 2.

It should be noted that λ is symmetric in θ. (This can be seen by showing that the joint moment generating function of the $2t+1$ summands making up M^* and S^* is a function of θ^2.)

APPENDIX

PROOF OF LEMMA 4.1, EQUATION (4.1). Let $N = \min\{n\colon l_n \ge a\}$:

$$
\begin{aligned}
e^a \mathrm{P}_0\Big(\max_{n<a/I-c\sqrt{a}} l_n \ge a\Big)
&= \mathrm{E}_1 \exp(-(l_N - a))\mathbb{1}\big(N < a/I - c\sqrt{a}\,\big)\\
&\le \mathrm{P}_1\Big(\max_{n<a/I-c\sqrt{a}} l_n \ge a\Big)\\
&= \mathrm{P}_1\left(\exists\, n \ge a/I - c\sqrt{a}\,,\ni \sum_{i=1}^{n}\left(Z_i - \theta^2\frac{X_{i-1}^2}{2}\right) \ge a - \theta^2\sum_{i=1}^{n}\frac{X_{i-1}^2}{2}\right)\\
&\le \mathrm{P}_1\left(\exists\, n \le a/I - c\sqrt{a}\,,\ni \sum_{i=1}^{n}\left(Z_i - \theta^2\frac{X_{i-1}^2}{2}\right) \ge \frac{Ic\sqrt{a}}{2}\right)\\
&\quad + \mathrm{P}_1\left(a - \theta^2\sum_{i=1}^{n}\frac{X_{i-1}^2}{2} < \frac{Ic\sqrt{a}}{2}\right)\\
&\le \frac{\mathrm{Var}_1\Big(l_{a/I-c\sqrt{a}} - \theta^2\sum_{i=1}^{a/I-c\sqrt{a}} X_{i-1}^2/2\Big)}{I^2ac^2/4}\\
&\quad + \mathrm{P}_1\left(a - \theta^2\sum_{i=1}^{a/I-c\sqrt{a}}\frac{X_{i-1}^2}{2} < \frac{Ic\sqrt{a}}{2}\right)
\end{aligned}
$$

where the last inequality follows from applying Doob's inequality to the P_1-martingale $l_n - \theta^2 \Sigma_{i=1}^n X_{i-1}^2/2$.

$$\mathrm{Var}_1\left(\sum_{i=1}^n \left(Z_i - \theta^2 \frac{X_{i-1}^2}{2}\right)\right) = \mathrm{E}_1\left(\sum_{i=1}^n \theta X_{i-1}(X_i - \theta X_{i-1})\right)^2$$
$$= \theta^2 \mathrm{E}_1\left(\sum_{i=1}^n X_{i-1}\epsilon_i\right)^2 = \theta^2 \mathrm{E}_1 \sum_{i=1}^n X_{i-1}^2 \le \frac{\theta^2}{1-\theta^2} n$$

and

$$\mathrm{P}_1\left(\frac{\theta^2}{2} \sum_{i=1}^{a/I - c\sqrt{a}} X_{i-1}^2 > a - \frac{Ic\sqrt{a}}{2}\right)$$
$$\le \mathrm{P}_1\left(\frac{\theta^2}{2} \sum_{i=1}^{a/I - c\sqrt{a}} \left(X_{i-1}^2 - \frac{1}{1-\theta^2}\right) > \frac{Ic\sqrt{a}}{2}\right).$$

Note that

$$\mathrm{E}_1 \sum_{i=1}^n \left(X_i^2 - \frac{1}{1-\theta^2}\right) = \sum_{i=1}^n \left(\sum_{j=0}^{i-1} \theta^{2j} - \frac{1}{1-\theta^2}\right) = -\frac{\theta^2 - \theta^{2n+2}}{(1-\theta^2)^2},$$

which is bounded in $1 \le n < \infty$, and that $\mathrm{Var}_1(\Sigma_{i=1}^n X_{i-1}^2) = O(n)$. Therefore, by Chebyshev's inequality,

$$\mathrm{P}_1\left(\frac{\theta^2}{2} \sum_{i=1}^{a/I - c\sqrt{a}} X_{i-1}^2 > a - \frac{Ic\sqrt{a}}{2}\right) \le \frac{O(a)}{I^2 a c^2/4}.$$

Hence, choosing c to be large enough yields (4.1). □

PROOF OF LEMMA 4.1, EQUATION (4.2).

$$\begin{aligned}
&e^a \mathrm{P}_0\left(\max_{n < a/I + c\sqrt{a}} l_n < a,\ \max_{1 \le n < \infty} l_n \ge a\right) \\
&\quad = e^a \mathrm{P}_0\left(a/I + c\sqrt{a} < N < \infty\right) \\
&\quad \le \mathrm{P}_1\left(l_{a/I + c\sqrt{a}} < a\right) \\
&\quad \le \frac{\mathrm{Var}_1\left(\Sigma_{i+1}^{a/I + c\sqrt{a}} \left(Z_i - (\theta^2/2) X_{i-1}^2\right)\right)}{I^2 c^2 a/4} + \frac{\mathrm{Var}_1\left((\theta^2/2)\Sigma_{i=1}^{a/I + c\sqrt{a}} X_{i-1}^2\right)}{I^2 c^2 a/4} \\
&\quad = \frac{O(a)}{I^2 c^2 a/4}
\end{aligned}$$

where the last equality follows from the same considerations as given above. Choosing c to be large enough yields (4.2). □

PROOF OF LEMMA 4.2. Without loss of generality, let $X_0 \sim N(0, 1/(1-\theta^2))$, $X_i = \theta X_{i-1} + \epsilon_i$ for $i = 1, \dots, n+1$, $\epsilon_i \sim N(0,1)$ iid. Denote $\underline{X} = (X_0, \dots, X_n, X_{n+1})'$. $\mathrm{E}_1 \underline{X} = \underline{0}$. Direct calculations show that

$$\mathrm{cov}_1(\underline{X}) = \frac{1}{1-\theta^2}\begin{pmatrix} 1 & \theta & \cdots & \theta^{n+1} \\ \theta & 1 & \cdots & \theta^n \\ \vdots & \vdots & \ddots & \vdots \\ \theta^{n+1} & \theta^n & \cdots & 1 \end{pmatrix} = \Sigma_{n+2}$$

Define

$$\Sigma_{11}^{-1} = \left\{ \left(\Sigma_{n+2}^{-1} \right)_{ij} : 1 \le i, j \le n \right\}.$$

Generally, if

$$A = B^{-1},\ A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

then $[A_{11} - A_{12} A_{22}^{-1} A_{22}]^{-1} = B_{11}$. Therefore

$$\mathrm{cov}_1(X_1, \dots, X_n | X_0, X_{n+1})^{-1} = \Sigma_{11}^{-1} = \Sigma_n^{-1}.$$

The eigenvalues of this matrix are obtained from Anderson (1971) 6.5.4, and are δ_i, $i = 1, \dots, n$.

Now

$$(X_0, X_i, X_{n+1})' \sim N\left(\underline{0}, \frac{1}{1-\theta^2}\begin{pmatrix} 1 & \theta^i & \theta^{n+1} \\ \theta^i & 1 & \theta^{n+1-i} \\ \theta^{n+1} & \theta^{n+1-i} & 1 \end{pmatrix}\right).$$

Since, if

$$(Y_1, Y_2, Y_3)' \sim N\left(\underline{\alpha}, \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}\right),$$

then

$$\mathrm{E}(Y_1 | Y_2, Y_3) = \alpha_1 + \Gamma_{12}\Gamma_{22}^{-1}\begin{pmatrix} Y_2 - \alpha_2 \\ Y_3 - \alpha_3 \end{pmatrix},$$

it follows that

$$\begin{aligned} \mathrm{E}_1(X_i | X_0, X_{n+1}) &= (1-\theta^2)(\theta^i, \theta^{n+1-i})\begin{pmatrix} 1 & \theta^{n+1} \\ \theta^{n+1} & 1 \end{pmatrix}^{-1}\begin{pmatrix} X_0 \\ X_{n+1} \end{pmatrix} \\ &= \mu_i. \end{aligned}$$

Now,

$$
\begin{aligned}
&\mathrm{E}_1\left(\exp\left(\gamma\sum_{i=1}^{n} Z_i\right)\,\middle|\, X_0, X_{n+1}\right)\\
&\quad= \int |\Sigma_n|^{-\gamma/2}\exp\left(-\frac{\gamma}{2}(\underline{x}-\underline{\mu})'\Sigma_n^{-1}(\underline{x}-\underline{\mu}) + (\gamma/2)\underline{x}'\underline{x}\right)(2\pi)^{-n/2}|\Sigma_n|^{-1/2}\\
&\quad\quad\times\exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})'\Sigma_n^{-1}(\underline{x}-\underline{\mu})\right)d\underline{x}\\
&\quad= \frac{|\Sigma_n^{-1}|^{(\gamma+1)/2}}{\left|(\gamma+1)\Sigma_n^{-1}-\gamma I_n\right|^{1/2}}\\
&\quad\quad\times\exp\left(-\frac{\gamma+1}{2}\underline{\mu}'\Sigma_n^{-1}\underline{\mu} + \frac{(\gamma+1)^2}{2}\underline{\mu}'\Sigma_n^{-1}\left((\gamma+1)\Sigma_n^{-1}-\gamma I_n\right)^{-1}\Sigma_n^{-1}\underline{\mu}\right)\\
&\quad= \frac{\left(\prod_{i=1}^{n}\delta_i\right)^{\gamma/2}}{\left(\prod_{i=1}^{n}\left(1+\gamma\frac{\delta_i-1}{\delta_i}\right)\right)^{1/2}}\exp\left(\frac{\gamma}{2}\underline{\mu}'\left[\sum_{j=0}^{\infty}\left(\frac{\gamma}{\gamma+1}\right)^{j}\Sigma_n^{j}\right]\underline{\mu}\right). \qquad\square
\end{aligned}
$$

PROOF OF LEMMA 4.3. Denote by A the 2×2 matrix

$$
\sum_{i=1}^{n}\begin{pmatrix} (\theta^{i}-\theta^{2n+2-i})^2 & (\theta^{i}-\theta^{2n+2-i})(\theta^{n+1-i}-\theta^{n+1+i})\\ (\theta^{i}-\theta^{2n+2-i})(\theta^{n+1-i}-\theta^{n+1+i}) & (\theta^{i}-\theta^{2n+2-i})^2\end{pmatrix}.
$$

Hence,

$$
\underline{\mu}'\underline{\mu} = \left(\frac{1-\theta^2}{1-\theta^{2n+2}}\right)^2 (X_0, X_{n+1})A\begin{pmatrix}X_0\\ X_{n+1}\end{pmatrix}.
$$

Now, $\max_{\|\underline{x}\|=1}\underline{x}'A\underline{x}$ equals the largest eigenvalue of A, which is in our case $a_{11}+a_{12}$. Therefore

$$
\begin{aligned}
\underline{\mu}'\underline{\mu} &\le \left(X_0^2+X_{n+1}^2\right)\left(\frac{1-\theta^2}{(1-\theta^{2n+2})^2}\right)^2\\
&\quad\times\sum_{i=1}^{n}\left[(\theta^{i}-\theta^{2n+2-i})^2 + (\theta^{i}-\theta^{2n+2-i})(\theta^{n+1-i}-\theta^{n+1+i})\right].
\end{aligned}
$$

Now

$$\sum_{i=1}^{n}\left(\theta^{i}-\theta^{2n+2-i}\right)^{2}=\frac{\theta^{2}}{1-\theta^{2}}-\frac{\theta^{2n+2}-\theta^{2n+4}+\theta^{4n+4}}{1-\theta^{2}}-2n\theta^{2n+2}$$

and

$$\sum_{i=1}^{n}\left(\theta^{i}-\theta^{2n+2-i}\right)\left(\theta^{n+1-i}-\theta^{n+1+i}\right)=n\theta^{n+1}\left(1+\theta^{2n+2}\right).$$

Therefore there exists a constant $\xi>0$ such that $\underline{\mu}'\underline{\mu}\leq\xi(X_0^2+X_{n+1}^2)$. The maximal eigenvalue of Σ_n^j is $[1-2\theta\cos(\pi n/(n+1))/(1+\theta^2)]^j$. Hence

$$\underline{\mu}'\Sigma_n^j\underline{\mu}\leq\xi\left(X_0^2+X_{n+1}^2\right)\left[1-\frac{2\theta}{1+\theta^2}\cos(\pi n/(n+1))\right]^j$$

and so for small enough γ there exists $\omega>0$ such that

$$\begin{aligned}\underline{\mu}'\left[\sum_{j=0}^{\infty}\left(\frac{\gamma}{\gamma+1}\right)^{j}\Sigma_n^j\right]\underline{\mu}&\leq\xi\left(X_0^2+X_{n+1}^2\right)\sum_{j=0}^{\infty}\left(\frac{\gamma}{\gamma+1}\right)^{j}\left[1-\frac{2\theta}{1+\theta^2}\right]^j\\&\leq\omega\left(X_0^2+X_{n+1}^2\right).\end{aligned}$$

□

PROOF OF LEMMA 4.4. Denote by $\mathscr{F}_{n-1}$ the sigma-algebra generated by the first $n-1$ observations and let $b<1/2$. Then

$$\begin{aligned}&\mathrm{E}\left(\exp\left(bX_n^2\right)\exp\left(s\sum_{i=1}^{n}\left(\theta X_{i-1}X_i-\theta^2X_{i-1}^2/2\right)\right)\middle|\mathscr{F}_{n-1}\right)\\&\quad=\exp\left(s\sum_{i=1}^{n-1}\left(\theta X_{i-1}X_i-\theta^2\frac{X_{i-1}^2}{2}\right)\right)(1-2b)^{-1/2}\\&\qquad\times\exp\left(\frac{1}{2}\theta^2\left(\frac{s^2}{1-2b}-s\right)X_{n-1}^2\right).\end{aligned}$$

Denote

$$f(b)=\frac{\theta^2}{2}\left(\frac{s^2}{1-2b}-s\right).$$

The solutions to the equation $f(b)=b$ are

$$b_{1,2}=\frac{1-\theta^2s\pm\sqrt{\left(1+\theta^2s\right)^2-4\theta^2s^2}}{4}.$$

Clearly,

$$f^{(n)}(b)\to b_1=\frac{1-\theta^2s-\sqrt{\left(1+\theta^2s\right)^2-4\theta^2s^2}}{4}\qquad\text{as } n\to\infty.$$

□

PROOF OF LEMMA 4.6. On the one hand

$$\mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J} l_n \geq a)}{\sum_{n\in J}\exp\{l_n\}}\right] - \mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J(k,t)} l_n \geq a)}{\sum_{n\in J}\exp\{l_n\}}\right]$$
$$\leq \mathrm{P}^{(k)}\Big(\max_{n>k+t} l_n - l_k > 0\Big) + \mathrm{P}^{(k)}\Big(\max_{1\leq n<k-t} l_n - l_k > 0\Big).$$

On the other hand,

$$\mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J} l_n \geq a)}{\sum_{n\in J}\exp\{l_n\}}\right] - \frac{1}{1+\varepsilon}\mathrm{E}^{(k)}\left[\frac{\mathbb{1}(\max_{n\in J(k,t)} l_n \geq a)}{\sum_{n\in J(k,t)}\exp\{l_n\}}\right]$$
$$\geq -\mathrm{P}^{(k)}\left(\sum_{n=k+t+1}^{\infty}\exp\{l_n - l_k\} \geq \frac{\varepsilon}{2}\right) - \mathrm{P}^{(k)}\left(\sum_{n=1}^{k-t-1}\exp\{l_n - l_k\} \geq \frac{\varepsilon}{2}\right).$$

Notice that

$$l_n - l_k = \begin{cases} \sum_{i=k+1}^{n} Z_i, & n > k, \\ -\sum_{i=n+1}^{k} Z_i, & n < k, \end{cases}$$

that for fixed $\eta > 0$ and large enough a,

$$\mathrm{P}^{(k)}\Big(\max_{n>k+t} l_n - l_k > 0\Big) \leq \mathrm{P}^{(k)}\left(\sum_{n=k+t+1}^{\infty}\exp\{l_n - l_k\} \geq \frac{\varepsilon}{2}\right)$$
$$\leq \sum_{n=k+t+1}^{\infty}\mathrm{P}^{(k)}\left(\sum_{i=k+1}^{n} Z_i > -\frac{\theta^2}{4}(n-k)\right)$$

and

$$\mathrm{P}^{(k)}\Big(\max_{1\leq n<k-t} l_n - l_k > 0\Big) \leq \mathrm{P}^{(k)}\left(\sum_{n=1}^{k-t-1}\exp\{l_n - l_k\} \geq \frac{\varepsilon}{2}\right)$$
$$\leq \sum_{n=1}^{k-t-1}\mathrm{P}^{(k)}\left(\sum_{i=n+1}^{k} Z_i < \eta(k-n)\right).$$

By virtue of Lemma 4.4, for $0 < s < 1/[|\theta|(2-|\theta|)]$, there exists $C(s)$ such that

$$\mathrm{P}^{(k)}\left(\sum_{i=k+1}^{n} Z_i > -\frac{\theta^2}{4}(n-k)\right)$$
$$\leq C(s)\left(\frac{\exp\{-\theta^2 s/2\}}{\left(1+\theta^2 s + \sqrt{(1+\theta^2 s)^2 - 4\theta^2 s^2}\right)/2}\right)^{(n-k)/2}$$
$$< C(s_\xi)\,\xi^{n-k}$$

for some $1 > \xi > 0$ and $s_\xi > 0$. Therefore

$$\sum_{n=k+t+1}^{\infty} \mathrm{P}^{(k)}\left(\sum_{i=k+1}^{n} Z_i > -\frac{\theta^2}{4}(n-k) \right) < C(s_\xi)\frac{\xi^{t+1}}{1-\xi}.$$

Note that $\log(1+y) - 1 + 1/(1+y) \geq 0$ for all $y > -1$, with equality iff $y = 0$. It follows that

$$\frac{d}{d\gamma}\left[\log\left(\frac{(\Pi_{i=1}^{n}\delta_i)^{\gamma/2}}{(\Pi_{i=1}^{n}(1+\gamma(\delta_i - 1)/\delta_i))^{1/2}} \right) \right]\Bigg|_{\gamma=0}$$
$$= \frac{1}{2}\sum_{i=1}^{n}\left[\log(1+(\delta_i - 1)) - 1 + \frac{1}{1+(\delta_i - 1)} \right] > 0.$$

Furthermore,

$$\log\left(\frac{(\Pi_{i=1}^{n}\delta_i)^{\gamma/2}}{(\Pi_{i=1}^{n}(1+\gamma(\delta_i - 1)/\delta_i))^{1/2}} \right)$$
$$\approx \frac{n}{2}\int_0^1 \Bigg[\gamma \log(1+\theta^2 - 2\theta\cos(\pi x)) - \log\left(1 + \frac{\gamma[\theta^2 - 2\theta\cos(\pi x)]}{1+\theta^2 - 2\theta\cos(\pi x)}\right) \Bigg]\, dx.$$

Hence, there exist $\rho > 0$, $s > 0$ such that, using the notation of Lemma 4.2,

$$\mathrm{E}_1 \exp\left(-s\sum_{i=1}^{n} Z_i \right) \leq \exp(-\rho n).$$

It follows from Lemma 4.2 that

$$\mathrm{P}^{(k)}\left(\sum_{i=n+1}^{k} Z_i < \eta(k-n) \right) \leq \exp(-(\rho - \eta s)(k-n)).$$

Choosing η small enough completes the proof of Lemma 4.6. □

PROOF OF LEMMA 4.7. The process $\exp\{l_n - l_k\}$ is a $\mathrm{P}^{(k)}$-martingale of mean one. Hence,

$$\mathrm{P}^{(k)}\left(\max_{n \in J(k,t)} (l_n - l_k) > \frac{\varepsilon\sqrt{a}}{2} \right) \leq \exp(-\varepsilon\sqrt{a}/2).$$

By virtue of Lemmas 4.2 and 4.3, in a manner similar to the proof of Lemma 4.6, it follows that for $s > 0$,

$$\mathrm{P}^{(k)}\big((l_k - l_{k-t}) > \varepsilon\sqrt{a}/2\big) \leq \exp\big(s\rho t - s\varepsilon\sqrt{a}/2\big). \qquad \square$$

PROOF OF LEMMA 4.10. The unconditional asymptotic distribution of $(\sum_{i=1}^{k-t} Z_i - (k-t)\tau)/(\sigma\sqrt{k-t})$ is $N(0,1)$ by virtue of Anderson (1971).

By virtue of Lemmas 4.2 and 4.3, the conditional asymptotic distribution is the same (apart from the set $\{X_{k-t} > o(n^{1/4})\}$, which becomes negligible).

To conclude the proof, one needs to show that the absolute value of the characteristic function $\chi(\lambda)$ is integrable. By Lemma 4.2, for small enough $\gamma > 0$,

$$
\begin{aligned}
|\chi(\lambda)| &\le \left| \prod_{j=1}^{n} \left(1 + i\lambda \frac{\delta_j - 1}{\delta_j} \right) \right|^{-1/2} \\
&= \prod_{j=1}^{n} \left(1 + \lambda^2 \frac{(\delta_j - 1)^2}{(\delta_j)^2} \right)^{-1/2} \\
&\le \prod_{\{j: |\theta - 2\cos(\pi j/(n+1))| > \gamma\}} \left(1 + \lambda^2 \left(\frac{\theta^2 - 2\theta\cos(\pi j/(n+1))}{1 + \theta^2 - 2\theta\cos(\pi j/(n+1))} \right)^2 \right)^{-1/2} \\
&\le \exp(-\text{const } \lambda^2)
\end{aligned}
$$

for an appropriate positive constant. □

Acknowledgment. The authors thank Professor Paul Shaman for some helpful comments.

REFERENCES

ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.

HOGAN, M. and SIEGMUND, D. (1986). Large derivations for the maxima of some random fields. *Adv. in Appl. Math.* **7** 2–22.

KESTEN, H. (1974). Renewal theory for functionals of a Markov Chain with general state space. *Ann. Probab.* **2** 355–386.

LALLEY, S. P. (1986). Renewal theorem for a class of stationary sequences. *Probab. Theory Related Fields* **72** 195–213.

SIEGMUND, D. (1976). Importance sampling in the Monte Carlo study of sequential tests. *Ann. Statist.* **4** 673–684.

SIEGMUND, F. (1985). *Sequential Analysis: Tests and Confidence Intervals*. Springer, New York.

VILLE, J. (1939). *Étude Critique de la Notion de Collectif*. Gauthier-Villars, Paris.

YAKIR, B. (1995). A note on the run length to false alarm of a change-point detection policy. *Ann. Statist.* **23** 272–281.

DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM
91905 ISRAEL
E-MAIL: msby@mscc.huji.ac.il
msmp@mscc.huji.ac.il