A CONTINUOUS TIME APPROXIMATION TO THE UNSTABLE FIRST-ORDER AUTOREGRESSIVE PROCESS:
THE CASE WITHOUT AN INTERCEPT

Pierre Perron
Princeton University and C.R.D.E.

Econometric Research Program
Research Memorandum No. 335

August 1988

*Assistant Professor, Department of Economics, Princeton University and Research Associate, Centre de Recherche et Développement en Économique (Université de Montréal). I have benefitted from comments by Peter Phillips. Research support is acknowledged from the Social Sciences and Humanities Research Council (Canada), the Fonds Concertés pour l'Aide à la Recherche (Québec) and the Natural Sciences and Engineering Research Council of Canada. This work was completed while the author was Assistant Professor at the Department of Economics, Université de Montréal.

Econometric Research Program
Princeton University
207 Dickinson Hall
Princeton, NJ 08544, USA
ABSTRACT

Consider the first-order autoregressive process \( y_t = \alpha y_{t-1} + e_t, \) \( y_0 \) a fixed constant, \( e_t \sim \text{i.i.d.} (0, \sigma^2) \) and let \( \hat{\alpha} \) be the least-squares estimator of \( \alpha \) based on a sample of size \( (T + 1) \) sampled at frequency \( h \). Consider also the continuous time Ornstein–Uhlenbeck process \( dy_t = \theta y_t dt + \sigma dw_t \) where \( w_t \) is a Weiner process and let \( \hat{\theta} \) be the continuous time maximum likelihood (conditional upon \( y_0 \)) estimator of \( \hat{\theta} \) based upon a single path of data of length \( N \). We first show that the exact distribution of \( N(\hat{\theta} - \theta) \) is the same as the asymptotic distribution of \( T(\hat{\alpha} - \alpha) \) as the sampling interval converges to zero. This asymptotic distribution permits explicit consideration of the effect of the initial condition \( y_0 \) upon the distribution of \( \hat{\alpha} \). We use this fact to provide an approximation to the finite sample distribution of \( \hat{\alpha} \) for arbitrary fixed \( y_0 \).

The moment-generating function of \( N(\hat{\theta} - \theta) \) is derived and used to tabulate the distribution and probability density function. We also derive the mean and mean-square error of \( \hat{\theta} \) as well as the power function. In each case, the adequacy of the approximation to the finite sample distribution of \( \hat{\alpha} \) is assessed for values of \( \alpha \) in the vicinity of one. The approximations are, in general, found to be excellent.

Key Words: Ornstein–Uhlenbeck process, moment-generating function, near-integrated processes, unit root, distribution theory.
1. INTRODUCTION

Consider the following first-order stochastic difference equation:

\[(1.1) \quad y_t = \alpha y_{t-1} + e_t\]

where \(y_0 = b\) a fixed constant and the \(e_t\) are independently and identically distributed \(N(0, \sigma^2)\) variates. The unrestricted maximum likelihood estimate of \(\alpha\) (conditional upon the initial observation \(y_0\)) based upon a sequence of observations of size \(T + 1\), \(\{y_t\}_{0}^{T}\), is the least squares estimator

\[\hat{\alpha} = \sum_{t=1}^{T} y_t y_{t-1} (\sum_{t=1}^{T} y_{t-1}^2)^{-1}.\]

The distribution of \(\hat{\alpha}\) has been extensively studied. One topic of concern has been the adequacy of various asymptotic approximations to the finite sample distribution.

Mann and Wald (1943) and Rubin (1950) showed that \(T^{1/2}(\hat{\alpha} - \alpha)(1 - \alpha^2)^{-1}\) has a limiting \(N(0,1)\) distribution when \(|\alpha| < 1\). White (1958) showed that when \(|\alpha| > 1\), the limiting distribution of \(\alpha T^{1/2}(\alpha^2 - 1)^{-1} (\hat{\alpha} - \alpha)\) is Cauchy provided that \(y_0 = 0\). White also considered the case: \(|\alpha| = 1\) and showed that the limiting distribution of \(T(\hat{\alpha} - \alpha)\) can be expressed in terms of the ratio of two functionals of a Weiner process (see also Phillips (1987)). In the latter case Rao (1976) obtained a computable expression.

The case of the unit root \(\alpha = 1\) has attracted a great deal of attention. The critical values of the asymptotic distribution of \(T(\hat{\alpha} - 1)\) have been derived by Dickey (1976) by simulation methods. Evans and Savin (1981a) obtained tabulated values of both the cumulative distribution function and the probability density function of the limiting distribution of \((T/\sqrt{2})(\hat{\alpha} - 1)\) by numerically integrating the limiting moment generating function derived by White (1958). A variety of procedures have also been suggested which permit a considerable weakening of the conditions imposed on the innovation sequence \(\{e_t\}\) while at the same time using the same critical values derived
in the earlier studies; see, for example, Dickey and Fuller (1979, 1981), Said and Dickey (1984) and Phillips (1987a).

The major point which emerges from these studies is that the asymptotic distributions have a discontinuity at one. Because the exact distribution of \( \hat{\alpha} \) is continuous for all values of \( \alpha \), this suggests that the limiting distribution inadequately approximates the finite sample distribution for \( \alpha \) near the discontinuity point \( \alpha = 1 \). This inadequacy has been well–documented by Evans and Savin (1981b, 1984) and by Phillips (1977) who also showed that Edgeworth expansions also perform poorly for \( \alpha \) less than but close to one.

Recently, a new class of models which specifically deal with the presence of a root close to, but not necessarily equal to, one, have been studied. Phillips (1988) introduced the concept of a near–integrated random process where the autoregressive parameter is defined by:

\[
(1.2) \quad \alpha = \exp(c/T).
\]

Here, the real–valued constant \( c \) is a measure of the deviation from the unit root case. The models (1.1) and (1.2) may also be described as having a root local to unity (see Cavanagh (1986)): as the sample size increases, the autoregressive parameter converges to unity. When \( c \) is negative, the process \( \{y_t\} \) is said to be (locally) stationary and when \( c \) is positive, it is said to be (locally) explosive. An expression for the limiting distribution of \( T(\hat{\alpha} - \alpha) \) under (1.2) has been derived by Phillips (1988), Cavanagh (1986) and Chan and Wei (1987), and the analysis has been extended to a multivariate framework in Jeganathan (1987). This class of models has been quite useful in studying various problems such as the power of tests of a unit root under local alternatives (Phillips (1987b), Phillips and Perron (1988) and Perron (1986)), the derivation of confidence intervals when \( \alpha \) is near unity (Cavanagh (1986)) and the calculation of the power of tests of a unit root with a continuum of observations (Perron (1987)).

Tabulations of the limiting distribution of \( T(\hat{\alpha} - \alpha) \) under (1.2) have been obtained by Nabeya and Tanaka (1987), Cavanagh (1986) and Perron (1988) using different procedures. These studies also provide some measures of the adequacy of this
limiting distribution to the finite sample distribution of $\hat{\alpha}$ when $\alpha$ is in the vicinity of 1. They show the approximation to be quite good in the case where $y_0 = 0$.

A feature of substantial interest is that the limiting distribution of $T(\hat{\alpha} - \alpha)$ under the near-integrated process (1.2) is invariant to the value of the initial observation $y_0$. As thoroughly documented by Evans and Savin (1981b), the finite sample distribution of $\hat{\alpha}$ is very sensitive to the value of $y_0$ when $\alpha$ is near 1 even for quite large sample sizes. Hence, the approximation provided by the 'near-integrated' limiting distribution is adequate only in the special case where $y_0 = 0$. When $y_0 \neq 0$, the approximation fails to capture the substantial effect a non-zero initial condition has on the finite sample distribution.

The purpose of this paper is to present a rather different approach in deriving an asymptotic approximation which permits the explicit consideration of the effects of different initial conditions on the distribution of $T(\hat{\alpha} - \alpha)$. Instead of the usual framework where the sample size increases to infinity keeping a fixed sampling interval, we use the so-called continuous records asymptotic analysis where $T$ is increased to infinity keeping the span of the data fixed, i.e. by letting the sampling interval converge to zero at the same rate as $T$ increases to infinity. This method is closely related to the exact distribution of estimators in continuous time.

Consider the following continuous time Ornstein-Uhlenbeck diffusion process

$$
(1.3) \quad dy_t = \theta y_t dt + \sigma_0 dw_t; \quad y_0 = b, \; t > 0.
$$

$w_t$ is the standard Weiner process. The maximum likelihood estimator of $\theta$, conditional upon $y_0 = b$, is

$$
\hat{\theta} = \frac{\int_0^N y_t dy_t}{\int_0^N y_t^2 dt}
$$

where $N$ is the span of the data. The discrete time representation of $y_t$ in (1.3) is given by
(1.4) \[ y_{th} = \exp(\theta h)y_{(t-1)h} + u_{th} \]

where \( u_{th} \sim N(0, \sigma^2_0(e^{2\theta h} - 1)/2\theta) \) and \( h \) is the sampling interval. (1.4) is in the form of (1.1) with \( \alpha = \exp(\theta h) \).

We derive the limiting distribution of \( T(\hat{\alpha} - \alpha) \) under (1.4) as \( h \to 0 \) and show that it is identical to the exact distribution of \( N(\hat{\theta} - \theta) \) under (1.3). Furthermore, this distribution contains an explicit dependence upon the initial condition \( y_0 \) and reduces to the near-integrated asymptotic distribution in the special case where \( y_0 = 0 \). These distributional results are considered in Section 2 which also discusses some interesting corollaries.

Section 3 derives the joint moment-generating function of \( (N^{-1} \int_0^N y_t^2 dt, N^{-2} \int_0^N y_t^2 dt) \). The expression obtained contains as a special case the limiting joint moment-generating function derived by Phillips for the 'near-integrated' framework by setting \( y_0 = 0 \). The result of White (1958) and Evans and Savin (1981a) is a special case obtained by setting \( y_0 = 0 \) and \( \theta = 0 \).

Section 4 uses the results of Section 3 to derive and tabulate the cumulative distribution and probability density function for selected combinations of the parameters \( \theta \) and \( y_0 \), and provides some measure of the adequacy of the approximation to the discrete-time finite sample distribution of \( \hat{\alpha} \).

Section 5 derives values for the bias and mean-squared error functions and discusses their use in approximating the exact mean and variance of \( \hat{\alpha} \). The issues of power are discussed in Section 6 with special emphasis on the power of tests for a unit root against alternatives close to one. Finally, Section 7 offers remarks about the usefulness of our method to other frameworks. A mathematical appendix contains the proofs of some theorems.

Throughout the paper, we provide a measure of the adequacy of the asymptotic approximation to the finite sample counterpart. The various percentage points of the
distribution, the moments of the distribution and the power functions are considered in turn. The results are quite striking, showing a very close approximation.

2. SOME DISTRIBUTIONAL RESULTS

We begin by restating our stochastic framework more formally. We consider an observable process \( \{y_t, t \geq 0\} \) defined on a probability space \((\Omega, \mathcal{F}, \mu^0)\). We denote by \( \{F_t, t \geq 0\} \) a non-decreasing family of sub-\(\sigma\)-algebras of \(\mathcal{F}\) such that \(y_t\) is \(F_t\) measurable. We denote by \(\{w_t, t \geq 0\}\) the standard Weiner process; that is, a stochastic process with independent increments defined on a probability space \((\Omega, \mathcal{F}, \mu_w)\) with \(w_0 = 0\) and where \(w(t) - w(s)\) is \(N(0, |t - s|)\). \(w_t\) is \(F_t\) measurable and the process \(w_t(s) = w(t + s) - w(t), (s \geq 0)\), is independent of \(F_t\) for any fixed \(t \geq 0\). The measure \(\mu^\theta\) is induced by the following diffusion type process postulated for \(y_t\):

\[
(2.1) \quad dy_t = \theta y_t dt + \sigma dw_t \quad ; \quad y_0 = b, t \geq 0.
\]

(2.1) is the standard Ornstein–Uhlenbeck process. \(\theta\) and \(\sigma\) are unknown parameters with \(-\infty < \theta < \infty\) and \(\sigma > 0\). The unique solution (\(\mu^\theta\) measure) in the mean squared sense to \(\{y_t\}\) is given by (see, e.g., Arnold (1974)):

\[
(2.2) \quad y_t = \exp(\theta t)b + \sigma \int_0^t \exp(\theta(t - s))dw_s.
\]

The integral that appears in (2.2) is a stochastic integral. After Phillips (1988), we adopt the following notation: \(J_\theta(t) = \int_0^t \exp(\theta(t - s))dw_s\). The solution to \(y_t\) is then written as:

\[
(2.3) \quad y_t = \exp(\theta t)b + \sigma J_\theta(t).
\]

Note that \(J_\theta(t) \sim N(0, e^{2\theta t} - 1)/2\theta\). Our concern is the estimation of the unknown parameter \(\theta\) given a single sample path of observations \(\{y_t, 0 \leq t \leq N\}\), where
N is the span of the data. The analog to the least-squares procedure in continuous time yields the following estimator:

\[
\hat{\theta}_N(y) = \int_0^N y_t \, dy_t / \int_0^N y_t^2 \, dt.
\]

\(\hat{\theta}(y)\) is also the unrestricted maximum likelihood estimator when \(b = 0\) (see, e.g., Liptser and Shiryaev (1978), 17.1.1). For simplicity of notation, we shall simply write \(\hat{\theta} = \hat{\theta}_N(y)\) and analyze the distributional properties of the standardized estimator \(N(\hat{\theta} - \theta)\).

The discrete-time representation of the process \(y_t\) is easily shown to be given by:

\[
y_{th} = \exp(\theta h) y_{(t-1)h} + u_{th}; \quad y_0 = b, \ t \geq 0
\]

where \(u_{th} \sim N(0, \sigma^2(e^{2\theta h} - 1)/2\theta)\) and \(h\) is the sampling interval. In this discrete-time framework the goal is to estimate the unknown quantity \(\alpha_h = \exp(\theta h)\) given a sequence of observations \(\{y_{th}, \ t = 0, 1, 2, ..., T\}\) where \(T = N/h\) is the total number of observations available (minus one given the initial condition \(y_0\)). The least square estimator (and maximum likelihood estimator conditional upon \(y_0\)) is

\[
\hat{\alpha}_h = \frac{\sum_{t=1}^T y_{th} y_{(t-1)h}}{\sum_{t=1}^T y_{(t-1)h}^2}
\]

We focus on the asymptotic distribution of \(T(\hat{\alpha}_h - \alpha_h)\) as \(h \to 0\) given a fixed span \(N\). For simplicity, we consider a limiting sequence \(\{h = h_1, h_2, ..., h_n\}\) such that \(T = N/h\) is integer-valued and we require that \(h_n \to 0\) as \(n \to \infty\).

The 'near-integrated' framework analyzed by Phillips (1988), Nabeya and Tanaka (1987), Perron (1988) and others specify the following process for \(y_t\):

\[
y_t = \alpha y_{t-1} + e_t, \quad y_0 = b, \ t \geq 0
\]
with \( e_t \sim N(0, \sigma^2) \) and \( \alpha = \exp(c/T) \). Here, the statistic of interest is \( T(\hat{\alpha} - \alpha) \) where \( \hat{\alpha} \) is the least-squares estimator of \( \alpha \) in (2.7) and the asymptotic distributional results are obtained letting \( T \to \infty \). Note the close relationship between the 'near-integrated' framework and the discrete-time representation of the Ornstein-Uhlenbeck process. The models are essentially the same if we interpret the parameter \( c \) as \( \theta N \) since one can write \( \alpha_h = \exp(\theta h) = \exp(\theta N/T) \).

The main result of this section ties together:

- the exact distribution of \( N(\hat{\theta} - \theta) \) given as sample path of observations \( \{y_t\}_{0}^{N} \);

- the asymptotic distribution of \( T(\hat{\alpha}_h - \alpha_h) \) with the continuous records asymptotic;

- the \( (T \to \infty) \) asymptotic distribution of \( T(\hat{\alpha} - \alpha) \) in the near-integrated framework.

**Theorem 1**

Let
\[
A(\gamma, c) = \gamma \int_{0}^{1} \exp(cr)dw_r + \int_{0}^{1} J_c(r)dr
\]

and
\[
B(\gamma, c) = \gamma^2 (\exp(2c) - 1)/2c + \gamma \int_{0}^{1} \exp(cr)J_c(r)dr + \int_{0}^{1} J_c(r)^2 dr,
\]

where \( J_c(r) = \int_{0}^{r} e^{c(r-s)}dw_s \) and \( w_r \) the standard Weiner process defined on \( C(0,1) \).

i) Let \( \{y_t, t \geq 0\} \) be a continuous-time stochastic process generated by (2.1) and let \( \hat{\theta} \) be the estimator of \( \theta \) defined by (2.4), then
\[
N(\hat{\theta} - \theta) \stackrel{d}{=} \frac{A(\gamma, c)}{B(\gamma, c)} \equiv Z(\gamma, c)
\]

with \( \gamma = b/\sigma N^{1/2} \) and \( c = \theta N \); and where \( \stackrel{d}{=} \) signifies equality in distribution.
ii) Let \( \{ y_t, t \geq 0 \} \) be a continuous time stochastic process generated by (2.1) and let \( \hat{\alpha}_h \) be defined by (2.6) with \( \{ y_{th}, t = 0, \ldots, T \} \) generated by (2.5), then as \( h \to 0 \) with \( T \to \infty \) and \( N \) fixed:

\[
T(\hat{\alpha}_h - \alpha_h) \to Z(\gamma, c)
\]

with, again, \( \gamma = b/\sigma N^{1/2} \) and \( c = \theta N \); and where \( \to \) denotes weak convergence in distribution.

iii) Let \( \{ y_t, t = 1, \ldots, T \} \) be a stochastic process generated by (2.7) and \( \hat{\alpha} \) be the least-squares estimator of \( \alpha \) from that regression, then as \( T \to \infty \):

\[
T(\hat{\alpha} - \alpha) \to Z(0, c).
\]

The proof of part (iii) is given in Phillips (1987b). The proof of part (ii) is closely related to the proof of Theorem 3.1 in Phillips (1987b) and is presented in the appendix.

To prove part (i), we consider a scale transformation from \( t \in (0, N) \) to \( r \in (0, 1) \) such that \( t \to Nr = t \). Then \( J_\theta(t) = \int_0^t e^{\theta(t-s)} dw_s \sim N(0, (e^{2\theta t} - 1)/2\theta) \) can be expressed as \( J_\theta(t) = N^{1/2} J_{c}(r) \) where \( c = \theta N \). Using (2.3), we obtain the solution

\[
y_r = \exp(\sigma c r) b + \sigma N^{1/2} J_{c}(r), \quad 0 \leq r \leq 1,
\]

with initial condition \( y_0 = b \). Now, the normalized estimator of \( \hat{\theta} \) can be expressed as:

\[
N(\hat{\theta} - \theta) = [\int_0^N y_t dy_t - \theta \int_0^N y_t^2 dt][N^{-1} \int_0^N y_t^2 dt]^{-1}
\]

\[
= \sigma \int_0^N y_t dw_t[N^{-1} \int_0^N y_t^2 dt]^{-1}
\]

using (2.1). Consider first the numerator of (2.9). Using (2.8) and the fact that \( w(t) = N^{1/2}w(r) \), we have:
\[ \int_0^N y_t \, dw_t = \int_0^1 [\exp(cr)b + \sigma N^{1/2} J_c(r)] d(N^{1/2} w_r) \]

(2.10) \[ = N^{1/2} b \int_0^1 \exp(cr) dw_r + N \sigma \int_0^1 J_c(r) dw_r. \]

The denominator of (2.9) can similarly be transformed as follows:

\[ \int_0^N y_t^2 \, dt = \int_0^1 [\exp(cr)b + \sigma N^{1/2} J_c(r)]^2 d(Nr) \]

(2.11) \[ = N b^2 (\exp(2c) - 1)/2c + 2 N^{3/2} c b \int_0^1 \exp(cr) J_c(r) \, dr \]

\[ + N^2 c^2 \int_0^1 J_c(r) \, dr, \]

using the fact that \( \int_0^1 \exp(2cr) \, dr = (\exp(2c) - 1)/2c. \) The result follows from (2.9) to (2.11).

The theorem suggests several avenues of investigation. First, the small–h (or continuous records) asymptotic distribution of \( T(\hat{\alpha}_h - \alpha_h) \) is the same as the exact distribution of the normalized continuous time estimator \( N(\hat{\theta} - \theta) \). This justifies studying the distributional properties of \( N(\hat{\theta} - \theta) \) as an approximation to the exact distribution of the discrete–time least–square estimator \( T(\hat{\alpha} - \alpha) \). Since the stochastic representation of the variable \( Z(\gamma,c) \) is explicitly affected by the value of the initial condition \( y_0 = b \), we can hope to capture the effects of different values of \( b \) on the finite sample distribution of the discrete–time least–squares estimator. More precisely, the distribution of the random variable \( Z(\gamma,c) \) depends only upon two parameters, namely \( c = \theta N \) and \( \gamma = b/\sigma N^{1/2} \). We shall therefore characterize the distributional properties of \( N(\hat{\theta} - \theta) \), in latter sections, directly as functions of these two quantities.

Second, we obtain from Theorem 1 that the asymptotic distribution of \( T(\hat{\alpha} - \alpha) \) in the near–integrated case is a special case of the exact distribution of \( N(\hat{\theta} - \theta) \) obtained by setting \( y_0 = b = 0 \). Hence, it appears that when \( b \neq 0 \), we can expect a better approximation to the exact distribution of \( \hat{\alpha} \) using the continuous record asymptotic rather than the near–integrated framework since in the latter case the effect of \( b \) vanishes asymptotically.
Several known distributional results can be obtained as special cases of the exact distribution of $N(\hat{\Theta} - \Theta)$. First when $b = 0$ we have

$$N(\hat{\Theta} - \Theta) = Z(0,c) = \int_0^1 J_c(r)dw(r) / \int_0^1 J_c(r)^2 dr$$

which is the asymptotic distribution of $T(\hat{\Theta} - \Theta)$ in the near-integrated case. Secondly, when $b = \Theta = 0$, we have:

$$N(\hat{\Theta} - \Theta) = (1/2)(w(1)^2 - 1)/ \int_0^1 w(r)^2 dr$$

which corresponds to the limiting distribution of $T(\hat{\Theta} - \Theta)$ in the unit root case where $\alpha = 1$, a result derived by Phillips (1987).

It is also straightforward to obtain different asymptotic results. Consider first the behavior of $N(\hat{\Theta} - \Theta)$ as $N \to \infty$. If $\Theta < 0$, we have $c \to -\infty$ and if $\Theta > 0$, $c \to +\infty$. Hence, the limiting distribution of $N(\hat{\Theta} - \Theta)$ as $N \to \infty$ depends upon the limiting distribution of the various functionals in $Z(\gamma,c)$ as $c \to \pm \infty$. These are stated in the following lemma.

**Lemma 1** Let $J_c(r) = \int_0^r c(r-s)dw_s$, then

i) $(-2c)\int_0^1 J_c(r)^2 dr \to 1$, as $c \to -\infty$;

ii) $(-2c)^{1/2} \int_0^1 J_c(r)dw_r \to N(0,1)$, as $c \to -\infty$;

iii) $(2c)^2e^{-2c} \int_0^1 J_c(r)^2 dr \to \sigma^2$, as $c \to +\infty$;

iv) $(2c)e^{-c} \int_0^1 J_c(r)^2 dw_r \to \xi \eta$, as $c \to +\infty$;

v) $(2c)^{3/2}e^{-2c} \int_0^1 \exp(cr)J_c(r)dr \to \eta$, as $c \to +\infty$;

vi) $(2c)^{1/2}e^{-c} \int_0^1 \exp(cr)dw_r \to 1$, as $c \to +\infty$;
where $\xi$ and $\eta$ are independent $N(0,1)$ variates. Parts (i) through (iv) of Lemma 1 are proved in Phillips (1987b) and the proof of parts (v) and (vi) are presented in the appendix.

Using Lemma 1, it is straightforward to deduce the following results concerning the asymptotic distribution of $N(\hat{\theta} - \theta)$ as $N \to \infty$:

**Corollary 1**

i) If $\theta < 0$, then for any fixed $y_0 = b$:

$$N^{1/2}(\hat{\theta} - \theta) \to N(0,-1/2\theta), \quad \text{as } N \to \infty;$$

ii) If $\theta > 0$, then for any fixed $y_0 = b$:

$$e^{\theta N(\hat{\theta} - \theta)/2\theta} \to [d\eta + \xi \eta]/[d + \eta]^2, \quad \text{as } N \to \infty;$$

where $\eta$ and $\xi$ are independent $N(0,1)$ variates and $d = b(2\theta)^{1/2}/\sigma$.

Remark: As a special case to part (ii), we have if $y_0 = b = 0$ and $\theta > 0$:

$$e^{\theta N(\hat{\theta} - \theta)/2\theta} \to \text{Cauchy}, \quad \text{as } N \to \infty.$$  

Part (i) of Corollary 1 is a standard result in the literature, see, for example, Brown and Hewitt (1975) and Basawa and Rao (1980). The above results provide a simple alternative derivation. Part (ii) appears not to have been derived previously. Note, however, that when $y_0 = b = 0$, the limiting distribution of $e^{\theta N(\hat{\theta} - \theta)/2\theta}$ is the same as the limiting distribution of $|\alpha|/\alpha^2 - 1)^{-1}(\alpha - \hat{\alpha})$ when $|\alpha| > 1$ (see White (1958)).

It is also possible to study the behavior of $N(\hat{\theta} - \theta)$ as $b \to \infty$. It is a straightforward consequence of Theorem 1 that $N(\hat{\theta} - \theta) \to 0$ as $b \to \infty$. Hence, one can expect the bias and variance of $\hat{\theta}$ to decrease as $b$ increases. The same behavior should hold concerning the discrete-time estimator $\hat{\alpha}$. These conjectures are verified in
Section 5. It is also possible to obtain the following approximating distribution of \(N(\hat{\theta} - \theta)\) for large \(\gamma\) (i.e. large \(b\)):

\[
N(\hat{\theta} - \theta) \sim \int_0^1 \exp(cr)dw_r[\gamma(\exp(2c) - 1)/2c]^{-1}
\]

Now \(\int_0^1 \exp(cr)dw(r) \sim N(0,(e^{2c} - 1)/2c)\) (see the appendix), hence,

\[
N(\hat{\theta} - \theta) \sim N(0,2c/\gamma^2(e^{2c} - 1)).
\]

Given that the distribution of \(N(\hat{\theta} - \theta)\) approximates the distribution of \(T(\hat{\alpha}_h - \alpha_h)\), we have the following approximation to the discrete–time estimator \(\hat{\alpha}_h\) for large \(b\) (or equivalently small \(\sigma\) asymptotic):

\[
(2.12) \quad T(\hat{\alpha}_h - \alpha_h) \sim N(0,2c/\gamma^2(e^{2c} - 1)).
\]

The approximation (2.12) generalizes a result derived by Phillips (1987) who considered the unit root case. Indeed, by taking the limit of (2.12) as \(c \to 0\), we have \(T(\hat{\alpha}_h - 1) \sim N(0,1/\gamma^2)\) which correspond to expression (27) in Phillips (1987).

As noted in Phillips (1987) for the unit root case, the usual \((T \to \infty)\) asymptotic distribution obscures the dependence of the distribution of \(\hat{\alpha}\) on the parameter \(\gamma_0/\sigma\). The same feature emerges in the \((T \to \infty)\) asymptotic distribution for the near–integrated case. The \((h \to 0)\) asymptotic distribution highlights the effects of the initial condition and may be a better guide to the finite sample behavior of \(T(\hat{\alpha} - \alpha)\) for values of \(\alpha\) in the vicinity of 1.

In order to use these distributional results as approximations to the finite sample behavior, we need a method to compute the distribution function of the random variable \(Z(\gamma,c)\) or equivalently of \(N(\hat{\theta} - \theta)\). To this effect, the next section derives a computable expression for this distribution.
3. THE EXACT DISTRIBUTION OF N(\(\hat{\theta} - \theta\))

Given that the exact distribution of N(\(\hat{\theta} - \theta\)) depends only upon the parameters c = \(\theta N\) and \(\gamma = b/\sigma N^{1/2}\), it is useful to transform the original model. The transformation \(t \in (0,N) \rightarrow r \in (0,1)\) with \(t = Nr\) yields

\[
N(\hat{\theta} - \theta) = \sigma N^{-1} \int_{0}^{N} y_t dw_t / N^{-2} \int_{0}^{N} y_t^2 dt
= \sigma N^{-1/2} \int_{0}^{1} y_r dw_r / N^{-1} \int_{0}^{1} y_r^2 dr
\]

where \(y_r\) is defined by (2.8). Now, let \(x_r = y_r / \sigma N^{1/2}\). Then,

\[
(3.1) \quad N(\hat{\theta} - \theta) = \int_{0}^{1} x_r dw_r / \int_{0}^{1} x_r^2 dr
\]

where

\[
(3.2) \quad x_r = \exp(cr)\gamma + J_c(r).
\]

The expression (3.2) is the solution \(x_r\) of the following stochastic differential equation:

\[
(3.3) \quad dx_r = cx_r dr + dw_r, \quad x_0 = \gamma, \quad 0 \leq r \leq 1,
\]

where \(c = \theta N\) and \(\gamma = b/\sigma N^{1/2}\).

To study the exact distribution of N(\(\hat{\theta} - \theta\)), we derive the joint moment generating function of \((\int_{0}^{1} x_r dw_r, \int_{0}^{1} x_r^2 dr)\) given that \(x_r\) is a random variable in the probability space \((\Omega, F, \mu_{\theta}^\gamma)\) generated according to the diffusion process (3.3). We denote this joint moment generating function:

\[
M_{c,\gamma}(v,u) = E[\exp(v\int_{0}^{1} x_r dw_r + u\int_{0}^{1} x_r^2 dr)]
\]
where the expectation is taken with respect to the measure $\mu_y^\theta \equiv \mu_x^c$. The main result of this section is an expression for $M_{c,\gamma}(v,u)$ which is contained in the following theorem.

**Theorem 2**

Let

$$\psi_c(v,u) = \left[ \frac{2\lambda e^{-(v+c)}}{[\lambda + v + c]e^{-\lambda} + [\lambda - (v + c)]e^{\lambda}} \right]^{1/2}$$

where

$$\lambda = (c^2 + 2cv - 2u)^{1/2}$$

then

$$M_{c,\gamma}(v,u) = \psi_c(v,u) \exp\{-\gamma^2/2\lambda(v + c - \lambda)(1 - \exp(v + c + \lambda)\psi_c^2(v,u))\}.$$

**Remark:** when $\gamma = 0$, $M_{c,\gamma}(v,u) = \psi_c(v,u)$ which is the joint moment generating function of $(\int_0^1 J(r)dw_r, \int_0^1 J(r)^2dr)$ derived by Phillips (1987b).

The proof of the theorem is inspired by the development in Lemma 17.3 of Liptser and Shiryaev (1978). Denote by $\mu_x^c$ and $\mu_x^\lambda$ the measures corresponding to the processes $x^c$ and $x^\lambda$ generated by the following stochastic differential equations:

\begin{align}
(3.4) \quad dx_r^c &= cx_r^c dr + dw_r, \quad x_0^c = \gamma \\
(3.5) \quad dx_r^\lambda &= \lambda x_r^\lambda dr + dw_r, \quad x_0^\lambda = \gamma.
\end{align}

The measures $\mu_x^c$ and $\mu_x^\lambda$ are equivalent (Liptser and Shiryaev (1978), Theorem 7.19) and the density or Radon–Nikodym derivative $d\mu_x^c/d\mu_x^\lambda$ evaluated with the random process $x^\lambda$ is given by:
\[
\frac{d\mu^c_x}{d\mu^\lambda_x}(x^\lambda) = \exp\{-(c - \lambda_1)\int_0^{1_x} dx^\lambda r - \frac{(c^2 - \lambda_2^2)}{2} \int_0^{1_x} (x^\lambda r)^2 dr\}.
\]

Hence, denoting by \(E_c\), the expectation operator taken with respect to the measure \(\mu^c_x\), we have:

\[
M_{c,\gamma}(v,u) = E_c[\exp\{v\int_0^{1_x} dw^r + u\int_0^{1_x} (x^c)^2 dr\}]
\]

\[
= E[\exp\{v\int_0^{1_x} dw^c + u\int_0^{1_x} (x^c)^2 dr\}]
\]

\[
= E[\exp\{v\int_0^{1_x} dw^c + u\int_0^{1_x} (x^c)^2 dr\} \cdot \frac{d\mu^c_x}{d\mu^\lambda_x}(x^\lambda)]
\]

\[
= E[\exp\{v\int_0^{1_x} dw^c + u\int_0^{1_x} (x^c)^2 dr\} + (c - \lambda_1)\int_0^{1_x} dx^\lambda r - \frac{(c^2 - \lambda_2^2)}{2} \int_0^{1_x} (x^\lambda r)^2 dr\}]
\]

Using the fact that \(\int_0^{1_x} dw^c = \int_0^{1_x} dx^\lambda r - \lambda_1 \int_0^{1_x} (x^\lambda r)^2 dr\) with (3.5) and rearranging:

\[
M_{c,\gamma}(v,u) = E[\exp\{(v + c - \lambda_1)\int_0^{1_x} dx^\lambda r + [u - vc - (c^2 - \lambda_2^2)/2] \int_0^{1_x} (x^\lambda r)^2 dr\}]
\]

Now, let \(\lambda = (c^2 + 2cv - 2u)^{1/2}\) and \(a = v + c - \lambda\). Then, \([u - vc - (c^2 - \lambda^2)/2] = 0\) and

\[
M_{c,\gamma}(v,u) = E[\exp[a\int_0^{1_x} dx^\lambda r]]
\]

By Ito's lemma applied to the function \((x_1^\lambda)^2\), we have \(\int_0^{1_x} dx^\lambda r = (x_1^\lambda - x_0^\lambda - 1)/2\).

Hence,

\[
M_{c,\gamma}(v,u) = E[\exp{(a/2)((x_1^\lambda)^2 - (x_0^\lambda)^2)}]
\]
(3.6) \quad \gamma = \exp((-a/2)(\gamma^2 + 1))E[\exp((a/2)(x_1^\lambda)^2)].

since \( x_0^\lambda = \gamma \), a fixed constant. Now, the solution to (3.5) is given by:

\[
x^\lambda_r = \exp(\lambda r)\gamma + \int_0^r \exp(\lambda(r-s))dw_s
\]

\[
= \exp(\lambda r)\gamma + J^\lambda(1)
\]

and, hence,

\[
x^\lambda_1 = \exp(\lambda)\gamma + J^\lambda(1).
\]

Now \( J^\lambda(1) \sim N(0,(e^{2\lambda} - 1)/2\lambda) \) and therefore \( x^\lambda_1 \) is \( N(\exp(\lambda)\gamma, (e^{2\lambda} - 1)/2\lambda) \). Let \( s^2 = (e^{2\lambda} - 1)/2\lambda \) and \( q = \exp(\lambda)\gamma \) then \( (x_1^\lambda)^2/s^2 \) is distributed as a non-central chi-squared variate with non-centrality parameter \( q \). Then

\[
E[\exp((a/2)(x_1^\lambda)^2)] = E[\exp((a/2)s^2(x_1^\lambda/s)^2)]
\]

\[
= (1 - as^2)^{-1/2}\exp((aq^2/2)(1 - as^2)^{-1})
\]

and

\[
M_{c,\gamma}(v,u) = \exp((-a/2)(\gamma^2 + 1))(1 - as^2)^{-1/2}\exp((aq^2/2)(1 - as^2)^{-1}).
\]

Upon substitution for \( s = (e^{2\lambda} - 1)/2\lambda \) and \( a = v + c - \lambda \), it is straightforward to verify that

\[
\psi_c(v,u) = (1 - as^2)^{-1/2}\exp(-a/2).
\]
It follows

\[ M_{c,y}(v,u) = \psi_c(v,u)\exp\{(-a\gamma^2/2) + (aq^2/2)(1 - as^2)^{-1}\} \]

\[ = \psi_c(v,u)\exp\{(-a\gamma^2/2)(1 - \exp(2\lambda + a)\psi_c(v,u)^2)\} \]

which reduces to the expression in Theorem 1 upon substitution for \( a = v + c - \lambda \). \( \square \)

The moment-generating function derived in Theorem 2 allows us to compute the cumulative distribution function as well as other distributional quantities using numerical integration. The next sections present the results of such computations.
4. THE COMPUTATION OF THE C.D.F. AND P.D.F. OF $N(\hat{\theta} - \theta)$

Denote the joint characteristic function of $(\int_0^{1-r} dw_r, \int_0^{1-r} x^2 dr)$ by $c_{C, \alpha}(v, u)$. Then,

$$(4.1) \quad c_{C, \alpha}(v, u) = M_{C, \alpha}(iv, u)$$

$$= E[\exp(iv\int_0^{1-r} dw_r + iu\int_0^{1-r} x^2 dr)].$$

The distribution function of $N(\hat{\theta} - \theta)$ can be obtained as follows. Let $F_{C, \alpha}(z) = P[N(\hat{\theta} - \theta) \leq z]$ and recall that $N(\hat{\theta} - \theta) = \int_0^{1-r} dw_r / \int_0^{1-r} x^2 dr$. Then by Theorem 1 of Gurland (1948), we have

$$F_{C, \alpha}(z) = \frac{1}{2} - \frac{1}{2\pi} \lim_{\varepsilon_1 \to 0, \varepsilon_2 \to \infty} \int_{\varepsilon_1 < |v| < \varepsilon_2} \frac{c_{C, \alpha}(v, -vz)}{v} dv$$

$$(4.2) \quad = \frac{1}{2} - \frac{1}{2\pi} \int_0^{\infty} \text{AIMAG}\left[\frac{c_{C, \alpha}(v, -vz)}{v}\right] dv$$

where AIMAG(·) denotes the imaginary part of the complex number. Further, the density function is derived as follows:

$$f_{C, \alpha}(z) = \frac{d}{dz} F_{C, \alpha}(z)$$

$$= \frac{1}{2\pi} \lim_{\varepsilon_1 \to 0, \varepsilon_2 \to \infty} \int_{\varepsilon_1 < |v| < \varepsilon_2} \left. \frac{\partial c_{C, \alpha}(v, u)}{\partial u} \right|_{u = -vz} dv$$

$$(4.3) \quad = \frac{1}{\pi} \int_0^{\infty} \text{Real}[c_{C, \alpha}(v, -vz)(A1 + A2/2)] dv$$

where $A1 = (-\gamma^2/2) \{(1 - \exp(iv + c + \lambda)[\psi_c(iv, -ivz)]^2)$
\[ + (iv + c - \lambda) \exp(iv + c + \lambda) \psi_c(iv, -ivz)(-\lambda^{-1} + A2) \]

and

\[ A2 = -\lambda^{-2} + D1/D2 \]

\[ D1 = \lambda^{-1} e^{-\lambda}[\lambda + (iv + c) - 1] + \lambda^{-1} e^{\lambda}[\lambda - (iv + c) + 1] \]

\[ D2 = [\lambda + (iv + c)]e^{-\lambda} + [\lambda - (iv + c)]e^{\lambda} \]

\[ \lambda = (c^2 + 2iv(c + z))^{1/2} \]

and \( \text{Real}(\cdot) \) denotes the real part of the complex number.

The expressions (4.2) and (4.3) can be numerically integrated to obtain values for the cumulative distribution function and the probability density function. When calculating these functions, we evaluate the integrals in the range \((0 + \varepsilon, \bar{V})\) where \(\bar{V}\) is an upper bound set such that the integrand evaluated at \(\bar{V}\) is less than \(\varepsilon\) \((\varepsilon\) was set at 1.0E−07 in each integration). The integrals are then evaluated in this range using the subroutine DCADRE of the International Mathematical and Statistical Library (IMSL) (the error criterion for this routine integration was also set at 1.0E−07). Special care, however, must be applied to the integration since it involves the square root of a complex valued function. The use of the principal value of the square root may not ensure the continuity of the intergrand. We must therefore integrate over the Reiman surface consisting here of two planes. The method used is described in more details in the Appendix to Perron (1988).

Figures 1 through 3 present the cumulative distribution functions of \(N(\hat{\theta} - \theta)\) for \(c = -5, 0\) and 2 respectively. Each figure contains four curves corresponding to the following values for \(\gamma\): 0, 0.5, 1.0 and 2.0. Figures 4 through 6 contain the corresponding results for the probability density function. Each curve was constructed by evaluating the c.d.f. (or p.d.f.) at 100 equidistant points in the relevant range.

Several striking features emerge from these graphs. First, the distributions are sharply non-normal. Second, the shapes are significantly affected by changes in the values of \(\gamma\) and \(c\). The most important feature is that the distribution becomes more closely centered and more concentrated near zero as \(\gamma\) and \(c\) increase. As \(c\) increases,
the distribution becomes less skewed to the left and shows much less variability around zero. A similar behavior occurs as $\gamma$ increases. Another interesting feature is the effect of increasing $\gamma$ for different values of $c$. For a given change in $\gamma$, the effect on the shape of the distribution is more pronounced the larger $c$ is. In other words, the effects of different values of $\gamma$ are smaller when $\theta$ is smaller (i.e. "more stationary") or when $N$ is larger, provided $\theta < 0$. When $\theta > 0$, the effect of the initial condition is considerable.

To summarize the main features of the various distributions, Table 1 presents selected percentage points of the distribution of $N(\hat{\theta} - \theta)$ for various values of $c$ and $\gamma$.

To assess the adequacy of the asymptotic approximation to the finite sample distribution, we proceed as follows. We simulated the least-squares estimator of $\hat{\alpha}$ from an AR(1) model with $N(0,1)$ errors. The number of replications was set at 20,000. We obtained the percentage points of the distribution of $T(\hat{\alpha} - \alpha)$ for selected values of $T$, $\alpha$ and $\gamma_0 = b$. The correspondence between the exact distribution and the asymptotic approximation is obtained by setting $c = T\ln(\alpha)$ (since $\alpha = \exp(\theta N/T)$) and $\gamma = b/T^{1/2}$ since $\sigma^2 = 1$. Table 2 presents the results for $T = 10, 25, 50$ and 100 with $c = -5, -2, 0$ and 2 and $\gamma = 0.5, 1.0, 2.0$. We omit the case where $\gamma = 0.0$ since it was thoroughly analysed in Nabeya and Tanaka (1987), Perron (1988) and Cavanagh (1986). These studies found the approximation to be adequate for various values of $c$. The approximation deteriorates as $c$ becomes smaller.

In our more general context, the approximation remains quite good especially for values of $T$ at 50 or 100. The approximation gets more accurate as 1) $c$ increases, 2) $\gamma$ increases and 3) if we consider the right tail rather than the left. It is also relatively less accurate in the extreme tails. For example, consider the case with $c = 2$, $\gamma = 2$ and the 5th percentage point. The asymptotic approximation is $-0.3294$ and the exact values range from $-0.338$ for $T = 10$ to $-0.325$ for $T = 100$, showing a high degree of accuracy even for a very small sample size.

On the other hand, when $c$ is small, the convergence to the asymptotic value is slower. For example, consider $c = -5$, $\gamma = 0.5$ and the 5th percentage point. The asymptotic critical value is $-8.755$ while the exact values range from $-5.442$ ($T = 10$) to $-8.314$ ($T = 100$). However, even for small values of $c$, it gets more accurate as $\gamma$
increases. For example, when $c = -5.0$, $\gamma = 2.0$, the asymptotic critical value for the 95th percentage point is 1.7357 and the exact values range from 1.464 ($T = 10$) to 1.675 ($T = 100$).

Overall, the asymptotic distribution approximates rather well the major changes in the distribution as $c$ and $\gamma$ are varied and for sample sizes typically available the approximation is rather good.
5. THE BIAS AND MEAN–SQUARED ERROR OF \( \hat{\theta} \)

The moment–generating function derived in Section 3 can be used to derive the moments of \( N(\hat{\theta} - \theta) \). Using Mehta and Swamy's (1978) result, we have:

(5.1) \[ N \cdot \text{Bias}(\hat{\theta}) = E[N(\hat{\theta} - \theta)] = \int_0^{\infty} \frac{\partial M_{c, \gamma}(v, -u)}{\partial v} \bigg|_{v=0} \, du \]

and

(5.2) \[ N^2 \cdot \text{MSE}(\hat{\theta}) = E[N(\hat{\theta} - \theta)]^2 = \int_0^{\infty} u \frac{\partial^2 M_{c, \gamma}(v, -u)}{\partial v^2} \bigg|_{v=0} \, du \]

where \( M_{c, \gamma}(v, u) = E[\exp(v \int_0^1 x_r \, dw_r + u \int_0^1 x_r^2 \, dr)] \)

is the joint moment–generating function of \((\int_0^1 x_r \, dw_r, \int_0^1 x_r^2 \, dr)\) derived in Theorem 2.

Straightforward computations give the following results:

\[ \frac{\partial M_{c, \gamma}(v, u)}{\partial v} \bigg|_{v=0} = M_{c, \gamma}(0, -u) [DH + [(DA/A) - (DB/B)]/2] \]

where \( DH = -\gamma^2 (1 - c/\lambda)(1 - \exp(c + \lambda)[\psi_c(0, -u)]^2)/2 \)

\[ + (1/2)(c - \lambda)\gamma^2 \exp(c + \lambda)[\psi_c(0, -u)]^2 ((1 + c/\lambda) + DA/A - DB/B) \]

\( A = 2\lambda e^{-c} \)

\( B = [\lambda + c]e^{-\lambda} + [\lambda - c]e^{\lambda} \)

\( DA = 2e^{-c}[c - \lambda^2]/\lambda \)

\( DB = (1 - c)[e^{-\lambda}(c + \lambda) + e^{\lambda}(c - \lambda)]/\lambda \)

\( \lambda = (c^2 + 2u)^{1/2} \)
and

\[
\left. \frac{\partial^2 M_{c, \gamma}(v, -u)}{\partial v^2} \right|_{v=0} = M_{c, \gamma}(0, -u) \{DH + [(DA/A) - (DB/B)]/2 \}^2 \\
+ M_{c, \gamma}(0, -u)\{D2H + (D2A - D2B)/2\}
\]

where

\[
D2A = -2c^2/\lambda^4
\]

\[
D2B = DBW/B - (DB)^2/B^2
\]

\[
D2W = ce^{-\lambda}[\lambda(c - 2) + c(c - 1) + c^2/\lambda]/\lambda^2
\]

\[
+ ce^{\lambda}[\lambda(c - 2) - c(c - 1) + c^2/\lambda]/\lambda^2
\]

and

\[
D2H = DH(1 - c/\lambda)/(c - \lambda) + H[c^2/\lambda^3(c - \lambda) - 1/\lambda^2]
\]

\[
+ [DH + (1 - c/\lambda)\gamma^2/2][(1 + c/\lambda) + DA/A - DB/B]
\]

\[
+ [H + (c - \lambda)\gamma^2/2][-c^2/\lambda^3 - 2c^2/\lambda^4 - DBW/B + (DB)^2/B^2].
\]

Figures 7 and 8 present, respectively, the normalized bias and mean–squared error of \( \hat{\theta} \) (i.e. the first two moments of \( N(\hat{\theta} - \theta) \)) as a function of \( c = \theta \N \) for various values of \( \gamma = b/\sigma N^{1/2} \). The conjectures discussed in Section 2 are verified. The bias and mean–squared error are smaller (in absolute value) as \( c \) increases. Furthermore, both decrease (in absolute value) as \( \gamma \) increases. An increase in \( \gamma \) has a larger reduction effect on both the bias and the mean squared error when \( c \) is small. But this is mainly due to the fact that the bias and MSE are already quite small for large \( c \) when \( \gamma = 0 \).

Both the bias and MSE converge quite rapidly to zero as \( c \) and/or \( \gamma \) increase. On the other hand, the bias appears to attain an upper limit (in absolute value) value when \( c \to \infty \) for any fixed \( \gamma \) and it approaches this limit fairly rapidly. The limit of the bias as \( c \to \infty \) is a decreasing (in absolute value) function of \( \gamma \).
To assess the relevance of the asymptotic approximation to the exact moments of \( \hat{\alpha} \) we consider applying our approximation to the experimental setting in Table III of Evans and Savin (1981b) which presents the exact mean and variance of a normalized estimator of \( \hat{\alpha} \); \( \alpha = 0.90, 0.99, 1.00, 1.01 \) and \( 1.05, y_0/\sigma = 0, 4, 16 \) and \( T = 25, 50, 100, 200 \) and 400. The normalization is defined as \( g(T)(\hat{\alpha} - \alpha) \) where

\[
g(T) = \begin{cases} 
T(1 - \alpha^2)^{-1} & \text{if } |\alpha| < 1; \\
T/\sqrt{2} & \text{if } |\alpha| = 1; \\
|\alpha|T(\alpha^2 - 1)^{-1} & \text{if } |\alpha| > 1.
\end{cases}
\]

To approximate the moments of \( g(T)(\hat{\alpha} - \alpha) \), we consider the correspondence \( c = T\ln(\alpha) \) and \( \gamma = y_0/\sigma T^{1/2} \) and use (5.1) and (5.2).

The results are presented in Tables 3 and 4 for the mean and the variance of \( g(T)(\hat{\alpha} - \alpha) \) respectively. The values in parentheses are taken from Table III of Evans and Savin (1981b).

The approximation is, in general, very good and certainly captures most of the changes in the distribution as \( \alpha \), \( y_0/\sigma \) and \( T \) are varied. As expected, the approximation is better as \( \alpha \) is closer to 1 and as \( T \) is larger. But it is also better as \( y_0/\sigma \) increases.

We can also compare our approximation to White's approximation when \( y_0 = 0 \) and \( \alpha < 1 \) (presented in Table III of Evans and Savin (1981b)). Our approximation is better when \( \alpha = 0.99 \) and slightly worse when \( \alpha = 0.90 \). Our approximation, however, has the advantage of capturing the effects of a non–zero initial condition.

Overall, our procedure appears to provide a very adequate approximation to the exact moments of \( \hat{\alpha} \) when \( \alpha \) is in the vicinity of one allowing the initial condition \( y_0/\sigma \) to be an arbitrary fixed value.
6. THE POWER FUNCTION

Using the statistic $N(\hat{\theta} - \theta_0)$, the results of Section 2 can also be used to analyze the power function for tests of the null hypothesis

(6.1) $H_0: \theta = \theta_0$

against various alternatives. Denote by $z^*$ the value such that $P_{\theta_0}[N(\hat{\theta} - \theta_0) \leq z^*] = \beta$. Then, the power function of a one-sided test with size $\beta$ for testing $\theta = \theta_0$ against $\theta < \theta_0$ is given by

(6.2) $P_{\theta}[N(\hat{\theta} - \theta_0) < z^*] = P_{\theta}[N(\hat{\theta} - \theta) < z^* + N(\hat{\theta} - \theta_0)]$

$= P_{\theta}[Z(\gamma,c) < z^* + (c - c_0)]$

with $c = N\theta$ and $c_0 = N\theta_0$. Expression (4.2) can be used to evaluate the power function (6.2) for various values of $c$, $\gamma$, $c_0$ and significance level $\beta$.

Much attention has recently been given to testing the null hypothesis of a unit root ($\alpha = 1$, $\theta = 0$). Letting $\theta_0 = 0$ and using (6.2), the power function of a one-sided test of the unit root hypothesis using the statistic $\hat{N}$ is given by

(6.3) $P_{\theta}[\hat{N} < z_0^*] = P_{\theta}[N(\hat{\theta} - \theta) < z_0^* - c]$

$= P_{\theta}[Z(\gamma,c) < z_0^* - c]$

which can be evaluated using (4.2). Here, $z_0^*$ is the critical value such that $P[Z(\gamma,0) \leq z_0^*] = \beta$ where $\beta$ is the size of the test.
Interestingly, (6.3) gives the local asymptotic power function of a test for a unit root using the statistic $T(\hat{\alpha} - 1)$ (in discrete–time) where the sequence of local alternatives is given by

$$\alpha = \exp(c/T)$$

and the asymptotic is derived as $h \to 0$, i.e. using the continuous records asymptotic instead of the usual $(T \to \infty)$ asymptotic. To elaborate, consider the following derivation:

$$\lim_{h \to 0} P[\alpha(T(\hat{\alpha} - 1) < z_0^*) = \lim_{h \to 0} P[\alpha(T(\hat{\alpha} - \alpha) < z_0^* + T(\alpha - 1)]$$

$$= P[Z(\gamma,c) < z_0^* - c]$$

since $T(\hat{\alpha} - \alpha) \to Z(\gamma,c)$ as $h \to 0$ (Theorem 1) and $T(\alpha - 1) \to c$ as $h \to 0$.

This framework generalizes the analysis of Phillips (1987b) who considers the local asymptotic power in the usual $(T \to \infty)$ asymptotic context. Here, the continuous records asymptotic ($h \to 0$) permits explicit consideration of the effect of a non–zero initial condition. Our results specialize to those of Phillips by setting $\gamma = 0$. Note finally, that when $\gamma = 0$, the critical values $z_0^*$ correspond to those derived by Dickey and Fuller (1979); see Fuller (1976).

Figure 9 presents the power function of a one–sided 5 % size test that $\theta = 0$ against both positive and negative alternative values of $c = \theta N$. Four curves are drawn corresponding to values of $\gamma$ at 0, 0.5, 1.0 and 2.0. Several features emerge. First, as expected from the studies of Evans and Savin (1981b) and others, the power is much higher against explosive alternatives ($c > 0$). Secondly, the power function is significantly influenced by different values of the startup condition $\gamma$. Indeed, the power increases as $\gamma$ increases. There is an interesting analogy with the bias function (see figure 7) in that changes in $\gamma$ have a larger effect on the power function for stationary alternatives. This is due to the fact that when $\gamma = 0$ the power function against explosive alternatives is already quite high compared to the power function under stationary alternatives.
To assess the adequacy of our asymptotic power function in the context of finite samples tests for a unit root against alternatives close to one, we considered the experimental framework adopted in Evans and Savin (1981b). The exact power of tests for a unit root against the alternatives $\alpha = 0.9, 0.95, 0.99, 1.01, 1.025$ and $1.05$ can be found in their Table IV. For each value of $\alpha$, they vary as the initial condition $y_0/\sigma = 0, 2$ and $4$ and $T = 25, 50, 100, 200$, and $400$. Our results are presented in Table 5 where, for ease of comparison, we have included in parentheses the exact values derived by Evans and Savin (1981b). The correspondence between the finite sample and asymptotic frameworks is again achieved by specifying $c = T \ln(\alpha)$ and $\gamma = y_0/\sigma T^{1/2}$.

The results are quite striking. For almost all combinations of parameters the approximation is indeed excellent. The worst fit occurs for a small sample size ($T = 25$) and large values of $\alpha$ ($1.05$). The conclusions reached by Evans and Savin (1981b) are well–captured by the asymptotic approximations.
7. CONCLUDING COMMENTS

This paper has considered an alternative asymptotic approximation to the least-squares estimator of the parameter in a first-order autoregressive model with a fixed initial condition. This approach is different in that we consider a framework using the continuous records asymptotics (h → 0) instead of the usual (T → ∞) asymptotic theory. The main advantage over the (T → ∞) near-integrated framework is that our method allows explicit consideration of the effects of non-zero initial conditions.

The continuous record asymptotic distribution for the discrete-time estimator is shown to be equivalent to the distribution of the continuous time estimator in an Ornstein-Uhlenbeck process. The exact distribution of this continuous time estimator was derived. It was found to yield interesting theoretical results for continuous time estimation. More importantly, it has proved to be a good approximation to the finite sample distribution of the discrete-time least squares estimator.

This study indicates the usefulness of the continuous records asymptotic distributional theory which can be used in other contexts. One such possible extension is the possibility to approximate the finite sample distribution of the least-squares estimator $\hat{\alpha}$ when a random initial condition is specified. Indeed, equation (3.6) can be used to derive the moment-generating function of $N(\hat{\theta} - \theta)$ with a random initial condition as well. An extensive study of the case where the stochastic process $\{y_t\}$ is stationary is presented in Perron (1988b).

Further possible extensions include the case where a fitted mean is included in the regression. Here, the parameter affecting the shape of the distribution is a normalization of the mean of the variables instead of the initial condition per se (see Evans and Savin (1984)). Again, the analysis can distinguish between a fixed and random initial condition. Such an analysis is presented in a companion paper, Perron (1988c). It remains to be seen whether analogs to Theorem 2 can be obtained in the context of more general models and different statistics such as the t-ratio.
MATHEMATICAL APPENDIX

Proof of Theorem 1, part (ii)

The proof of Theorem 1, part (ii) is closely related to the proof of Theorem 3.1 in Phillips (1987b). We consider a triangular array of random variables \( \{ y_{nt} \}_{t=1}^{T_n} \). For a given \( n \), the sequence \( \{ y_{nt} \}_{n=1}^{T_n} \) is generated by (see equation (2.5))

\[
(A.1) \quad y_{nt} = \exp(\theta h_n) y_{n(t-1)} + u_{nt}, \quad t = 1, \ldots, T_n
\]

where the innovation sequence \( \{ u_{nt} \}_{t=1}^{T_n} \) is i.i.d. normal with mean 0 and variance \( \sigma^2 (\exp(2\theta h_n) - 1)/2\theta \). We have \( T_n = N/h_n \) and \( T_n \to \infty, h_n \to 0 \) as \( n \to \infty \).

Let the sequence \( \{ \epsilon_{nt} \}_{t=1}^{T_n} \) be defined by \( \epsilon_{nt} = a(h_n)^{-1/2} u_{nt} \) where \( a(h_n) = (\exp(2\theta h_n) - 1)/2\theta \). Then \( \{ \epsilon_{nt} \}_{t=1}^{T_n} \) is a triangular array of i.i.d. \( N(0, \sigma^2) \) variates. Note that \( \lim_{h_n \to 0} h_n^{-1} a(h_n) = 1 \). Now from (A.1), we have:

\[
(A.2) \quad y_{nt} = \sum_{j=1}^{t} \exp((t-j)\theta h_n) u_{nj} + \exp(t\theta h_n) y_0
\]

\[
= a(h_n)^{1/2} \sum_{j=1}^{t} \exp((t-j)\theta h_n) \epsilon_{nj} + \exp(t\theta h_n) y_0.
\]

Now define the following partial sums \( Z_{nt} = \sum_{j=1}^{t} \epsilon_{nj} \) with \( Z_{n0} = 0 \) and the random elements

\[
X_{T_n}(r) = T_n^{r} \sigma^{-1} Z_{T_n r} = T_n^{r} \sigma^{-1} Z_{nj-1} \quad (j-1)/T_n \leq r < j/T_n
\]

\[
X_{T_n}(1) = T_n^{1/2} \sigma^{-1} Z_{T_n}
\]

\[
(j = 1, \ldots, T_n)
\]
then, (see Phillips (1987a)), $X^{\nu}_{n}(r) \rightarrow w(r)$, a Weiner process on $C[0,1]$.

To prove Theorem 1, we shall derive the limiting distribution of the following quantities: a) $y_{T_n}$; b) $\sum_{t=1}^{T_n} y_{tn}$; c) $\sum_{t=1}^{T_n} y_{nt}^2$; and d) $\sum_{t=1}^{T_n} y_{nt-1}^1 u_{nt}^1$

A) $y_{T_n}$:

$$y_{T_n} = N^{1/2} T^{-1/2}_n h^{1/2}_n \sum_{j=1}^{T_n} \exp((T_n - j)\theta h_n) u_{nj} + \exp(T_n \theta h_n) y_0$$

$$= N^{1/2} h^{1/2}_n a(h_n)^{1/2} \sum_{j=1}^{T_n} \exp((1 - j/T_n)\theta N) \int_{(j-1)/T_n}^{j/T_n} dX_{T_n}(s) + \exp(\theta N) y_0$$

$$= N^{1/2} h^{1/2}_n a(h_n)^{1/2} \sum_{j=1}^{T_n} \exp((1 - s)\theta N) dX_{T_n}(s) + \exp(\theta N) y_0$$

$$= N^{1/2} a(h_n)^{1/2} \int_0^1 \exp((1 - s)\theta N) dX_{T_n}(s) + \exp(\theta N) y_0$$

$$\rightarrow N^{1/2} \sigma f_{c}(1) + \exp(\theta N) y_0, \quad \text{as } n \rightarrow \infty.$$  

B) $\sum_{t=1}^{T_n} y_{tn}$:

$$N^{-1/2} T^{-1}_n \sum_{t=1}^{T_n} y_{tn} = h^{-1/2}_n T^{-3/2}_n \sum_{t=1}^{T_n} \sum_{j=1}^{t} \exp((t - j)\theta h_n) u_{nj}$$

$$+ h^{1/2}_n T^{-3/2}_n \sum_{t=1}^{T_n} \exp(t\theta h_n) y_0$$

$$= \sigma a(h_n)^{1/2} h^{-1/2}_n T^{-1}_n \sum_{t=1}^{T_n} \sum_{j=1}^{t} \exp((t - j)\theta h_n) \int_{(j-1)/T_n}^{j/T_n} dX_{T_n}(s)$$
\[ + h_n^{-1/2} T_n^{-3/2} \sum_{t=1}^{T_n} \exp(t \theta h_n) y_0 \]

\[ = \sigma (h_n) \frac{1}{2} h_n^{-1/2} \sum_{t=1}^{T_n} \sum_{j=1}^{T_n} \int_{(j-1)/T_n}^{j/T_n} \exp((t-j) \theta h_n) dX_{T_n}(s) \]

\[ + h_n^{-1/2} T_n^{-3/2} \sum_{t=1}^{T_n} \exp(t \theta h_n) y_0 \]

\[ = \sigma (h_n) \frac{1}{2} h_n^{-1/2} \sum_{t=1}^{T_n} \sum_{j=1}^{T_n} \int_{(j-1)/T_n}^{j/T_n} \int_{(t-1)/T_n}^{t/T_n} \exp((r-s) \theta N) dX_{T_n}(s) \] \[ \text{dr} \]

\[ + h_n^{-1/2} T_n^{-3/2} \sum_{t=1}^{T_n} \exp(t \theta h_n) y_0 \]

\[ = \sigma (h_n) \frac{1}{2} h_n^{-1/2} \int_{0}^{1} \int_{0}^{r} \exp((r-s) \theta N) dX_{T_n}(s) \text{dr} \]

\[ + h_n^{-1/2} T_n^{-3/2} \sum_{t=1}^{T_n} \exp(t \theta h_n) y_0 \]

\[ = \sigma (h_n) \frac{1}{2} h_n^{-1/2} \int_{0}^{1} \left( X_{T_n}(r) + \theta N \int_{0}^{r} \exp((r-s) \theta N) dX_{T_n}(s) \right) \text{dr} \]

\[ + h_n^{-1/2} T_n^{-3/2} \sum_{t=1}^{T_n} \exp(t \theta h_n) y_0 \]

and \[ N^{-1/2} T_n^{-1} \sum y_{nt} \to \sigma \int_{0}^{1} c(r) \text{dr} + y_0 N^{-3/2} (\exp(\theta N) - 1) / \theta \] as \( n \to \infty \).

C) \[ \sum_{t=1}^{T_n} y_{nt}^2 : \]

\[ N^{-1} T_n^{-1} \sum_{t=1}^{T_n} y_{nt}^2 = N^{-1} T_n^{-1} \sum_{t=1}^{T_n} \left( \sum_{j=1}^{T_n} \exp((t-j) \theta h_n) u_{nj} + \exp(t \theta h_n) y_0 \right)^2 \]
\[
= N^{-1} T_n^{-1} \sum_{t=1}^{T_n} a(h_n^t) \sum_{j=1}^{T_n} \exp((t-j)\theta h_n^j) e_{n_j}^2
\]

\[
+ 2 N^{-1} T_n^{-1} y_0 a(h_n^{1/2}) \sum_{t=1}^{T_n} \exp(t\theta h_n^t) \sum_{j=1}^{T_n} \exp((t-j)\theta h_n^j) e_{n_j}^2
\]

\[
+ N^{-1} T_n^{-1} \sum_{t=1}^{T_n} \exp(2t\theta h_n^t) y_0^2
\]

\[
= \sigma^2 h_n^{-1} a(h_n^t) T_n^{-1} \sum_{t=1}^{T_n} (\sum_{j=1}^{T_n} \exp((t-j)\theta h_n^j)) \int_{(j-1)/T_n}^{j/T_n} dX_T(s)^2
\]

\[
+ 2\sigma h_n^{-1} a(h_n^{1/2}) T_n^{-3/2} y_0 \sum_{t=1}^{T_n} \exp(t\theta h_n^t) \sum_{j=1}^{T_n} \exp((t-j)\theta h_n^j) \int_{(j-1)/T_n}^{j/T_n} dX_T(s)
\]

\[
+ N^{-1} T_n^{-1} \sum_{t=1}^{T_n} \exp(2t\theta h_n^t) y_0^2
\]

\[
= \sigma^2 h_n^{-1} a(h_n^t) \sum_{t=1}^{T_n} \int_{(t-1)/T_n}^{t/T_n} \text{dr}(\sum_{j=1}^{T_n} \int_{(j-1)/T_n}^{j/T_n} \exp((r-s)\theta N) dX_T(s)^2)
\]

\[
+ 2\sigma h_n^{-1} a(h_n^{1/2}) N^{-1/2} y_0 \sum_{t=1}^{T_n} \sum_{j=1}^{T_n} \int_{(t-1)/T_n}^{t/T_n} \exp(r\theta N) dr \int_{(j-1)/T_n}^{j/T_n} \exp((r-s)\theta N) dX_T(s)
\]

\[
+ N^{-1} T_n^{-1} \sum_{t=1}^{T_n} \exp(2t\theta h_n^t) y_0^2
\]

\[
= \sigma^2 h_n^{-1} a(h_n^t) \int_0^1 \text{dr} \left( \int_0^1 \exp((r-s)\theta N) dX_T(s) \right)^2
\]

\[
+ 2\sigma h_n^{-1} a(h_n^{1/2}) N^{-1/2} y_0 \int_0^1 \exp(r\theta N) dX_T(s)
\]
\[ + N^{-1}T^{-1} \sum_{t=1}^{T} \exp(2t\theta h_n)y_0^2 \]
\[ = \sigma^2 h_n^{-1}a(h_n) \int_0^1 (X_{T_n}(r) + \theta N \int_0^r \exp((r-s)\theta N)X_{T_n}(s)ds) dr \]
\[ + 2\sigma h_n^{-1/2}a(h_n)^{1/2}N^{-1/2}y_0 \int_0^1 \exp(r\theta N)(X_{T_n}(r) + \theta N \int_0^r \exp((r-s)\theta N)X_{T_n}(s)ds) dr \]
\[ + N^{-1}T^{-1} \sum_{t=1}^{T} \exp(2t\theta h_n)y_0^2 \]
\[ - \sigma^2 \int_0^1 c(r)^2 dr + 2\sigma y_0 N^{-1/2} \int_0^1 \exp(rc)J_c(r)dr + y_0^2 N^{-2} (\exp(2\theta N) - 1)/2\theta \]

as \( n \to \infty \)

D) \( \sum_{t=1}^{T} y_{nt-1}u_{nt}^2 \):

Note that by squaring (A.1) we obtain

\[ y_{nt}^2 = \exp(2\theta h_n)y_{nt-1}^2 + 2\exp(\theta h_n)y_{nt-1}u_{nt} + u_{nt}^2 \]

\[ y_{nt}^2 - y_{nt-1}^2 = (\exp(2\theta h_n) - 1)y_{nt-1}^2 + 2\exp(\theta h_n)y_{nt-1}u_{nt} + u_{nt}^2 \]

summing both sides over \( t=1, ..., T_n \), we get

\[ y_{T_n}^2 - y_0^2 = (\exp(2\theta h_n) - 1)\sum y_{nt-1}^2 + 2\exp(\theta h_n)\sum y_{nt-1}u_{nt} + \sum u_{nt}^2 \]
and

$$N^{-1}\sum_{nt} u_{nt} = (1/2)\exp(-\theta h_n) [N^{-1}y_n^2 T_n - N^{-1}y_0^2 - T_n (\exp(2\theta h_n) - 1)N^{-1}T_n^{-1}\sum_{nt} y_{nt-1}^2 - N^{-1}\sum u_{nt}^2].$$

Now,

$$N^{-1}\sum u_{nt}^2 = N^{-1}a(h)\sum e_{nt}^2 = h_n^{-1}a(h)T_n^{-1}\sum e_{nt}^2 \rightarrow \sigma^2,$$

and

$$N^{-1}\sum y_{nt-1} u_{nt} \rightarrow (1/2)[(\sigma J_c(1) + N^{-1/2}\exp(\theta N)y_0)^2 - y_0^2/N - 2\theta N[\psi] - \sigma^2]$$

where

$$\psi = \sigma^2 \int_0^1 J_c(r)^2 dr + 2\sigma y_0 N^{-1/2} \int_0^1 \exp(cr) J_c(r) dr + y_0^2 N^{-2}(\exp(2\theta N) - 1)/2\theta$$

Using the fact that:

$$T_n (\alpha_n - \alpha_n) = N^{-1}\sum_{nt} y_{nt-1} u_{nt} \left( N^{-1}T_n^{-1}\sum y_{nt-1}^2 \right)^{-1}$$

$$T_n (\alpha_n - \alpha_n) \rightarrow (1/2)[(\sigma J_c(1) + N^{-1/2}\exp(\theta N)y_0)^2 - y_0^2/N - 2\theta N[\psi] - \sigma^2]/\psi$$

$$= \left[ \int_0^1 J_c(r) dw_r + (y_0/\sigma N^{1/2})(\exp(c) J_c(1) - 2c \int_0^1 \exp(cr) J_c(r) dr) / \right]$$

$$\left[ \int_0^1 J_c(r)^2 dr + (2y_0/\sigma N^{1/2}) \int_0^1 \exp(cr) J_c(r) dr \right]$$

$$+ (y_0^2 / N\sigma^2)(\exp(2c) - 1)/2c.$$
Note that $J_c(1) \sim N(0, (\exp(2c) - 1)/2c)$ and $\int_0^1 \exp(rc) J_c(r) dr \sim N(0, (\exp(4c) - 1 - 4c \exp(2c))/8c^3)$. Therefore $[\exp(c)J_c(1) - 2c \int_0^1 \exp(rc) J_c(r) dr]$ is Gaussian with mean 0 and variance given by $\exp(2c) \text{Var}[J_c(1)] + 4c^2 \text{Var}[\int_0^1 \exp(rc) J_c(r) dr] - 2E[J_c(1) \int_0^1 \exp(rc) J_c(r) dr]$.

Now, $E[J_c(1) \int_0^1 \exp(rc) J_c(r) dr]$

$= \int_0^1 \exp(rc) E[J_c(1) J_c(r)] \, dr$

$= \int_0^1 \exp(rc) [\exp((r+1)c) - \exp((1-r)c)]/2c \, dr$

$= \exp(c) [\exp(2c) - 1 - 2c]/(2c)^2$.

Therefore $[\exp(c)J_c(1) - 2c \int_0^1 \exp(rc) J_c(r) dr]$ has variance given by $(\exp(2c) - 1)/2c$.

**Proof of Lemma 1**

To prove part (v), we first note that $w(s)$ is Gaussian with mean zero and variance $s$, then $J_c(r)$ and $\int_0^1 \exp(rc) J_c(r) dr$ are also Gaussian with mean zero for any $c$. Now $J_c(r) \sim N(0, (\exp(2cr) - 1)/2c)$ and using the fact that $E(w(s)w(t)) = \min(t,s)$, simple calculations yield

$E[J_c(r) J_c(k)] = [\exp((r+k)c) - \exp((M-m)c)]/2c$

where $M = \max(r,k)$ and $m = \min(r,k)$. This result can be used to show that

$\int_0^1 \exp(rc) J_c(r) dr \sim N(0, v)$

where $v = (\exp(4c) - 1 - 4c \exp(2c))/8c^3$.

Part (v) follows by noting that $(2c)^3 \exp(-4c)v \to 1$ as $c \to +\infty$. 
To prove part (vi), note that \( \int_0^1 \exp(cr) dw_r = \exp(c) \int_0^1 \exp(-c(1 - r)) dw_r - \exp(c) J^{-c}(1) \).

Now, \( J^{-c}(1) \sim N(0,(\exp(-2c) - 1)/(-2c)) \), hence \( \int_0^1 \exp(cr) dw_r \sim N(0,(\exp(2c) - 1)/2c) \).

The result follows by noting that \( (2c)^{1/2} \exp(-c) \int_0^1 \exp(cr) dw_r = (2c)^{1/2} \exp(-c) \cdot N(0,(\exp(2c) - 1)/2c) = N(0,(1 - \exp(-2c))) \rightarrow N(0,1) \) as \( c \rightarrow +\infty \).
BIBLIOGRAPHY


### Table 1

**Selected Percentage Points of the Distribution of N(\(\hat{\theta} - \theta\))**

<table>
<thead>
<tr>
<th>c</th>
<th>1 %</th>
<th>2.5 %</th>
<th>5 %</th>
<th>10 %</th>
<th>90 %</th>
<th>95 %</th>
<th>97.5 %</th>
<th>99 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>c = -5.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>= 0.5</td>
<td>-13.8955</td>
<td>-10.9832</td>
<td>-8.7550</td>
<td>-6.4899</td>
<td>2.2695</td>
<td>2.8257</td>
<td>3.2413</td>
<td>3.6788</td>
</tr>
<tr>
<td>= 1.0</td>
<td>-9.5405</td>
<td>-7.6095</td>
<td>-6.1140</td>
<td>-4.5703</td>
<td>1.9279</td>
<td>2.4017</td>
<td>2.7578</td>
<td>3.1215</td>
</tr>
<tr>
<td>= 2.0</td>
<td>-4.8739</td>
<td>-3.9531</td>
<td>-3.2206</td>
<td>-2.4413</td>
<td>1.3836</td>
<td>1.7357</td>
<td>2.0123</td>
<td>2.3035</td>
</tr>
<tr>
<td>c = -2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = 0.0</td>
<td>-15.0342</td>
<td>-11.6602</td>
<td>-9.1377</td>
<td>-6.6406</td>
<td>1.4622</td>
<td>1.926*</td>
<td>2.3095</td>
<td>2.7849</td>
</tr>
<tr>
<td>= 0.5</td>
<td>-12.2749</td>
<td>-9.5498</td>
<td>-7.5038</td>
<td>-5.4726</td>
<td>1.3047</td>
<td>1.6839</td>
<td>2.005*</td>
<td>2.3921</td>
</tr>
<tr>
<td>= 1.0</td>
<td>-8.1023</td>
<td>-6.3512</td>
<td>-5.0258</td>
<td>-3.6952</td>
<td>1.0808</td>
<td>1.3688</td>
<td>1.5996</td>
<td>1.8594</td>
</tr>
<tr>
<td>= 2.0</td>
<td>-3.8162</td>
<td>-3.0438</td>
<td>-2.4456</td>
<td>-1.8281</td>
<td>0.7943</td>
<td>0.9940</td>
<td>1.1599</td>
<td>1.3310</td>
</tr>
<tr>
<td>c = -1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = 0.0</td>
<td>-14.3901</td>
<td>-11.0813</td>
<td>-8.6182</td>
<td>-6.2072</td>
<td>1.1752</td>
<td>1.5887</td>
<td>1.9502</td>
<td>2.4005</td>
</tr>
<tr>
<td>= 0.5</td>
<td>-11.6481</td>
<td>-8.9854</td>
<td>-7.0027</td>
<td>-5.0566</td>
<td>1.007*</td>
<td>1.3498</td>
<td>1.6426</td>
<td>2.0031</td>
</tr>
<tr>
<td>= 1.0</td>
<td>-7.5197</td>
<td>-5.8323</td>
<td>-4.5689</td>
<td>-3.3203</td>
<td>0.8076</td>
<td>1.056*</td>
<td>1.2535</td>
<td>1.4819</td>
</tr>
<tr>
<td>= 2.0</td>
<td>-3.3477</td>
<td>-2.6348</td>
<td>-2.0925</td>
<td>-1.5445</td>
<td>0.5908</td>
<td>0.7512</td>
<td>0.8784</td>
<td>1.012*</td>
</tr>
<tr>
<td>c = 0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = 0.0</td>
<td>-13.6919</td>
<td>-10.4399</td>
<td>-8.0383</td>
<td>-5.7133</td>
<td>0.9280</td>
<td>1.2854</td>
<td>1.6122</td>
<td>2.0325</td>
</tr>
<tr>
<td>= 0.5</td>
<td>-10.9570</td>
<td>-8.3535</td>
<td>-6.4315</td>
<td>-4.5709</td>
<td>0.7643</td>
<td>1.0481</td>
<td>1.3042</td>
<td>1.6338</td>
</tr>
<tr>
<td>= 1.0</td>
<td>-6.8480</td>
<td>-5.2202</td>
<td>-4.0196</td>
<td>-2.8568</td>
<td>0.5876</td>
<td>0.7780</td>
<td>0.9384</td>
<td>1.1302</td>
</tr>
<tr>
<td>= 2.0</td>
<td>-2.7392</td>
<td>-2.0886</td>
<td>-1.6093</td>
<td>-1.1454</td>
<td>0.4116</td>
<td>0.5271</td>
<td>0.6211</td>
<td>0.7260</td>
</tr>
<tr>
<td>c = 1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = 0.0</td>
<td>-12.9229</td>
<td>-9.7183</td>
<td>-7.3710</td>
<td>-5.1243</td>
<td>0.7101</td>
<td>1.0079</td>
<td>1.2966</td>
<td>1.6832</td>
</tr>
<tr>
<td>= 0.5</td>
<td>-10.1753</td>
<td>-7.6177</td>
<td>-5.7469</td>
<td>-3.9627</td>
<td>0.5544</td>
<td>0.7771</td>
<td>0.9932</td>
<td>1.2871</td>
</tr>
<tr>
<td>= 1.0</td>
<td>-6.0273</td>
<td>-4.4386</td>
<td>-3.2793</td>
<td>-2.1696</td>
<td>0.4017</td>
<td>0.5366</td>
<td>0.6571</td>
<td>0.8093</td>
</tr>
<tr>
<td>= 2.0</td>
<td>-1.7104</td>
<td>-1.1842</td>
<td>-0.8640</td>
<td>-0.5948</td>
<td>0.2610</td>
<td>0.3360</td>
<td>0.3994</td>
<td>0.4725</td>
</tr>
<tr>
<td>c = 2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = 0.0</td>
<td>-12.0557</td>
<td>-8.8762</td>
<td>-6.5587</td>
<td>-4.3431</td>
<td>0.5191</td>
<td>0.7585</td>
<td>1.0067</td>
<td>1.3550</td>
</tr>
<tr>
<td>= 0.5</td>
<td>-9.2554</td>
<td>-6.7080</td>
<td>-4.8444</td>
<td>-3.0206</td>
<td>0.3756</td>
<td>0.5401</td>
<td>0.7143</td>
<td>0.9680</td>
</tr>
<tr>
<td>= 1.0</td>
<td>-4.8763</td>
<td>-3.1253</td>
<td>-1.7404</td>
<td>-0.8782</td>
<td>0.2493</td>
<td>0.3357</td>
<td>0.4182</td>
<td>0.5294</td>
</tr>
<tr>
<td>= 2.0</td>
<td>-0.5897</td>
<td>-0.4307</td>
<td>-0.3294</td>
<td>-0.2374</td>
<td>0.1466</td>
<td>0.1896</td>
<td>0.2272</td>
<td>0.2712</td>
</tr>
<tr>
<td>c = 5.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ = 0.0</td>
<td>-7.8066</td>
<td>-3.0640</td>
<td>-1.0714</td>
<td>-0.3734</td>
<td>0.1316</td>
<td>0.2212</td>
<td>0.3410</td>
<td>0.5473</td>
</tr>
<tr>
<td>= 0.5</td>
<td>-2.2400</td>
<td>-0.5435</td>
<td>-0.2103</td>
<td>-0.0969</td>
<td>0.0611</td>
<td>0.0974</td>
<td>0.1505</td>
<td>0.2571</td>
</tr>
<tr>
<td>= 1.0</td>
<td>-0.0826</td>
<td>-0.0579</td>
<td>-0.0438</td>
<td>-0.0315</td>
<td>0.0283</td>
<td>0.0385</td>
<td>0.0491</td>
<td>0.0650</td>
</tr>
<tr>
<td>= 2.0</td>
<td>-0.0275</td>
<td>-0.0225</td>
<td>-0.0186</td>
<td>-0.0142</td>
<td>0.0137</td>
<td>0.0177</td>
<td>0.0214</td>
<td>0.0259</td>
</tr>
</tbody>
</table>

**Note:** The integral (4.2) which evaluates \(P[N(\hat{\theta} - \theta) \leq z]\) is undefined for \(z = -c\) and the numerical integration is accordingly ill-behaved for values of \(z\) around \(-c\). The * entry indicates where such a situation occurs. While all digits reported are accurate for the other entries, in these cases the accuracy is not as precise.
Table 2
Selected Percentage Points of the Distribution of $T(\hat{\alpha} - \alpha)$

$y_t = \alpha y_{t-1} + \epsilon_t$, $y_0/\sigma = T^{1/2}$, $\alpha = \exp(c/T)$

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>2.5%</th>
<th>5%</th>
<th>10%</th>
<th>90%</th>
<th>95%</th>
<th>97.5%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c=-5.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma=0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma=1.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T=10$</td>
<td>-5.642</td>
<td>-4.756</td>
<td>-3.941</td>
<td>-3.050</td>
<td>1.693</td>
<td>2.079</td>
<td>2.395</td>
<td>2.713</td>
</tr>
<tr>
<td>$T=50$</td>
<td>-8.488</td>
<td>-6.930</td>
<td>-5.616</td>
<td>-4.228</td>
<td>1.862</td>
<td>2.335</td>
<td>2.677</td>
<td>3.038</td>
</tr>
<tr>
<td>$\gamma=2.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c=-2.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma=0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T=10$</td>
<td>-7.854</td>
<td>-6.626</td>
<td>-5.414</td>
<td>-4.196</td>
<td>1.336</td>
<td>1.748</td>
<td>2.123</td>
<td>2.564</td>
</tr>
<tr>
<td>$T=50$</td>
<td>-10.951</td>
<td>-8.626</td>
<td>-6.962</td>
<td>-5.217</td>
<td>1.321</td>
<td>1.708</td>
<td>2.053</td>
<td>2.472</td>
</tr>
<tr>
<td>$\gamma=1.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T=10$</td>
<td>-5.583</td>
<td>-4.572</td>
<td>-3.712</td>
<td>-2.882</td>
<td>1.049</td>
<td>1.337</td>
<td>1.568</td>
<td>1.847</td>
</tr>
<tr>
<td>$T=50$</td>
<td>-7.518</td>
<td>-6.045</td>
<td>-4.751</td>
<td>-3.580</td>
<td>1.075</td>
<td>1.363</td>
<td>1.592</td>
<td>1.872</td>
</tr>
<tr>
<td>$T=100$</td>
<td>-7.732</td>
<td>-6.225</td>
<td>-4.935</td>
<td>-3.678</td>
<td>1.068</td>
<td>1.349</td>
<td>1.573</td>
<td>1.836</td>
</tr>
<tr>
<td>$\gamma=2.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T=10$</td>
<td>-3.015</td>
<td>-2.449</td>
<td>-1.967</td>
<td>-1.507</td>
<td>0.756</td>
<td>0.939</td>
<td>1.097</td>
<td>1.262</td>
</tr>
<tr>
<td>$T=25$</td>
<td>-3.402</td>
<td>-2.698</td>
<td>-2.211</td>
<td>-1.672</td>
<td>0.773</td>
<td>0.978</td>
<td>1.138</td>
<td>1.303</td>
</tr>
<tr>
<td>$T=50$</td>
<td>-3.620</td>
<td>-2.913</td>
<td>-2.382</td>
<td>-1.789</td>
<td>0.785</td>
<td>0.983</td>
<td>1.152</td>
<td>1.326</td>
</tr>
<tr>
<td>$T=100$</td>
<td>-3.779</td>
<td>-2.991</td>
<td>-2.437</td>
<td>-1.834</td>
<td>0.785</td>
<td>0.986</td>
<td>1.138</td>
<td>1.316</td>
</tr>
<tr>
<td>$c=0.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma=0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T=10$</td>
<td>-7.892</td>
<td>-6.392</td>
<td>-5.230</td>
<td>-3.897</td>
<td>0.864</td>
<td>1.188</td>
<td>1.508</td>
<td>1.915</td>
</tr>
<tr>
<td>$T=25$</td>
<td>-9.589</td>
<td>-7.578</td>
<td>-5.897</td>
<td>-4.347</td>
<td>0.791</td>
<td>1.098</td>
<td>1.372</td>
<td>1.765</td>
</tr>
<tr>
<td>$T=50$</td>
<td>-9.954</td>
<td>-7.873</td>
<td>-6.128</td>
<td>-4.454</td>
<td>0.797</td>
<td>1.092</td>
<td>1.348</td>
<td>1.688</td>
</tr>
<tr>
<td>$T=100$</td>
<td>-10.457</td>
<td>-8.159</td>
<td>-6.417</td>
<td>-4.587</td>
<td>0.772</td>
<td>1.052</td>
<td>1.313</td>
<td>1.608</td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th></th>
<th>1 %</th>
<th>2.5 %</th>
<th>5 %</th>
<th>10 %</th>
<th>90 %</th>
<th>95 %</th>
<th>97.5 %</th>
<th>99 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma = 1.0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T=10</td>
<td>-5.253</td>
<td>-4.209</td>
<td>-3.364</td>
<td>-2.469</td>
<td>0.634</td>
<td>0.843</td>
<td>1.001</td>
<td>1.221</td>
</tr>
<tr>
<td>T=25</td>
<td>-5.892</td>
<td>-4.638</td>
<td>-3.677</td>
<td>-2.653</td>
<td>0.597</td>
<td>0.793</td>
<td>0.961</td>
<td>1.167</td>
</tr>
<tr>
<td>T=50</td>
<td>-6.557</td>
<td>-5.128</td>
<td>-3.995</td>
<td>-2.788</td>
<td>0.598</td>
<td>0.798</td>
<td>0.955</td>
<td>1.156</td>
</tr>
<tr>
<td>T=100</td>
<td>-6.629</td>
<td>-5.178</td>
<td>-4.063</td>
<td>-2.891</td>
<td>0.588</td>
<td>0.777</td>
<td>0.935</td>
<td>1.125</td>
</tr>
<tr>
<td>( \gamma = 2.0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T=10</td>
<td>-2.433</td>
<td>-1.873</td>
<td>-1.462</td>
<td>-1.057</td>
<td>0.430</td>
<td>0.549</td>
<td>0.648</td>
<td>0.758</td>
</tr>
<tr>
<td>T=25</td>
<td>-2.589</td>
<td>-1.996</td>
<td>-1.569</td>
<td>-1.102</td>
<td>0.417</td>
<td>0.531</td>
<td>0.630</td>
<td>0.737</td>
</tr>
<tr>
<td>T=50</td>
<td>-2.715</td>
<td>-2.055</td>
<td>-1.582</td>
<td>-1.131</td>
<td>0.416</td>
<td>0.534</td>
<td>0.629</td>
<td>0.738</td>
</tr>
<tr>
<td>T=100</td>
<td>-2.721</td>
<td>-2.107</td>
<td>-1.613</td>
<td>-1.151</td>
<td>0.411</td>
<td>0.527</td>
<td>0.617</td>
<td>0.718</td>
</tr>
<tr>
<td>( c = 2.0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma = 0.5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T=10</td>
<td>-7.598</td>
<td>-5.796</td>
<td>-4.461</td>
<td>-2.832</td>
<td>0.459</td>
<td>0.655</td>
<td>0.867</td>
<td>1.163</td>
</tr>
<tr>
<td>T=25</td>
<td>-8.423</td>
<td>-6.302</td>
<td>-4.686</td>
<td>-2.962</td>
<td>0.406</td>
<td>0.583</td>
<td>0.770</td>
<td>1.054</td>
</tr>
<tr>
<td>T=50</td>
<td>-8.953</td>
<td>-6.493</td>
<td>-4.751</td>
<td>-3.016</td>
<td>0.401</td>
<td>0.565</td>
<td>0.745</td>
<td>1.008</td>
</tr>
<tr>
<td>T=100</td>
<td>-9.207</td>
<td>-6.595</td>
<td>-4.801</td>
<td>-3.000</td>
<td>0.377</td>
<td>0.550</td>
<td>0.718</td>
<td>0.957</td>
</tr>
<tr>
<td>( \gamma = 1.0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T=10</td>
<td>-4.359</td>
<td>-2.564</td>
<td>-1.421</td>
<td>-0.788</td>
<td>0.290</td>
<td>0.386</td>
<td>0.481</td>
<td>0.611</td>
</tr>
<tr>
<td>T=25</td>
<td>-4.306</td>
<td>-2.870</td>
<td>-1.617</td>
<td>-0.833</td>
<td>0.263</td>
<td>0.356</td>
<td>0.445</td>
<td>0.558</td>
</tr>
<tr>
<td>T=50</td>
<td>-4.703</td>
<td>-3.036</td>
<td>-1.676</td>
<td>-0.860</td>
<td>0.262</td>
<td>0.348</td>
<td>0.432</td>
<td>0.543</td>
</tr>
<tr>
<td>T=100</td>
<td>-5.037</td>
<td>-3.163</td>
<td>-1.735</td>
<td>-0.887</td>
<td>0.252</td>
<td>0.340</td>
<td>0.418</td>
<td>0.526</td>
</tr>
<tr>
<td>( \gamma = 2.0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T=10</td>
<td>-0.562</td>
<td>-0.437</td>
<td>-0.338</td>
<td>-0.247</td>
<td>0.165</td>
<td>0.214</td>
<td>0.257</td>
<td>0.311</td>
</tr>
<tr>
<td>T=25</td>
<td>-0.585</td>
<td>-0.432</td>
<td>-0.334</td>
<td>-0.241</td>
<td>0.154</td>
<td>0.200</td>
<td>0.237</td>
<td>0.288</td>
</tr>
<tr>
<td>T=50</td>
<td>-0.596</td>
<td>-0.445</td>
<td>-0.335</td>
<td>-0.242</td>
<td>0.151</td>
<td>0.196</td>
<td>0.235</td>
<td>0.279</td>
</tr>
<tr>
<td>T=100</td>
<td>-0.597</td>
<td>-0.424</td>
<td>-0.325</td>
<td>-0.238</td>
<td>0.149</td>
<td>0.192</td>
<td>0.226</td>
<td>0.272</td>
</tr>
</tbody>
</table>
Table 3
Expectations of $g(T)(\hat{\alpha} - \alpha)$

Based on the asymptotic distribution (exact values in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 0.90$</th>
<th>$\alpha = 0.99$</th>
<th>$\alpha = 1.00$</th>
<th>$\alpha = 1.01$</th>
<th>$\alpha = 1.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>$-0.8933$</td>
<td>$-2.5718$</td>
<td>$-1.2597$</td>
<td>$-4.4482$</td>
<td>$-2.0080$</td>
</tr>
<tr>
<td>50</td>
<td>$-0.6416$</td>
<td>$-1.8457$</td>
<td>$-1.2597$</td>
<td>$-2.7778$</td>
<td>$-2.4096$</td>
</tr>
<tr>
<td>100</td>
<td>$-0.4570$</td>
<td>$-1.3342$</td>
<td>$-1.2597$</td>
<td>$-2.1295$</td>
<td>$-3.6704$</td>
</tr>
<tr>
<td>200</td>
<td>$-0.3240$</td>
<td>$-0.9675$</td>
<td>$-1.2597$</td>
<td>$-2.2841$</td>
<td>$-5.2877$</td>
</tr>
<tr>
<td>400</td>
<td>$-0.2293$</td>
<td>$-0.6975$</td>
<td>$-1.2597$</td>
<td>$-3.3706$</td>
<td>$-5.3839$</td>
</tr>
</tbody>
</table>

$y(0)/\sigma = 0.0$

<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 0.90$</th>
<th>$\alpha = 0.99$</th>
<th>$\alpha = 1.00$</th>
<th>$\alpha = 1.01$</th>
<th>$\alpha = 1.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>$-0.5509$</td>
<td>$-1.5492$</td>
<td>$-0.7480$</td>
<td>$-2.5895$</td>
<td>$-0.9968$</td>
</tr>
<tr>
<td>50</td>
<td>$-0.4878$</td>
<td>$-1.4021$</td>
<td>$-0.9467$</td>
<td>$-2.0429$</td>
<td>$-1.3231$</td>
</tr>
<tr>
<td>100</td>
<td>$-0.3943$</td>
<td>$-1.1555$</td>
<td>$-1.0835$</td>
<td>$-1.7800$</td>
<td>$-1.8444$</td>
</tr>
<tr>
<td>200</td>
<td>$-0.3001$</td>
<td>$-0.8981$</td>
<td>$-1.1657$</td>
<td>$-2.0108$</td>
<td>$-2.3509$</td>
</tr>
<tr>
<td>400</td>
<td>$-0.2205$</td>
<td>$-0.6712$</td>
<td>$-1.2110$</td>
<td>$-2.9477$</td>
<td>$-2.9467$</td>
</tr>
</tbody>
</table>

$y(0)/\sigma = 4.0$

<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 0.90$</th>
<th>$\alpha = 0.99$</th>
<th>$\alpha = 1.00$</th>
<th>$\alpha = 1.01$</th>
<th>$\alpha = 1.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>$-0.0805$</td>
<td>$-0.1677$</td>
<td>$-0.0732$</td>
<td>$-0.2244$</td>
<td>$-0.0455$</td>
</tr>
<tr>
<td>50</td>
<td>$-0.1060$</td>
<td>$-0.2605$</td>
<td>$-0.1509$</td>
<td>$-0.2556$</td>
<td>$-0.0308$</td>
</tr>
<tr>
<td>100</td>
<td>$-0.1288$</td>
<td>$-0.3678$</td>
<td>$-0.2968$</td>
<td>$-0.3150$</td>
<td>$-0.0060$</td>
</tr>
<tr>
<td>200</td>
<td>$-0.1423$</td>
<td>$-0.4311$</td>
<td>$-0.5120$</td>
<td>$-0.3955$</td>
<td>$-0.0001$</td>
</tr>
<tr>
<td>400</td>
<td>$-0.1399$</td>
<td>$-0.4287$</td>
<td>$-0.7480$</td>
<td>$-0.4107$</td>
<td>$-0.0045$</td>
</tr>
</tbody>
</table>

$y(0)/\sigma = 16.0$
Table 4
Standard Deviations of \( g(T)(\hat{\alpha} - \alpha) \)

Based on the asymptotic distribution (exact values in parentheses)

\[
\begin{array}{cccccccc}
\text{y(0)/}\sigma = 0.0 & & & & & & & \\
T & \alpha = 0.90 & \alpha = 0.99 & \alpha = 1.00 & \alpha = 1.01 & \alpha = 1.05 \\
25 & 1.7438 & (1.5135) & 4.5919 & (4.1480) & 2.2486 & (2.0401) & 7.9463 & (7.2601) & 3.7365 & (3.5359) \\
50 & 1.4274 & (1.3004) & 3.3049 & (3.1293) & 2.2486 & (2.1356) & 4.9795 & (4.7771) & 5.4288 & (5.4447) \\
100 & 1.2482 & (1.1657) & 2.4199 & (2.3552) & 2.2486 & (2.1898) & 3.8966 & (3.8620) & 18.4035 & (18.8851) \\
200 & 1.1523 & (1.0876) & 1.8307 & (1.8113) & 2.2486 & (2.2186) & 4.6999 & (4.7802) & 228.4101 & (234.2883) \\
400 & 1.1030 & (1.0452) & 1.4611 & (1.4594) & 2.2486 & (2.2334) & 12.0295 & (12.3768) & - & - \\
\text{y(0)/}\sigma = 4.0 & & & & & & & \\
T & \alpha = 0.90 & \alpha = 0.99 & \alpha = 1.00 & \alpha = 1.01 & \alpha = 1.05 \\
25 & 1.2425 & (1.0374) & 2.8652 & (2.6100) & 1.3852 & (1.2683) & 4.8429 & (4.4536) & 2.1914 & (2.0004) \\
50 & 1.2305 & (1.0706) & 2.5471 & (2.4084) & 1.7088 & (1.6240) & 3.7469 & (3.5833) & 3.8048 & (3.5898) \\
100 & 1.1813 & (1.0611) & 2.1203 & (2.0554) & 1.9401 & (1.8888) & 3.3340 & (3.2820) & 13.2287 & (12.7345) \\
200 & 1.1321 & (1.0402) & 1.7228 & (1.6956) & 2.0825 & (2.0543) & 4.3122 & (4.3375) & 161.2751 & (5.6014) \\
400 & 1.0974 & (1.0231) & 1.4256 & (1.4158) & 2.1622 & (2.1475) & 11.2652 & (11.4476) & - & - \\
\text{y(0)/}\sigma = 16 & & & & & & & \\
T & \alpha = 0.90 & \alpha = 0.99 & \alpha = 1.00 & \alpha = 1.01 & \alpha = 1.05 \\
25 & 0.4425 & (0.3173) & 0.6374 & (0.5686) & 0.2585 & (0.2552) & 0.8714 & (0.8164) & 0.2396 & (0.2274) \\
50 & 0.5862 & (0.4264) & 0.7560 & (0.6835) & 0.4155 & (0.4098) & 0.8121 & (0.7567) & 0.2311 & (0.2179) \\
100 & 0.7441 & (0.5554) & 0.8961 & (0.8166) & 0.6714 & (0.6631) & 0.9412 & (0.8692) & 0.2299 & (0.2119) \\
200 & 0.8861 & (0.6876) & 1.0199 & (0.9346) & 1.0141 & (1.0046) & 1.5026 & (1.3650) & 0.6096 & (0.2090) \\
400 & 0.9844 & (0.8021) & 1.0937 & (1.0122) & 1.3852 & (1.3769) & 4.0157 & (3.6065) & - & - \\
\end{array}
\]
<table>
<thead>
<tr>
<th>( T )</th>
<th>( \alpha = 0.90 )</th>
<th>( \alpha = 0.95 )</th>
<th>( \alpha = 0.99 )</th>
<th>( \alpha = 1.00 )</th>
<th>( \alpha = 1.01 )</th>
<th>( \alpha = 1.025 )</th>
<th>( \alpha = 1.05 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.0789 (0.0779)</td>
<td>0.0494 (0.0506)</td>
<td>0.0456 (0.0472)</td>
<td>0.0500</td>
<td>0.0584 (0.0570)</td>
<td>0.0832 (0.0759)</td>
<td>0.1946 (0.1645)</td>
</tr>
<tr>
<td>50</td>
<td>0.1913 (0.1848)</td>
<td>0.0769 (0.0770)</td>
<td>0.0440 (0.0461)</td>
<td>0.0500</td>
<td>0.0730 (0.0706)</td>
<td>0.1992 (0.1843)</td>
<td>0.7073 (0.6904)</td>
</tr>
<tr>
<td>100</td>
<td>0.5893 (0.5615)</td>
<td>0.1835 (0.1818)</td>
<td>0.0461 (0.0462)</td>
<td>0.0500</td>
<td>0.1379 (0.1322)</td>
<td>0.7164 (0.7046)</td>
<td>0.9755 (1.0000)</td>
</tr>
<tr>
<td>200</td>
<td>0.9925 (0.9870)</td>
<td>0.5665 (0.5487)</td>
<td>0.0627 (0.0622)</td>
<td>0.0500</td>
<td>0.5330 (0.5179)</td>
<td>1.0000 (1.0000)</td>
<td>1.0000 (1.0000)</td>
</tr>
<tr>
<td>400</td>
<td>1.0000 (1.0000)</td>
<td>0.9897 (0.9862)</td>
<td>0.1290 (0.1284)</td>
<td>0.0500</td>
<td>0.9402</td>
<td>1.0000 (1.0000)</td>
<td>1.0000 (1.0000)</td>
</tr>
</tbody>
</table>

| \( y(0)/\sigma = 2.0 \) |
|---|---|---|---|---|---|---|---|
| \( T \) | \( \alpha = 0.90 \) | \( \alpha = 0.95 \) | \( \alpha = 0.99 \) | \( \alpha = 1.00 \) | \( \alpha = 1.01 \) | \( \alpha = 1.025 \) | \( \alpha = 1.05 \) |
| 25 | 0.0920 (0.0918) | 0.0524 (0.0536) | 0.0451 (0.0467) | 0.0500 | 0.0604 (0.0614) | 0.0949 (0.0887) | 0.2646 (0.2406) |
| 50 | 0.2160 (0.2119) | 0.0829 (0.0825) | 0.0439 (0.0463) | 0.0500 | 0.0763 (0.0745) | 0.2331 (0.2198) | 0.7612 (0.7513) |
| 100 | 0.6232 (0.5960) | 0.1951 (0.1915) | 0.0465 (0.0469) | 0.0500 | 0.1465 (0.1412) | 0.7444 (0.7341) | 0.9798 (1.0000) |
| 200 | 0.9942 (0.9900) | 0.5834 (0.5676) | 0.0637 (0.0633) | 0.0500 | 0.5510 (0.5369) | 0.9790 (1.0000) | 1.0000 (1.0000) |
| 400 | 1.0000 (1.0000) | 0.9908 (0.9877) | 0.1307 (0.1304) | 0.0500 | 0.9425 | 1.0000 (1.0000) | 1.0000 (1.0000) |

| \( y(0)/\sigma = 8.0 \) |
|---|---|---|---|---|---|---|---|
| \( T \) | \( \alpha = 0.90 \) | \( \alpha = 0.95 \) | \( \alpha = 0.99 \) | \( \alpha = 1.00 \) | \( \alpha = 1.01 \) | \( \alpha = 1.025 \) | \( \alpha = 1.05 \) |
| 25 | 0.4933 (0.4853) | 0.1541 (0.1565) | 0.0447 (0.0448) | 0.0500 | 0.0965 (0.0918) | 0.3184 (0.3083) | 0.8395 (0.8421) |
| 50 | 0.7035 (0.6816) | 0.2257 (0.2242) | 0.0465 (0.0462) | 0.0500 | 0.1532 (0.1475) | 0.6648 (0.6616) | 0.9835 (0.9852) |
| 100 | 0.9494 (0.9350) | 0.4135 (0.4069) | 0.0548 (0.0554) | 0.0500 | 0.3271 (0.3207) | 0.9375 (0.9360) | 0.9989 (1.0000) |
| 200 | 1.0000 (0.9998) | 0.8037 (0.7925) | 0.0794 (0.0803) | 0.0500 | 0.7457 (0.7430) | 0.9951 (1.0000) | 1.0000 (1.0000) |
| 400 | 1.0000 (1.0000) | 0.9986 (0.9980) | 0.1590 (0.1610) | 0.0500 | 0.9681 | 1.0000 (1.0000) | 1.0000 (1.0000) |

Table 5

Power function for testing the hypothesis \( \alpha = 1.00 \) against alternatives near 1.00

Based on the asymptotic distribution (exact values in parentheses)
Significance level 0.05
CUMULATIVE DISTRIBUTION FUNCTIONS

Figure 1

N(0 - δ)

γ = 0.5
γ = 0.0
γ = 1.0
γ = 2.0

c = 5.0

-15 -13 -11 -9 -7 -5 -3 -1 1 3 5

0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
PROBABILITY DENSITY FUNCTIONS

$N(\hat{\theta} - \theta)$

$\gamma = 2.0$

$\gamma = 1.0$

$\gamma = 0.5$

$\gamma = 0.0$

$C = -5.0$
Figure 5

PROBABILITY DENSITY FUNCTIONS

N(\theta - \theta)

\gamma = 2.0
\gamma = 1.0
\gamma = 0.5
\gamma = 0.0

C = 0.0
Figure 6

PROBABILITY DENSITY FUNCTIONS

$N(\theta - \theta)$

$\gamma = 2.0$

$\gamma = 1.0$

$\gamma = 0.5$

$\gamma = 0.0$
Figure 7

BIAS FUNCTIONS

$E[N(\hat{\theta} - \theta)]$

$\gamma = 5.0$

$\gamma = 2.0$

$\gamma = 1.0$

$\gamma = 0.5$

$\gamma = 0.0$

Non-centrality parameter $C$
Figure 8

MEAN–SQUARED ERROR

$E[N(\hat{\theta} - \theta)]^2$

Non-centrality parameter $C$
Figure 9

Power functions

5% one-tailed test of \( \theta = 0 \)

- \( \gamma = 2.0 \)
- \( \gamma = 1.0 \)
- \( \gamma = 0.5 \)
- \( \gamma = 0.1 \)

Non-centrality parameter \( \lambda \)