The Annals of Statistics
1978, Vol. 6, No. 6, 1337–1344

TESTING THE MEAN OF A NORMAL POPULATION UNDER DEPENDENCE[1]

BY W. ALBERS[2]

Technological University Twente, Enschede

Modifications of the t-test are considered which are robust under certain violations of the independence assumption. The additional number of observations these modified tests require in order to obtain under independence the same power as the t-test, is obtained asymptotically.

1. Introduction. In testing the mean of a normal population it is well known that the t-test may be invalidated if the observations are dependent. In fact, the sensitivity of statistical procedures to violations of the independence assumption has been studied by several authors. Gastwirth, Rubin and Wolff (1967) showed that the sign test is no longer distribution-free even when the observations are from two stationary processes with the same spectrum. Gastwirth and Rubin (1971) studied the effect of serial correlation of the observations on the level of the one-sample t-test, sign test and Wilcoxon test. Serfling (1968) considered the two-sample Wilcoxon test under strongly mixing processes. The problem has also received attention in the engineering literature, see, e.g., Modestino (1969). Finally, the effect of dependence on robust estimators has been studied by Gastwirth and Rubin (1975).

In the present paper a modification of the t-test is considered which has robustness of validity under m-dependence. As concerns the price for this robustness, it is shown that in case of independence the modified test asymptotically requires mu_α^2 more observations than the ordinary t-test. Here α is the size of the test and u_α is the upper α-point of the standard normal distribution function.

Finally it is demonstrated that similar results hold for autoregressive processes. In particular, it is shown that the required additional number of observations under independence again tends to mu_α^2, where, in this case, m is the order of the autoregressive equation.

2. A modified t-test for dependent observations. Let $X_1, \cdots, X_N$ be normally distributed random variables (rv's) with expectation $EX_i = \mu$ and variance $\sigma^2(X_i) = \sigma^2 > 0$, $i = 1, \cdots, N$. If the X_i are also independent, the usual test for $H_0 : \mu = 0$ against $H_1 : \mu > 0$ is Student's t-test, which we shall denote by ψ_t. It

Received March 1976; revised June 1977.

[1]This research was supported by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

[2]This paper was written while the author was visiting the Statistics Department of the University of California at Berkeley.

AMS 1970 *subject classifications*. 62F05, 62F20.

Key words and phrases. Normal population, mean, t-test, m-dependence, serial correlation coefficients, asymptotic expansions, deficiency, autoregressive processes.

rejects the hypothesis for large values of

$$t = N^{\frac{1}{2}}\bar{X}/S, \tag{2.1}$$

where $\bar{X} = N^{-1}\Sigma_{i=1}^{N}X_i$ and $S^2 = (N-1)^{-1}\Sigma_{i=1}^{N}(X_i - \bar{X})^2$.

If the observations are not independent, however, it is well known that ψ_t may be invalidated, if the same critical value as under independence is used. In this section we consider the following special case of m-dependence: let $(X_1, \cdots, X_N)$ be jointly normally distributed with $EX_i = \mu$ and $\sigma^2(X_i) = \sigma^2 > 0$, $i = 1, \cdots, N$. Let R_N be the $N \times N$ matrix of the correlation coefficients $\rho(X_i, X_j)$ of X_i and X_j, $i, j, = 1, \cdots, N$. (Hence the covariance matrix of $(X_1, \cdots, X_N)$ is $\sigma^2 R_N$.) Suppose that the elements of R_N satisfy, for some fixed positive integer m,

$$\begin{aligned} \rho(X_i, X_j) &= \rho_{|i-j|}, \qquad 1 \leqslant |i-j| \leqslant m, \\ &= 0, \qquad |i-j| > m, \qquad i, j = 1, \cdots, N, \end{aligned} \tag{2.2}$$

where ρ_k, $k = 1, \cdots, m$, are constants such that R_N is positive definite and moreover $1 + 2\Sigma_{k=1}^{m}\rho_k > 0$.

To determine the performance of ψ_t in this situation, we note in the first place that $N^{\frac{1}{2}}(\bar{X} - \mu)/\sigma$ is normally distributed with mean 0 and variance $1 + 2\Sigma_{k=1}^{m}(1 - kN^{-1})\rho_k$. Together with the fact that $S^2 \to_P \sigma^2$ this implies in view of Slutsky's theorem (see, e.g., Cramér (1946), page 254), that $N^{\frac{1}{2}}(\bar{X} - \mu)/S$ is asymptotically normal with mean zero and variance $1 + 2\Sigma_{k=1}^{m}(1 - kN^{-1})\rho_k$. Now suppose that the critical value of ψ_t is selected such that the size is α under independence and let u_α be such that $\alpha = 1 - \Phi(u_\alpha)$, where Φ is the standard normal distribution function (df). Then for general ρ_k, $k = 1, \cdots, m$, the size of ψ_t equals $1 - \Phi(u_\alpha\{1 + 2\Sigma_{k=1}^{m}(1 - kN^{-1})\rho_k\}^{-\frac{1}{2}}) + o(1)$, which may differ substantially from α if some ρ_k are nonzero.

To obtain a test in this case which has asymptotically the correct size for all possible ρ_k and not merely for $\rho_1 = \cdots = \rho_m = 0$, the following approach can be used. First we need appropriate estimators of the ρ_k. The natural choice seems to be the so-called serial correlation coefficients (s.c.c.'s)

$$\hat{\rho}_k = \frac{\Sigma_{i=1}^{N}X_iX_{i+k} - N\bar{X}^2}{\Sigma_{i=1}^{N}X_i^2 - N\bar{X}^2} = \frac{\Sigma_{i=1}^{N}(X_i - \bar{X})(X_{i+k} - \bar{X})}{(N-1)S^2}, \qquad k = 1, 2, \cdots, \tag{2.3}$$

where by definition X_{N+k} equals X_k. These s.c.c.'s have certain optimality properties. For example, according to Walsh (1962), page 71, in the normal case the most powerful permutation test for testing independence against $\rho_1 > 0$ is based on $\hat{\rho}_1$. In (2.3) we have given the circular version of the s.c.c. Another possibility is the noncircular s.c.c.

$$\{\Sigma_{i=1}^{N-k}(X_i - \bar{X})(X_{i+k} - \bar{X})\}/\{(N-1)S^2\}.$$

Fortunately it is irrelevant here which version is preferred, since the results of the paper are the same for both choices, as can be verified easily.

Next we define the statistic W_m on the set $B = \{1 + 2\Sigma_{k=1}^m \hat{\rho}_k > 0\}$ by

$$W_m^2 = S^2(1 + 2\Sigma_{k=1}^m \hat{\rho}_k). \tag{2.4}$$

Let W_m be bounded, but otherwise arbitrary, on B^c. Now we have the following result:

LEMMA 2.1. *Let $(X_1, \cdots, X_N)$ be jointly normally distributed with $EX_i = \mu$ and $\sigma^2(X_i) = \sigma^2 > 0$ and suppose that* (2.2) *holds. Then, for all x,*

$$P\big(N^{\frac{1}{2}}(\bar{X} - \mu)/W_m \leqslant x\big) \to \Phi(x) \tag{2.5}$$

as $N \to \infty$, where the convergence is uniform for all σ^2 and ρ_k, $k = 1, \cdots, m$, such that σ^2 and $(1 + 2\Sigma_{k=1}^m \rho_k)$ are bounded away from zero. Consequently, the test $\psi_V^{(m)}$ which rejects $H_0 : \mu = 0$ for large values of

$$V_m = N^{\frac{1}{2}}\bar{X}/W_m \tag{2.6}$$

has asymptotically the same size for all ρ_k such that $(1 + 2\Sigma_{k=1}^m \rho_k)$ is bounded away from zero.

PROOF. By using Chebyshev's inequality for the $2r$th moment we obtain that $P(|S^2 - \sigma^2| \geqslant \varepsilon) = O(N^{-r})$, for all positive integers r and all $\varepsilon > 0$. Likewise $P(|(N-1)^{-1}\Sigma_{i=1}^N (X_i - \bar{X})(X_{i+k} - \bar{X}) - \rho_k\sigma^2| \geqslant \varepsilon) = O(N^{-r})$. Hence $P(|\hat{\rho}_k - \rho_k| \geqslant \varepsilon) = O(N^{-r})$ and therefore

$$P\big(|1 + 2\Sigma_{k=1}^m \hat{\rho}_k - \big(1 + 2\Sigma_{k=1}^m (1 - kN^{-1})\rho_k\big)| \geqslant 2m\varepsilon\big) = O(N^{-r}).$$

Moreover, as $1 + 2\Sigma_{k=1}^m \rho_k$ is bounded away from zero, we also have that $P(B^c) = O(N^{-r})$. These results imply together with (2.4) that

$$W_m^2 - \sigma^2\big(1 + 2\Sigma_{k=1}^m (1 - kN^{-1})\rho_k\big) \to_P 0.$$

As $N^{\frac{1}{2}}(\bar{X} - \mu)$ is normal with mean zero and variance $\sigma^2(1 + 2\Sigma_{k=1}^m (1 - kN^{-1})\rho_k)$, it now follows from Slutsky's theorem that (2.5) holds. From this result, the conclusion about the test $\psi_V^{(m)}$ is immediate. □

Note that $\psi_V^{(0)}$ is ψ_t. Let $\xi_\alpha^{(m)}$ be the critical value of $\psi_V^{(m)}$ that leads to size α, then $\xi_\alpha^{(m)}$ is asymptotically independent of σ^2 and ρ_k, $k = 1, \cdots, m$, in the sense that $\xi_\alpha^{(m)} \to u_\alpha$ as $N \to \infty$, uniformly for all σ^2 and $(1 + 2\Sigma_{k=1}^m \rho_k)$ bounded away from zero. In the next section (see formula (3.4)) we shall give an approximation $\bar{\xi}_\alpha^{(m)}$ such that $|\xi_\alpha^{(m)} - \bar{\xi}_\alpha^{(m)}| = O(N^{-2})$ under independence. To conclude the present section, we consider the question whether $\psi_V^{(m)}$ is optimal in some sense.

LEMMA 2.2. *Under the conditions of Lemma* 2.1 *the test $\psi_V^{(m)}$ is asymptotically equivalent to the optimal t-test for known ρ_k, $k = 1, \cdots, m$.*

PROOF. Let $U_N = (u_{ij})$ be the inverse of R_N. If $\rho_1, \cdots, \rho_m$, and hence R_N, are known, it follows from Scheffé (1959) (see pages 20–30) that the optimal t-test is based on

$$L = \big(\tilde{S}^2\Sigma_{i=1}^N\Sigma_{j=1}^N u_{ij}\big)^{-\frac{1}{2}}\Sigma_{i=1}^N\big(\Sigma_{j=1}^N u_{ij}\big)X_i, \tag{2.7}$$

where $\tilde{S}^2 = \{\Sigma_{i=1}^N\Sigma_{j=1}^N u_{ij}X_iX_j - (\Sigma_{i=1}^N\Sigma_{j=1}^N u_{ij}X_i)^2/(\Sigma_{i=1}^N\Sigma_{j=1}^N u_{ij})\}/(N-1)$. Now R_N is a so-called Laurent matrix, i.e., the elements in any diagonal running parallel to the principal diagonal are the same. The inverse of such a Laurent matrix is approximately a Laurent matrix also, and it approaches Laurent form exactly as $N \to \infty$. The fact that the inverse is not exactly a Laurent matrix may be ascribed to the "end-effects" of a finite series (see Whittle (1951), page 34). Moreover, as the covariance matrix $\sigma^2 R_N$ corresponds to a moving average scheme, its inverse $\sigma^{-2}U_N$ is—again apart from end-effects—the covariance matrix of an autoregressive scheme (see Whittle (1951), pages 34–35). This implies in particular that the elements of U_N converge to zero at an exponential rate as they become more remote from the principal diagonal (see again Whittle (1951), page 35). Together with the fact that U_N is asymptotically Laurent, this shows that $\Sigma_{j=1}^N u_{ij}$ is approximately constant in i. (To be more precise, for each $\varepsilon > 0$ there exists a constant C such that $|\Sigma_{j=1}^N u_{ij} - N^{-1}\Sigma_{i=1}^N \Sigma_{j=1}^N u_{ij}| \leqslant \varepsilon$ for $C \leqslant i \leqslant N - C$.) As the X_i possess moments of arbitrary high order, it follows in view of (2.7) that the optimal t-test is asymptotically equivalent to the test based on $(\Sigma_{i=1}^N\Sigma_{j=1}^N u_{ij})^{\frac{1}{2}}\overline{X}/\tilde{S}$. Using once more the (approximate) Laurent form of R_N and U_N, we obtain that $(N^{-1}\Sigma_{i=1}^N\Sigma_{j=1}^N u_{ij})(1 + 2\Sigma_{k=1}^m\rho_k) \to 1$ as $N \to \infty$. Hence there also exists asymptotic equivalence between the optimal t-test and the test based on $\{\tilde{S}^2(1 + 2\Sigma_{k=1}^m\rho_k)\}^{-\frac{1}{2}}N^{\frac{1}{2}}\overline{X}$. Now note that $\tilde{S}^2$ and $S^2 = (N-1)^{-1}\Sigma_{i=1}^N(X_i - \overline{X})^2$ are consistent estimators of σ^2 and that $\hat{\rho}_k$ is consistent for ρ_k, then the desired result follows in view of (2.4) and (2.6). □

3. Comparing the performance of $\psi_V^{(m)}$ and ψ_t under independence. The price we have to pay for the robustness of validity of $\psi_V^{(m)}$ is, of course, that if all ρ_k happen to be 0, i.e., if $X_1, \cdots, X_N$ are indeed independent, we could have done better by using ψ_t rather than $\psi_V^{(m)}$. Since our willingness to use $\psi_V^{(m)}$ for a certain m instead of ψ_t will heavily depend on the height of this price, we will devote this section to a detailed comparison of the performance of $\psi_V^{(m)}$ and ψ_t in the case of independence. Hence, in the remainder of this section we will always assume that $\rho_k = 0$ for all k.

In the first place we note that it immediately follows from the previous section that $\psi_V^{(m)}$ has power

$$\pi_V^{(m)}(\mu) = 1 - \Phi\left(u_\alpha - N^{\frac{1}{2}}\mu/\left\{\sigma^2(1 + 2\Sigma_{k=1}^m\rho_k)\right\}^{\frac{1}{2}}\right) + o(1)$$

against alternatives μ. As $\rho_k = 0$, $k = 1, \cdots, m$, this coincides to $o(1)$ with the power $\pi_t(\mu)$ of ψ_t and therefore $\psi_V^{(m)}$ has asymptotic relative efficiency $e = 1$ with respect to ψ_t when $X_1, \cdots, X_N$ are independent. This means that if ψ_t requires N observations to attain a certain power and $\psi_V^{(m)}$ requires k_N observations to reach that same power, then $k_N/N \to 1$ as $N \to \infty$.

In view of this result, the price of using $\psi_V^{(m)}$ for some m instead of ψ_t might be reasonable. To obtain more information on this point, we use the approach suggested by Hodges and Lehmann (1970). They compare statistical procedures by studying the asymptotic behaviour of the so-called deficiency $d_N = k_N - N$, the additional number of observations required by the less effective procedure. In one of their examples (see page 792), the deficiency of the t-test with respect to the test based on $N^{\frac{1}{2}}\bar{X}/\sigma$ is obtained. By a similar derivation we shall obtain $d_N(V_m, t)$.

First, we note that $\bar{X}$ is sufficient for μ, whereas the distribution of S^2 or $\hat{\rho}_k$ does not depend on μ. Under these circumstances it follows from Lehmann (1959), Ch. 5.1, Corollary 1 and Example 1 that $\bar{X}$ and $(S^2, \hat{\rho}_k)$ are independent. Together with (2.6) this leads to

$$\begin{aligned}\pi_V^{(m)}(\mu) &= P_\mu(V_m \geqslant \xi_\alpha^{(m)}) \\ &= EP_\mu\big(N^{\frac{1}{2}}\bar{X} \geqslant \xi_\alpha^{(m)} W_m | S^2, \hat{\rho}_1, \cdots, \hat{\rho}_m\big) \\ &= 1 - E\Phi\big(\{\xi_\alpha^{(m)} W_m - N^{\frac{1}{2}}\mu\}/\sigma\big),\end{aligned} \tag{3.1}$$

where W_m is given by (2.5) and P_μ denotes probability under μ. As $W_m/\sigma - 1 = O_P(N^{-\frac{1}{2}})$, we can expand the last expression in (3.1) around $\xi_\alpha^{(m)} - N^{\frac{1}{2}}\mu/\sigma$.

Before thus obtaining an expansion for $\pi_V^{(m)}(\mu)$, we first derive some moments of $(W_m/\sigma - 1)$. In doing so, we may assume $\mu = 0$ and $\sigma = 1$ without loss of generality since S/σ and $\hat{\rho}_k$ are translation and scale invariant. In the same way as it was shown that $P(B^c) = O(N^{-r})$, it can be demonstrated that there exists $\varepsilon > 0$ such that $P(B_\varepsilon^c) = O(N^{-r})$, where $B_\varepsilon = \{1 + 2\Sigma_{k=1}^m \hat{\rho}_k \geqslant \varepsilon\} \subset B$. Now W_m is bounded on B_ε^c and therefore $E|W_m - 1|^n I_{B_\varepsilon^c} = O(N^{-r})$, where $I_{B_\varepsilon^c}$ is the indicator function of B_ε^c. Hence $E(W_m - 1)^n = E(W_m - 1)^n I_{B_\varepsilon} + O(N^{-r})$.

On B_ε we have

$$\begin{aligned}W_m &= \{1 + (S^2 - 1) + 2\Sigma_{k=1}^m \hat{\rho}_k S^2\}^{\frac{1}{2}} \\ &= \Sigma_{a=0}^{b-1}\binom{\frac{1}{2}}{a}\{(S^2 - 1) + 2\Sigma_{k=1}^m \hat{\rho}_k S^2\}^a + O\big(|(S^2 - 1) + 2\Sigma_{k=1}^m \hat{\rho}_k S^2|^b\big), \\ &\qquad b = 1, 2, \cdots .\end{aligned}$$

Furthermore, straightforward calculations show that the following expressions are all $O(N^{-2})$: $E(S^2 - 1)$, $E(\hat{\rho}_k S^2) + N^{-1}$, $E(S^2 - 1)^2 - 2N^{-1}$, $E(\hat{\rho}_k^2 S^4) - N^{-1}$, $E\{(S^2 - 1)\hat{\rho}_k S^2\}$, $E(\hat{\rho}_j \hat{\rho}_k S^4)$, $j, k = 1, \cdots, m$, $j \neq k$. The same holds for all expectations containing three or more factors $S^2 - 1$ or $\rho_k S^2$ and for all absolute expectations containing four or more of these factors. These results enable us to find $E(W_m - 1)^n I_{B_\varepsilon}$ to $O(N^{-2})$, and hence $E(W_m - 1)^n$ to $O(N^{-2})$. It turns out that

$$\begin{aligned}EW_m - 1 &= -\frac{(6m + 1)}{4N} + O(N^{-2}), \\ E(W_m - 1)^2 &= \frac{(2m + 1)}{2N} + O(N^{-2}), \\ E(W_m - 1)^n &= O(N^{-2}),\end{aligned} \tag{3.2}$$

where $n = 3, 4$.

Now we use a Taylor expansion to order 3 for the last expression in (3.1), around $\xi_\alpha^{(m)} - N^{\frac{1}{2}}\mu/\sigma$. Then we apply the results in (3.2) to this expansion and note that the error term is $O(N^{-2})$, uniformly in μ, as ϕ'' is bounded. Using the fact that $\phi'(x) = -x\phi(x)$ and once more the expansion of Φ (cf. Hodges and Lehmann (1970)), we arrive at

$$(3.3)\quad \pi_V^{(m)}(\mu) = 1 - \Phi\Big(\xi_\alpha^{(m)} - N^{\frac{1}{2}}\mu/\sigma + (4N)^{-1}\big\{(2m+1)(\xi_\alpha^{(m)})^2 N^{\frac{1}{2}}\mu/\sigma - (6m+1)\xi_\alpha^{(m)} - (2m+1)(\xi_\alpha^{(m)})^3\big\}\Big) + O(N^{-2}).$$

From $\pi_V^{(m)}(0) = \alpha$ it then follows that under independence $\xi_\alpha^{(m)} = \bar{\xi}_\alpha^{(m)} + O(N^{-2})$, where

$$(3.4)\qquad \bar{\xi}_\alpha^{(m)} = u_\alpha + (4N)^{-1}\{(6m+1)u_\alpha + (2m+1)u_\alpha^3\}.$$

Using (3.3) and (3.4) we finally obtain the following expansion

$$(3.5)\qquad \pi_V^{(m)}(\mu) = 1 - \Phi\left(u_\alpha - \frac{N^{\frac{1}{2}}\mu}{\sigma}\left\{1 - \frac{(2m+1)u_\alpha^2}{4N}\right\}\right) + O(N^{-2}).$$

From the definition of deficiency, the fact that $\psi_t = \psi_V^{(0)}$ and (3.5) we now obtain

$$(3.6)\qquad d_N(V_m, t) = mu_\alpha^2 + O(N^{-1}).$$

This means that in the normal case protection against the effects of m-dependence asymptotically requires mu_α^2 additional observations. Now u_α^2 itself is a reasonably small number. For example, as α decreases from 0.05 to 0.01, u_α^2 increases from 2.7 to 5.4. Hence, especially for small m, one might consider paying the price for using $\psi_V^{(m)}$ instead of ψ_t if the independence of the observations is doubted.

Some final remarks are:

(i) In (2.7) we considered the statistic L of the optimal t-test for known ρ_k, $k = 1, \cdots, m$. Using the $\hat{\rho}_k$ we can estimate the u_{ij} in L. Let $\hat{L}$ be the resulting statistic and let $\psi_{\hat{L}}$ be the corresponding test. Then it follows from Lemma 2.2 that $\psi_{\hat{L}}$ is asymptotically equivalent to $\psi_V^{(m)}$. Note, however, that since deficiency is a second-order property this does not imply that both tests have the same deficiency with respect to ψ_t. But, as V_m resembles $t = N^{\frac{1}{2}}\bar{X}/S$ more closely than $\hat{L}$ (V_m and t, e.g., have the same numerator) and ψ_t is optimal for $\rho_k = 0$, $d_N(V_m, t)$ should certainly not be larger than $d_N(\hat{L}, t)$.

(ii) A test for which the deficiency with respect to ψ_t tends to 0 as $N \to \infty$ can be derived from $\psi_V^{(m)}$ as follows: for some $0 < \delta < \frac{1}{2}$, replace W_m^2 in (2.4) by $(W_m^*)^2$ which equals S^2 if $|\hat{\rho}_k| \leqslant N^{-\frac{1}{2}+\delta}$ for $k = 1, \cdots, m$ and which equals W_m^2 otherwise. As $P(|\hat{\rho}_k| > N^{-\frac{1}{2}+\delta}) = O(N^{-r})$ for $0 < \delta < \frac{1}{2}$ and $r > 0$ when $\rho_k = 0$, $k = 1, \cdots, m$, the statistic $V_m^* = N^{\frac{1}{2}}\bar{X}/W_m^*$ then coincides with t except on a set of probability $O(N^{-r})$ and hence $d_N(V_m^*, t) = O(N^{-r+1})$. On the other hand, the test $\psi_{V^*}^{(m)}$ based on V_m^* is asymptotically equivalent to $\psi_V^{(m)}$ for all fixed ρ_k such that $\rho_k \neq 0$ for at least one k.

(iii) As the distribution of $\hat{\rho}_1$ concentrates around zero rather slowly as $N \to \infty$ (cf. Anderson (1971), page 319), one probably needs moderately large sample sizes before the asymptotic results show a satisfactory agreement with the exact values. To improve this, one could modify W_m^2 in several ways, e.g. by replacing $\hat{\rho}_k$ by $(1 - k/N)(\hat{\rho}_k - E\hat{\rho}_k)$, where E denotes expectation under the assumption that all ρ_k are zero.

4. Autoregressive processes. In this section we shall briefly show that the approach of Sections 2 and 3 also applies to autoregressive processes. Let $(X_1, \cdots, X_N)$ again be jointly normally distributed rv's with $EX_i = \mu$ and $\sigma^2(X_i) = \sigma^2 > 0$, $i = 1, \cdots, N$. However, suppose now that the X_i form a stationary solution of an autoregressive equation of order m, i.e.,

$$\Sigma_{k=0}^{m} a_k(X_{i-k} - \mu) = Z_i, \qquad i = m+1, \cdots, N. \tag{4.1}$$

Here $a_0 = 1, a_1, \cdots, a_m$ are certain constants and $Z_{m+1}, \cdots, Z_N$ are independent identically normally distributed rv's with $EZ_i = 0$ and $\sigma^2(Z_i) = \tau^2 > 0$, $i = m + 1, \cdots, N$. Moreover the a_k are such that all roots of the equation $\Sigma_{k=0}^{m} a_k w^{m-k} = 0$ lie inside the unit disk. This latter condition is necessary and sufficient for the existence of a stationary solution of (4.1) (see, e.g., Feller (1966), page 90). Note that it implies in particular that $\Sigma_{k=0}^{m} a_k \neq 0$.

Using results of Anderson (1971) (Section 5.2) we obtain that $N^{\frac{1}{2}}(\bar{X} - \mu)$ is normally distributed with mean 0 and variance

$$\left(\sigma^2 \Sigma_{k=0}^{m} a_k \rho_k\right) / \left(\Sigma_{k=0}^{m} a_k\right)^2 + O(mN^{-1}). \tag{4.2}$$

Hence, like in Section 2, ψ_t should be replaced by a test $\psi_{\tilde{V}}^{(m)}$ based on a statistic $\tilde{V} = N^{\frac{1}{2}}\bar{X}/\tilde{W}$, where $\tilde{W}^2$ is some consistent estimator of the variance in (4.2). To estimate this expression we not only need the s.c.c.'s $\hat{\rho}_k$, but also consistent estimators $\hat{a}_k$ for a_k, $k = 1, \cdots, m$. These are given by Anderson (1971) (see Section 5.4) and can be calculated recursively: let $\hat{b}_1, \cdots, \hat{b}_{m-1}$ be the estimated coefficients of the $(m-1)$st order equation, then

$$\hat{a}_m = -\frac{\hat{\rho}_m + \Sigma_{k=1}^{m-1} \hat{\rho}_{m-k} \hat{b}_k}{1 + \Sigma_{k=1}^{m-1} \hat{\rho}_k \hat{b}_k}, \tag{4.3}$$

$$\hat{a}_k = \hat{b}_k + \hat{b}_{m-k} \hat{a}_m, \qquad k = 1, \cdots, m-1.$$

Now define, in analogy to (2.4),

$$\tilde{W}_m^2 = S^2 \Sigma_{k=0}^{m} \hat{a}_k \hat{\rho}_k / \left(\Sigma_{k=0}^{m} \hat{a}_k\right)^2 \tag{4.4}$$

on the set $\{\Sigma_{k=0}^{m} \hat{a}_k \hat{\rho}_k > 0\}$ and let $\tilde{W}_m^2$ be bounded but arbitrary otherwise. Then results similar to those of Section 2 hold for the test $\psi_{\tilde{V}}^{(m)}$. As an example, for $m = 1$, (4.3) gives $\hat{a}_1 = -\hat{\rho}_1$ and therefore $\tilde{W}_1^2 = S^2(1 + \hat{\rho}_1)/(1 - \hat{\rho}_1)$.

The deficiency $d_N(\tilde{V}_m, t)$ of $\psi_{\tilde{V}}^{(m)}$ with respect to ψ_t under independence can be obtained largely in the same way in which $d_N(V_m, t)$ was found in Section 3. In the first place, we remark that $\hat{a}_k$ only depends on $\hat{\rho}_1, \cdots, \hat{\rho}_m$. Consequently, $\tilde{W}_m^2$ in

(4.4) is a function of $S^2, \hat{\rho}_1, \cdots, \hat{\rho}_m$ only and, therefore, $\tilde{W}_m^2$ and $\bar{X}$ are independent. Hence (3.1) continues to hold if V, V_m and W_m are replaced by $\tilde{V}$, $\tilde{V}_m$ and $\tilde{W}_m$.

Furthermore, if $\rho_k = 0$, $k = 1, \cdots, m$, it follows from (4.3) that $\hat{a}_k$ can be expanded in powers of $\hat{\rho}_1, \cdots, \hat{\rho}_m$, except on a set of probability $O(N^{-r})$, with r arbitrarily large. Straightforward computation yields that on the complement of this set $\hat{a}_k = -\hat{\rho}_k + \frac{1}{2}\{1 + (-1)^k\}\hat{\rho}_{k/2}^2 + R_k + O(\Sigma_{k=1}^m \hat{\rho}_k^4)$, $k = 1, \cdots, m$, where each R_k consists of a finite number of terms either of the form $\hat{\rho}_i\hat{\rho}_j$ with $i \neq j$ or of the form $\hat{\rho}_i\hat{\rho}_j\hat{\rho}_k$. Together with (4.4) this leads to

$$\tilde{W}_m^2 = S^2\{1 + 2\Sigma_{k=1}^m \hat{\rho}_k + 2\Sigma_{k=[m/2]+1}^m \hat{\rho}_k^2 + R + O(\Sigma_{k=1}^m \hat{\rho}_k^4)\}, \tag{4.5}$$

again except on a set of probability $O(N^{-r})$. Here $[m/2]$ denotes the integer part of $m/2$ and R is of the same form as the R_k above.

Noting the similarity of (2.5) and (4.5) and applying the moment results of Section 3, we obtain that (3.2) continues to hold for $\tilde{W}_m$ instead of W_m, provided that we replace $(6m + 1)$ by $4m + (-1)^m$. Hence, making the same replacement in (3.4) gives the critical value of $\psi_{\tilde{V}}^{(m)}$ to $O(N^{-2})$, while the results for the power and the deficiency in (3.5) and (3.6) respectively, remain valid in the present case. Thus, the price for robustness against autoregressive departures from the independence assumption also tends to mu_α^2 additional observations.

REFERENCES

[1] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.
[2] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton.
[3] Feller, W. (1966). *An Introduction to Probability Theory and its Applications, II*. Wiley, New York.
[4] Gastwirth, J. L. and Rubin, H. (1971). Effect of dependence on the level of some one-sample tests. *J. Amer. Statist. Assoc.* **66** 816–820.
[5] Gastwirth, J. L. and Rubin, H. (1975). The behavior of robust estimators on dependent data. *Ann. Statist.* **3** 1070–1100.
[6] Gastwirth, J. L., Rubin, H. and Wolff, S. S. (1967). The effect of autoregressive dependence on a nonparametric test. *IEEE Trans. Information Theory* **IT-13** 311–313.
[7] Hodges, J. L. Jr. and Lehmann, E. L. (1970). Deficiency. *Ann. Math. Statist.* **41** 783–801.
[8] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.
[9] Modestino, J. W. (1969). Nonparametric and adaptive detecting on dependent data. Technical Report 27, Dept. of Electrical Engineering, Princeton University.
[10] Scheffé, H. (1959). *The Analysis of Variance*. Wiley, New York.
[11] Walsh, J. E. (1962). *Handbook of Nonparametric Statistics I*. Van Nostrand, New York.
[12] Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Almquist, Uppsala.

Twente University of Technology
Department of Mathematics
P.O. Box 217
Enschede, The Netherlands