The Annals of Statistics
1978, Vol. 6, No. 4, 836–845

ONE-SAMPLE RANK TESTS UNDER AUTOREGRESSIVE DEPENDENCE[1]

BY W. ALBERS

Technological University Twente, Enschede

One-sample linear rank tests are considered for the case where the observations are not independent but come from an autoregressive process. It is proposed to apply the tests under these circumstances to certain transformations of the observations, rather than to the observations themselves. Then the tests have asymptotically the same properties as under independence, both under the hypothesis and under contiguous location alternatives. In particular, they are asymptotically distribution-free.

1. Introduction. In the standard formulation of the one-sample problem an independent sample from a common distribution is given, on the basis of which the hypothesis has to be tested that this distribution is symmetric about zero. Widely used tests for this problem are one-sample linear rank tests. These tests have the desirable property that under the hypothesis the distribution of their test statistic does not depend on the generally unknown underlying distribution.

Unfortunately, this nice property does not continue to hold if we drop the assumption that the observations are independent. This sensitivity of nonparametric tests to dependence has been noted 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).

For dependent observations in general, it is not clear how distribution-free tests can be obtained (cf. the remarks in Section 4 of Gastwirth and Rubin (1971)). For special types of dependence, however, it may be possible to find such tests. In this paper we shall consider the case where the observations come from an autoregressive process with independent symmetric errors.

Then we have the following situation: let m be a fixed nonnegative integer. For $N = 1, 2, \cdots$, consider $(m + N)$ identically distributed random variables

Received July 1976; revised August 1977.

[1] This research was supported by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.). This paper was written while the author was visiting the Statistics Department of the University of California at Berkeley.

AMS 1970 *subject classifications.* 62G10, 62G20.

Key words and phrases. One-sample problem, rank tests, contiguous location alternatives, autoregressive processes.

(rv's) $X_{-m+1}, \cdots, X_{-1}, X_0, X_1, \cdots, X_N$. Suppose that the X_i form a stationary solution of the following autoregressive equation:

$$Z_i = \sum_{k=0}^{m} a_k X_{i-k}, \qquad i = 1, \cdots, N. \tag{1.1}$$

Here the Z_i are independent identically distributed (i.i.d) rv's, $a_0 = 1$ and $a_1, \cdots, a_m$ are constants such that all roots of $\sum_{k=0}^{m} a_k w^{m-k} = 0$ lie inside the unit disk. This latter condition ensures the existence of a stationary solution of (1.1) (see, e.g., Anderson (1971), Theorem 5.2.1, page 170). Note that it in particular implies that $\sum_{k=0}^{m} a_k > 0$. Let us also assume that the Z_i are symmetrically distributed if the X_i are symmetrically distributed.

In Section 2 we first consider the case where $a_1, \cdots, a_m$ are known. Then it is easy to obtain a distribution-free test, as the Z_i can be found from the X_i through (1.1). As $EZ_i = EX_i(\sum_{k=0}^{m} a_k)$, with $\sum_{k=0}^{m} a_k > 0$, we can test the hypothesis that the distribution of the X_i is symmetric about zero on the basis of $(Z_1, \cdots, Z_N)$. But since the Z_i are i.i.d., this is just the standard one-sample problem and any linear rank test will do.

Unfortunately, the solution above is not only simple, but also unrealistic, as $a_1, \cdots, a_m$ are genenerally unknown. In the general case we proceed as follows: first we note that consistent estimators $\hat{a}_k$ for a_k, $k = 1, \cdots, m$, based on $X_{-m+1}, \cdots, X_N$, are given by Anderson (1971). Then we introduce

$$\hat{Z}_i = \sum_{k=0}^{m} \hat{a}_k X_{i-k}, \qquad i = 1, \cdots, N \tag{1.2}$$

and propose to use linear rank tests based on $(\hat{Z}_1, \cdots, \hat{Z}_N)$. In Section 3 we show that such tests have the same asymptotic behavior as those based on $(Z_1, \cdots, Z_N)$, both under the hypothesis and under contiguous location alternatives. Hence, in particular, such tests are asymptotically distribution-free.

In the above we have considered one-sample rank tests for the hypothesis of symmetry about a specific value. Such tests can also be used for the hypothesis of symmetry in general, by applying them to $X_i - \hat{\mu}$ rather than to X_i, where $\hat{\mu}$ is some estimator of the center of symmetry. However, it should be noted that the result above does not extend without additional conditions to the latter case. In fact, Gastwirth (1971) has shown that the sign test based on $X_i - \bar{X}$, where $\bar{X}$ is the sample mean, is not even asymptotically distribution-free for independent X_i.

2. Rank tests for autoregressive processes. We shall consider the hypothesis of symmetry and contiguous location alternatives. More precisely, let $\theta (= \theta_N)$ be such that, for some positive constant C,

$$0 \leqq \theta \leqq CN^{-\frac{1}{2}}. \tag{2.1}$$

Assume that the joint distribution of $(m + 1)$ consecutive $(X_i - \theta)$ has a density f_{m+1} such that, for all $(x_1, \cdots, x_{m+1})$,

$$f_{m+1}(-x_1, \cdots, -x_{m+1}) = f_{m+1}(x_1, \cdots, x_{m+1}). \tag{2.2}$$

This implies that the distribution function (df) F of $(X_i - \theta)$ has a density f

that is symmetric about zero. Moreover, suppose that $0 < \sigma^2(X_i) = \sigma^2 < \infty$. Then it follows from (1.1) and (2.2) that the df G of $(Z_i - \theta \sum_{k=0}^m a_k)$ has a density g that is also symmetric about zero. Assume that $\sigma^2(Z_i) = \tau^2 > 0$. Let ρ_k be the correlation coefficient of X_i and X_{i+k}. Since the X_i form a stationary solution of (1.1), we have that X_i is independent of $Z_{i+1}, Z_{i+2}, \cdots$ (see Anderson (1971), Corollary 5.2.1, page 170). Hence

$$\tau^2 = E\{(Z_i - \theta \textstyle\sum_{k=0}^m a_k)(X_i - \theta)\} = \sigma^2 \sum_{k=0}^m a_k \rho_k . \tag{2.3}$$

Let J be a square-integrable function on (0, 1) and let $R_1, \cdots, R_N$ be the ranks of $|Z_1|, \cdots, |Z_N|$. We define the linear rank statistic

$$S = \textstyle\sum_{i=1}^N J\left(\frac{R_i}{N+1}\right) \operatorname{sign} Z_i , \tag{2.4}$$

where sign $x = 1$ for $x > 0$, sign $x = 0$ for $x = 0$ and sign $x = -1$ for $x < 0$. The properties of the level α test ψ_S which rejects H_0: $\theta = 0$ for large values of S are well known. From Hájek and Šidák (1967) (see Theorem V.1.7, page 166) it follows that its critical value ξ_α satisfies

$$\xi_\alpha = N^{\frac{1}{2}} u_\alpha \{\textstyle\int_0^1 J^2(t)\, dt\}^{\frac{1}{2}} + o(N^{\frac{1}{2}}) , \tag{2.5}$$

where u_α is given by $\alpha = 1 - \Phi(u_\alpha)$, in which Φ is the standard normal df. Furthermore, let $\pi_S(\theta)$ be the power of ψ_S and let $\Psi_{\tilde{g}}(t) = -\tilde{g}'(\tilde{G}^{-1}([1 + t]/2))/\tilde{g}(\tilde{G}^{-1}([1 + t]/2))$, where $\tilde{g}(x) = g(x/\tau)$ is the standard density of type g. If $\int_0^1 \Psi_{\tilde{g}}^2(t)\, dt < \infty$, then Theorem VI.2.5 on page 220 of Hájek and Šidák implies that

$$\pi_S(\theta) = 1 - \Phi\left\{u_\alpha - \frac{N^{\frac{1}{2}}\theta}{\sigma} \cdot \frac{\sum_{k=0}^m a_k}{(\sum_{k=0}^m a_k \rho_k)^{\frac{1}{2}}} \cdot \frac{\int_0^1 J(t)\Psi_{\tilde{g}}(t)\, dt}{(\int_0^1 J^2(t)\, dt)^{\frac{1}{2}}}\right\} + o(1) , \tag{2.6}$$

where we have used (2.3) and the fact that $EZ_i = \theta \sum_{k=0}^m a_k$.

Next we shall introduce consistent estimators $\hat{a}_k$ for a_k. First we define the so-called serial correlation coefficients

$$\hat{\rho}_k = \frac{\sum_{i=-m+1}^{N-k} X_i X_{i+k} - (m + N)\bar{X}^2}{\sum_{i=-m+1}^{N} X_i^2 - (m + N)\bar{X}^2} , \qquad k = 1, \cdots, m , \tag{2.7}$$

where $\bar{X} = (m + N)^{-1} \sum_{i=-m+1}^N X_i$. As $0 < \sigma^2 < \infty$, it follows that $\hat{\rho}_k$ is consistent for ρ_k, $k = 1, \cdots, m$. From these $\hat{\rho}_k$, Anderson (1971) (see Section 5.4) derives the $\hat{a}_k$ as follows: in view of (1.1) we have that $R^{(m)}a = -r^{(m)}$, where $a^T = (a_1, \cdots, a_m)$, $R^{(n)}$ is the covariance matrix of n consecutive X_i and $(r^{(n)})^T = (\rho_1, \cdots, \rho_n)$, $n = 1, \cdots, m$. Let $\hat{R}^{(n)}$ and $(\hat{r}^{(n)})^T$ be obtained from $R^{(n)}$ and $(r^{(n)})^T$ through replacing ρ_k by $\hat{\rho}_k$ everywhere, for $k = 1, \cdots, n$. Since $\hat{R}^{(n)}$ is positive definite (see Anderson (1971), page 187), the obvious choice for $\hat{a}^T = (\hat{a}_1, \cdots, \hat{a}_m)$ then is $\hat{a} = -(\hat{R}^{(m)})^{-1}\hat{r}^{(m)}$. From this result $\hat{a}_k$ can be calculated recursively: let $\tilde{a}^T = (\tilde{a}_1, \cdots, \tilde{a}_{m-1})$ be defined by $\tilde{a} = -(\hat{R}^{(m-1)})^{-1}\hat{r}^{(m-1)}$, then

$$\hat{a}_m = \frac{-\hat{\rho}_m + \sum_{k=1}^{m-1} \hat{\rho}_{m-k}\tilde{a}_k}{1 + \sum_{k=1}^{m-1} \hat{\rho}_k \tilde{a}_k} , \qquad \hat{a}_k = \tilde{a}_k + \tilde{a}_{m-k}\hat{a}_m , \tag{2.8}$$

$k = 1, \cdots, m-1$. Finally, Anderson also shows that the estimators $\hat{a}_k$ thus obtained are consistent under the present conditions.

As a first application we note that (2.7) and (2.8) immediately yield a consistent estimator $\hat{\pi}_S(\theta)$ of $\pi_S(\theta)$ in (2.6). More important, however, is the fact that (2.8) enables us to evaluate the $\hat{Z}_i$ defined in (1.2). Let $\hat{R}_1, \cdots, \hat{R}_N$ be the ranks of $|\hat{Z}_1|, \cdots, |\hat{Z}_N|$. Then we introduce the level α test $\psi_{\hat{S}}$ which rejects $H_0: \theta = 0$ for large values of

$$\hat{S} = \sum_{i=1}^N J\left(\frac{\hat{R}_i}{N+1}\right) \operatorname{sign} \hat{Z}_i . \tag{2.9}$$

In the next section we shall show that $\psi_{\hat{S}}$ is asymptotically equivalent to ψ_S.

3. Asymptotic equivalence of ψ_S and $\psi_{\hat{S}}$. Let E_θ and P_θ denote expectation and probability under $F(x-\theta)$ for $\theta \geqq 0$. Then our main result is

THEOREM 3.1. *Suppose that θ satisfies* (2.1) *and that the X_i are such that* (1.1) *and* (2.2) *hold. Assume that $E|X_i|^r < \infty$ for some $r > 20$ and that the density g of Z_i is bounded. Finally, suppose that J is differentiable on* (0, 1) *and that its derivative J' is bounded. Then*

$$\sup_x |P_\theta(N^{-\frac{1}{2}}\hat{S} \leqq x) - P_\theta(N^{-\frac{1}{2}}S \leqq x)| = o(1) , \tag{3.1}$$

for $\theta = 0$. If moreover g is such that $\int_0^1 \Psi_g^2(t)\,dt < \infty$, then (3.1) *also holds for $\theta > 0$.*

REMARKS. (1) Let $\hat{\xi}_\alpha$ and $\pi_{\hat{S}}(\theta)$ be the critical value and the power of $\psi_{\hat{S}}$, respectively. Then (3.1) obviously implies that (2.5) and (2.6) remain valid if we replace ξ_α by $\hat{\xi}_\alpha$ and $\pi_S(\theta)$ by $\pi_{\hat{S}}(\theta)$.

(2) As an example in which the conditions of the theorem are clearly satisfied, we mention the case where the X_i (and hence the Z_i) are normally distributed and $J(t) = t$ or $J(t) = 1$, corresponding to Wilcoxon's signed rank test and the sign test, respectively.

To prove Theorem 3.1, we shall use Lemmas 3.1–3.5 below.

LEMMA 3.1. *Suppose that θ satisfies* (2.1) *and that the X_i are such that* (1.1) *and* (2.2) *hold. Let J be square-integrable and assume that*

$$E_0(\hat{S} - S)^2 = o(N) . \tag{3.2}$$

Then the conclusions of Theorem 3.1 *hold.*

PROOF. From Chebyshev's inequality and (3.2) it follows that $N^{-\frac{1}{2}}(\hat{S} - S) \to_{P_0} 0$. Hence by Slutsky's theorem (see, e.g., Cramér (1946), page 254), $N^{-\frac{1}{2}}\hat{S}$ and $N^{-\frac{1}{2}}S$ have the same limit distribution under the hypothesis, i.e., (3.1) holds for $\theta = 0$. If $\int_0^1 \Psi_g^2(t)\,dt < \infty$, we have contiguity and therefore $N^{-\frac{1}{2}}(\hat{S} - S) \to_{P_0} 0$ implies $N^{-\frac{1}{2}}(\hat{S} - S) \to_{P_\theta} 0$. Hence, by the same argument, (3.1). holds for $\theta > 0$. □

In view of Lemma 3.1 it suffices to show that (3.2) holds under the conditions of Theorem 3.1. In doing so, we are concerned with the hypothesis only and therefore we shall in the sequel simply use E and P instead of E_0 and P_0.

We introduce the follwoing notation: let $\hat{S} - S = \sum_{p=1}^{3} T_p$, where

$$(3.3)\qquad \begin{aligned} T_1 &= \sum_{i=1}^{N} \left\{J\left(\frac{\hat{R}_i}{N+1}\right) - J\left(\frac{R_i}{N+1}\right)\right\} \operatorname{sign} Z_i\,, \\ T_2 &= \sum_{i=1}^{N} J\left(\frac{R_i}{N+1}\right)\{\operatorname{sign} \hat{Z}_i - \operatorname{sign} Z_i\}\,, \\ T_3 &= \sum_{i=1}^{N} \left\{J\left(\frac{\hat{R}_i}{N+1}\right) - J\left(\frac{R_i}{N+1}\right)\right\}\{\operatorname{sign} \hat{Z}_i - \operatorname{sign} Z_i\}\,. \end{aligned}$$

Furthermore, for $s = 0, 1, \cdots, N$, let $Z^{(s)} = (Z_1^{(s)}, \cdots, Z_N^{(s)})$, where

$$(3.4)\qquad \begin{aligned} Z_i^{(s)} &= -Z_i\,, \quad i \leqq s\,, \\ &= Z_i\,, \quad\ \ i > s\,, \qquad\qquad i = 1, \cdots, N\,. \end{aligned}$$

In particular, $Z^{(0)} = Z$ and $Z^{(N)} = -Z$. Under (1.1) and (2.2) the Z_i are i.i.d. and symmetrically distributed about zero and therefore the distribution of $Z^{(s)}$ does not depend on s. In particular, if $Eh(Z)$ exists for some function h, then

$$(3.5)\qquad Eh(Z^{(s)}) = Eh(Z)\,,$$

for $s = 0, \cdots, N$. Let $X^{(s)} = (X_{-m+1}^{(s)}, \cdots, X_N^{(s)})$ be the stationary solution of $Z_i^{(s)} = \sum_{k=0}^{m} a_k X_{i-k}^{(s)}$. Again, $X^{(0)} = X$ and $X^{(N)} = -X$. Furthermore, replace X by $X^{(s)}$ in (2.7) and call the result $\hat{\rho}_k^{(s)}$, $k = 1, \cdots, m$. Likewise, replace the $\hat{\rho}_k$ by $\hat{\rho}_k^{(s)}$ in (2.8) and let $\hat{a}_k^{(s)}$, $k = 1, \cdots, m$, be the result. Now $\hat{\rho}_k^{(0)} = \hat{\rho}_k^{(N)} = \hat{\rho}_k$, $\hat{a}_k^{(0)} = \hat{a}_k^{(N)} = \hat{a}_k$, $k = 1, \cdots, m$. Define $\hat{Z}^{(s)} = (\hat{Z}_1^{(s)}, \cdots, \hat{Z}_N^{(s)})$ by $\hat{Z}_i^{(s)} = \sum_{k=0}^{m} \hat{a}_k^{(s)} X_{i-k}^{(s)}$, $i = 1, \cdots, N$, then $\hat{Z}^{(0)} = \hat{Z}$ and $\hat{Z}^{(N)} = -\hat{Z}$. Finally, let $\hat{R}^{(s)} = (\hat{R}_1^{(s)}, \cdots, \hat{R}_N^{(s)})$ be the vector of ranks for $|\hat{Z}^{(s)}| = (|\hat{Z}_1^{(s)}|, \cdots, |\hat{Z}_N^{(s)}|)$, then $\hat{R}^{(0)} = \hat{R}^{(N)} = \hat{R}$. In passing we note that these facts, together with (3.5), immediately yield

$$(3.6)\qquad E\hat{S} = \tfrac{1}{2} \sum_{i=1}^{N} E\left\{J\left(\frac{\hat{R}_i^{(0)}}{N+1}\right) \operatorname{sign} \hat{Z}_i^{(0)} + J\left(\frac{\hat{R}_i^{(N)}}{N+1}\right) \operatorname{sign} \hat{Z}_i^{(N)}\right\} = 0\,.$$

To investigate the relation between X_i and $X_i^{(s)}$, we use the fact that (1.1) implies (see Anderson (1971), page 167)

$$(3.7)\qquad X_i = \sum_{k=0}^{i-1} b_k Z_{i-k} + \sum_{k=i}^{i+m-1} b_{ik}^* X_{i-k}\,, \qquad i = -m+1, \cdots, N\,,$$

where the b_k are given by $\sum_{k=0}^{\infty} b_k z^k = (\sum_{k=0}^{m} a_k z^k)^{-1}$ and the b_{ik}^* by $\sum_{k=i}^{i+m-1} b_{ik}^* z^k = \sum_{k=i}^{\infty} b_k z^k \sum_{k=0}^{m} a_k z^k$. Let $w_1, \cdots, w_m$ be the roots of $\sum_{k=0}^{m} a_k w^{m-k} = 0$; then it follows that b_k can be written as $b_k = \sum_{j=1}^{m} c_j w_j^{k+1}$ for certain constants c_j. By assumption all w_j lie inside the unit disk and hence, uniformly in k,

$$(3.8)\qquad |b_k| = O(e^{-(k+1)c})\,,$$

where $c = -\log\{\max_{1 \leqq j \leqq k} |w_j|\} > 0$. To deal with the b_{ik}^*, we note that $|b_{ik}^*| = O(|b_i|)$.

From (3.4), (3.7), and (3.8) it follows that, for all $\delta > 0$,

$$|X_i^{(s)} - X_i \operatorname{sign}(i-s)| = O(N^{-\lambda} \max_{1\le i\le N} |Z_i|) \tag{3.9}$$

for λ arbitrarily large and uniformly in i and s such that $i \geqq N^\delta$ and $|i-s| \geqq N^\delta$. Now we have the following set of preliminary results.

LEMMA 3.2. *Suppose that the X_i satisfy* (1.1) *and* (2.2). *Let $E|X_i|^r < \infty$ for some $r > 4$. Then for $\varepsilon > 2/r$*

$$P(\max_{1\le i\le N} |Z_i| \geqq N^\varepsilon) = o(N^{-1}), \tag{3.10}$$

$$P(\max_{-m+1\le i\le N} |X_i| \geqq N^\varepsilon) = o(N^{-1}), \tag{3.11}$$

$$P(\max_{1\le k\le m} |\hat{a}_k - a_k| \geqq N^{-\frac{1}{2}+\varepsilon}) = o(N^{-1}), \tag{3.12}$$

$$P(\max_{1\le k\le m} |\hat{a}_k^{(s)} - \hat{a}_k| \geqq N^{-1+2\varepsilon}) = o(N^{-1}). \tag{3.13}$$

PROOF. As $Z_i = \sum_{k=0}^m a_k X_{i-k}$, the fact that $E|X_i|^r < \infty$ implies that $E|Z_i|^r < \infty$. Hence by Chebyshev's inequality, the df G of Z_i satisfies $1 - G(x) \leqq C_1 x^{-r}$ for $x > 0$ and some positive constant C_1. As the Z_i are independent, $P(\max_{1\le i\le N} |Z_i| \geqq x) = 1 - \{2G(x) - 1\}^N = 1 - \{1 - 2(1 - G(x))\}^N \leqq 1 - \{1 - 2C_1 x^{-r}\}^N \leqq 1 - \{1 - 2NC_1 x^{-r}\} = O(Nx^{-r})$ for $x > (2C_1)^{1/r}$. If we choose $x = N^\varepsilon$ with $\varepsilon > 2/r$, then $O(Nx^{-r}) = o(N^{-1})$ and hence (3.10) follows. In view of (1.1), (3.7), and (3.8), $\max_{-m+1\le i\le N} |X_i|$ and $\max_{1\le i\le N} |Z_i|$ are of the same order and therefore (3.11) also holds.

Using Chebyshev's inequality, (3.7) and (3.8), we obtain that $P(|\bar{X}| > N^{-\frac{1}{2}+\varepsilon}) = O(N^{-r\varepsilon})$ and thus

$$P(\bar{X}^2 > N^{-1+\varepsilon}) = O(N^{-r\varepsilon/2}) = o(N^{-1}), \tag{3.14}$$

for $\varepsilon > 2/r$. In the same way, we find that, for $\varepsilon > 2/r$, $P(|(m+N)^{-1} \sum_{i=-m+1}^N X_i^2 - \sigma^2| > N^{-\frac{1}{2}+\varepsilon}) = O(N^{-r\varepsilon/2}) = o(N^{-1})$. As $\sigma^2 > 0$, these results and a similar one for $(m+N)^{-1} \sum_{i=-m+1}^{N-k} X_i X_{i+k}$ imply in view of (2.7) that $P(|\hat{\rho}_k - \rho_k| > N^{-\frac{1}{2}+\varepsilon}) = o(N^{-1})$ for $\varepsilon > 2/r$. Together with (2.8), this proves (3.12).

Finally, because of (3.9), we have that $(m+N)^{-1} \sum_{i=-m+1}^N (X_i^{(s)})^2 - (m+N)^{-1} \sum_{i=-m+1}^N X_i^2 = O(\{N^{-1+\delta} + N^{-\lambda}\} \max_{1\le i\le N} Z_i^2)$ for all $\delta > 0$ and λ arbitrarily large. A similar result holds for $(m+N)^{-1} \sum_{i=-m+1}^{N-k} X_i^{(s)} X_{i+k}^{(s)}$. Combining these facts with (3.10), (3.14), and (2.7), we obtain that $P(|\hat{\rho}_k^{(s)} - \hat{\rho}_k| > N^{-1+2\varepsilon}) = o(N^{-1})$ for $\varepsilon > 2/r$. This implies (3.13). □

Next we use the results of this lemma to obtain bounds for $|\hat{R}_i - R_i|$ and $|\hat{R}_i^{(s)} - \hat{R}_i|$. Let $u(x) = 1$ for $x \geqq 0$ and $u(x) = 0$ for $x < 0$. Then

$$\hat{R}_i - R_i = \sum_{j=1}^N \{u(|\hat{Z}_i| - |\hat{Z}_j|) - u(|Z_i| - |Z_j|)\}. \tag{3.15}$$

LEMMA 3.3. *Suppose that the X_i satisfy* (1.1) *and* (2.2). *Let $E|X_i|^r < \infty$ for some $r > 20$ and assume that g is bounded. Then, for every $\delta > 0$, there exists $\varepsilon < 1/10$ such that*

$$P(\max_{1\le i\le N} |\hat{R}_i - R_i| > N^{\frac{1}{2}+2\varepsilon}) = o(N^{-1}), \tag{3.16}$$

$$E|\hat{R}_i^{(s)} - \hat{R}_i| = O(N^{3\varepsilon}), \tag{3.17}$$

uniformly in all i and s such that $i \geqq N^\delta$ and $|i-s| \geqq N^\delta$.

PROOF. As $|u(x+a)-u(x)| \leqq u(|a|-|x|)$, we have in view of (3.14) that $|\hat{R}_i - R_i| \leqq \sum_{j=1}^N u\{|\hat{Z}_i - Z_i| + |\hat{Z}_j - Z_j| - ||Z_i| - |Z_j||\}$. From $(\hat{Z}_i - Z_i) = \sum_{k=1}^m (\hat{a}_k - a_k)X_{i-k}$, together with (3.11) and (3.12), it then follows that there exists $\varepsilon_1 < 1/10$ such that

$$|\hat{R}_i - R_i| \leqq \sum_{j=1}^N u(2mN^{-\frac{1}{2}+2\varepsilon_1} - ||Z_i| - |Z_j||) , \tag{3.18}$$

except on a set of probability $o(N^{-1})$, uniformly in i. As g is bounded, $P(||Z_j| - |x|| \leqq 2mN^{-\frac{1}{2}+2\varepsilon_1}) = O(N^{-\frac{1}{2}+2\varepsilon_1})$, uniformly in x. Using a well-known bound for binomial probabilities (see, e.g., Okamoto (1958)), it follows that for $\varepsilon_1 < \varepsilon < 1/10$, given $|Z_i| = |z_i|$,

$$\begin{aligned} P(\{\# \text{ indices } j \text{ for which } ||Z_j| - |z_i|| \leqq 2mN^{-\frac{1}{2}+2\varepsilon_1}\} \geqq N^{\frac{1}{2}+2\varepsilon}) \\ \leqq \exp\{-2N[N^{-\frac{1}{2}+2\varepsilon} - O(N^{-\frac{1}{2}+2\varepsilon_1})]^2\} = o(N^{-2}) , \end{aligned} \tag{3.19}$$

uniformly in $|z_i|$. Together, (3.18) and (3.19) imply (3.16).

To prove (3.17), we first note that

$$\begin{aligned} E|\hat{R}_i^{(s)} - \hat{R}_i| &\leqq E \sum_{j=1}^N |u(|\hat{Z}_i^{(s)}| - |\hat{Z}_j^{(s)}|) - u(|\hat{Z}_i| - |\hat{Z}_j|)| \\ &\leqq \sum_{j=1}^N P(||\hat{Z}_i| - |\hat{Z}_j|| \leqq ||\hat{Z}_i^{(s)}| - |\hat{Z}_i|| + ||\hat{Z}_j^{(s)}| - |\hat{Z}_j||) . \end{aligned} \tag{3.20}$$

Now $||\hat{Z}_i^{(s)}| - |\hat{Z}_i|| = ||\sum_{k=0}^m \hat{a}_k^{(s)} X_{i-k}^{(s)}| - |\sum_{k=0}^m \hat{a}_k X_{i-k}|| \leqq |\sum_{k=1}^m (\hat{a}_k^{(s)} - \hat{a}_k)X_{i-k}^{(s)}| + \min\{|\sum_{k=0}^m \hat{a}_k(X_{i-k}^{(s)} - X_{i-k})|, |\sum_{k=0}^m \hat{a}_k(X_{i-k}^{(s)} + X_{i-k})|\}$. Using (3.13), (3.11), (3.9), and (3.10), respectively, we arrive at

$$P(||\hat{Z}_i^{(s)}| - |\hat{Z}_i|| \geqq mN^{-1+3\varepsilon} + mN^{-\lambda+\varepsilon}) = o(N^{-1}) , \tag{3.21}$$

for $i \geqq N^\delta$ and $|i - s| \geqq N^\delta$. In view of (3.20) and (3.21) it remains to show that $\sum_{j=1}^N P(||\hat{Z}_i| - |\hat{Z}_j|| \leqq 2(m+1)N^{1+3\varepsilon}) = O(N^{3\varepsilon})$.

To this end, we introduce some notation. Let $X_{i,p} = X_i - b_{i-p}Z_p$ for $i \geqq p$, then $X_{i,p}$ and Z_p are independent. Replace X_i in (2.7) by $X_{i,p}$ for $i \geqq p$ and denote the result by $\hat{\rho}_{k,p}$. Let $\hat{a}_{k,p}$ be the corresponding estimators obtained through (2.8). Then $\hat{a}_{k,p}$ and Z_p are independent and $|\hat{a}_{k,p} - \hat{a}_k| = O(N^{-1}|Z_p| \max_{1\leqq i\leqq N} |Z_i|)$. Obviously this procedure can be repeated to obtain $\hat{a}_{k,p,q}$, independent of Z_p and Z_q, and such that, uniformly in p and q,

$$|\hat{a}_{k,p,q} - \hat{a}_k| = O(N^{-1} \max_{1\leqq i\leqq N} Z_i^2) . \tag{3.22}$$

Furthermore, for every $\delta > 0$ and all $i \geqq N^\delta$, let $X_{i,\delta} = \sum_{k=0}^{N^\delta - m} b_k Z_{i-k}$. In view of (3.7) we then have, uniformly in i

$$|X_{i,\delta} - X_i| = O(N^{-\lambda} \max_{1\leqq i\leqq N}|Z_i|) . \tag{3.23}$$

Using (3.11), (3.22), and (3.23), we obtain that for $i, j \geqq N^\delta$ the following holds: $||\hat{Z}_i| - |\hat{Z}_j|| \geqq ||Z_i + \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{i-k,\delta}| - |Z_j + \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta}|| - O(N^{-1+3\varepsilon_2} + N^{-\lambda+\varepsilon_2})$, for some ε_2 such that $2/r < \varepsilon_2 < \varepsilon$, except on a set of probability $o(N^{-1})$. Hence it now remains to show that

$$\begin{aligned} \sum_{j=1}^N P(||Z_i + \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{i-k,\delta}| - |Z_j + \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta}|| \\ \leqq 2(m+2)N^{-1+3\varepsilon}) = O(N^{3\varepsilon}) . \end{aligned}$$

But that is easy: for $i, j \geqq N^\delta$, $|i-j| \geqq N^\delta$, both $\sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{i-k,\delta}$ and $\sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta}$ are independent of Z_i and Z_j and therefore the conditional probability

$$P(||Z_i + \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{i-k,\delta}| - |Z_j + \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta}|| \leqq 2(m+2)N^{-1+3\varepsilon} \mid \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{i-k,\delta} = x, \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta} = y)$$

equals $P(||Z_i + x| - |Z_j + y|| \leqq 2(m+2)N^{-1+3\varepsilon})$. Since Z_i and Z_j are independent and their density g is bounded, this latter probability is $O(N^{-1+3\varepsilon})$, uniformly in x and y and the desired result follows. □

We shall also use the results of Lemma 3.2 to obtain bounds for $|\text{sign } \hat{Z}_i - \text{sign } Z_i|$ and $|\text{sign } \hat{Z}_i^{(s)} - \text{sign } \{(i-s)\hat{Z}_i\}|$.

LEMMA 3.4. *Under the conditions of Lemma 3.3 we have that for every $\delta > 0$ there exists $\varepsilon < 1/10$ such that*

$$E|\text{sign } \hat{Z}_i - \text{sign } Z_i| = O(N^{-\frac{1}{2}+2\varepsilon}), \quad \textit{uniformly in } i, \tag{3.24}$$

$$E|\text{sign } \hat{Z}_i - \text{sign } Z_i||\text{sign } \hat{Z}_j^{(s)} - \text{sign } \{(j-s)\hat{Z}_j\}| = O(N^{-\frac{3}{2}+5\varepsilon}), \tag{3.25}$$

uniformly in i, j, and s such that $i, j \geqq N^\delta$ and $i \leqq s - N^\delta \leqq j - 2N^\delta$ or $i \geqq s + N^\delta \geqq j + 2N^\delta$.

PROOF. In view of (3.11), (3.12) and the boundedness of g, we have that $E|\text{sign } \hat{Z}_i - \text{sign } Z_i| \leqq 2P(|\hat{Z}_i - Z_i| \geqq |Z_i|) \leqq 2P(|Z_i| \leqq mN^{-\frac{1}{2}+2\varepsilon}) + o(N^{-1}) = O(N^{-\frac{1}{2}+2\varepsilon})$, for some $\varepsilon < 1/10$. Hence (3.24) follows.

To prove (3.25), we note in the first place that its left-hand side is at most $4P(|\hat{Z}_i - Z_i| \geqq |Z_i| \wedge |\hat{Z}_j^{(s)} - \hat{Z}_j \text{ sign }(j-s)| \geqq |\hat{Z}_j|)$. As $|\hat{Z}_j^{(s)} - \hat{Z}_j \text{ sign }(j-s)| \leqq |\sum_{k=0}^m (\hat{a}_k^{(s)} - \hat{a}_k)X_{j-k}^{(s)}| + |\sum_{k=0}^m \hat{a}_k(X_{j-k}^{(s)} - X_{j-k} \text{ sign }(j-s))|$, application of (3.11), (3.12), (3.13), (3.9), and (3.10), respectively, shows that there exists $\varepsilon < 1/10$ such that the expectation in (3.25) is less than or equal to

$$4P(|Z_i| \leqq mN^{-\frac{1}{2}+2\varepsilon} \wedge |\hat{Z}_j| \leqq (m+1)N^{-1+3\varepsilon}) + o(N^{-1}), \tag{3.26}$$

uniformly in the values of i, j, and s indicated above.

Next we use an argument similar to the one used at the end of the previous lemma. From (3.22) and (3.23) it follows that $|\hat{Z}_j| \geqq |Z_j + \sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta}| - O(N^{-1+3\varepsilon_1} + N^{-\lambda+\varepsilon_1})$, except on a set of probability $o(N^{-1})$, for some $2/r < \varepsilon_1 < \varepsilon$. Moreover Z_i, Z_j and $\sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta}$ are mutually independent for $i, j \geqq N^\delta$, $|i-j| \geqq N^\delta$. Hence the expression in (3.26) is at most

$$\begin{aligned} &4P(|Z_i| \leqq mN^{-\frac{1}{2}+2\varepsilon}) \times P(|Z_j + \textstyle\sum_{k=1}^m (\hat{a}_{k,i,j} - a_k)X_{j-k,\delta}| \leqq (m+2)N^{-1+3\varepsilon}) + o(N^{-1}) \\ &\quad \leqq O(N^{-\frac{1}{2}+2\varepsilon}) \sup_x \{P(|Z_j + x| \leqq (m+2)N^{-1+3\varepsilon})\} + o(N^{-1}) = O(N^{-\frac{3}{2}+5\varepsilon}). \quad \square \end{aligned}$$

Using Lemmas 3.3 and 3.4 we can now finally show:

LEMMA 3.5. *Suppose that the conditions of Theorem 3.1 are satisfied. Then* (3.2) *holds.*

PROOF. In view of (3.3) it suffices to show that $ET_p^2 = o(N)$, $p = 1, 2, 3$.

First we consider $ET_1^2 = E\sum_{i=1}^N \sum_{j=1}^N A_{ij}$, where

$$A_{ij} = \left\{J\left(\frac{\hat{R}_i}{N+1}\right) - J\left(\frac{R_i}{N+1}\right)\right\}\left\{J\left(\frac{\hat{R}_j}{N+1}\right) - J\left(\frac{R_j}{N+1}\right)\right\} \operatorname{sign} Z_i Z_j .$$

Let $\delta > 0$ be arbitrarily small and let Σ' denote summation over the indices i and j such that $i, j \geqq N^\delta$, $|i-j| \geqq 2N^\delta$. Let Σ'' denote summation over the remaining indices, i.e., $T_1^2 = (\Sigma' + \Sigma'')A_{ij}$. For the terms in Σ'' we simply use the boundedness of J' together with (3.16): $|EA_{ij}| = O(N^{-2}E|\hat{R}_i - R_i||\hat{R}_j - R_j|) = O(N^{-1+4\varepsilon}) + o(N^{-1}) = O(N^{-1+4\varepsilon})$, for some $\varepsilon < 1/10$. As there are only $O(N^{1+\delta})$ terms in Σ'',

$$\Sigma'' EA_{ij} = O(N^{4\varepsilon+\delta}) = o(N) . \tag{3.27}$$

For the terms in Σ', we use (3.4), (3.5) and the fact that $|Z_1^{(s)}|, \cdots, |Z_N^{(s)}|$ have the same ranks $R_1, \cdots, R_N$ for all s:

$$\begin{aligned} EA_{ij} = \tfrac{1}{2}E\Bigg\{\Bigg[&\left\{J\left(\frac{\hat{R}_i}{N+1}\right) - J\left(\frac{R_i}{N+1}\right)\right\}\left\{J\left(\frac{\hat{R}_j}{N+1}\right) - J\left(\frac{R_j}{N+1}\right)\right\} \\ &- \left\{J\left(\frac{\hat{R}_i^{(s)}}{N+1}\right) - J\left(\frac{R_i}{N+1}\right)\right\} \\ &\times \left\{J\left(\frac{\hat{R}_j^{(s)}}{N+1}\right) - J\left(\frac{R_j}{N+1}\right)\right\}\Bigg] \operatorname{sign}\{Z_i Z_j\}\Bigg\} , \end{aligned}$$

where s is chosen between i and j such that $\min(|i-s|, |j-s|) \geqq N^\delta$. Using the relation $|(c_i - d_i)(c_j - d_j) - (b_i - d_i)(b_j - d_j)| \leqq |(c_i - b_i)(c_j - d_j) + (b_i - d_i)(c_j - b_j)|$ and the boundednes of J', we arrive at

$$|EA_{ij}| = O(N^{-2}E\{|\hat{R}_i^{(s)} - \hat{R}_i||\hat{R}_j^{(s)} - R_j| + |\hat{R}_i - R_i||\hat{R}_j^{(s)} - \hat{R}_j|\}) . \tag{3.28}$$

From (3.4) and (3.5) it also follows that $E|\hat{R}_i^{(s)} - \hat{R}_i||\hat{R}_j^{(s)} - R_j| = E|\hat{R}_i - \hat{R}_i^{(s)}||\hat{R}_j - R_j|$. Hence we can apply (3.16) and (3.17) to both terms in (3.28), thus obtaining that $|EA_{ij}| = O(N^{-\frac{3}{2}+5\varepsilon}) + o(N^{-1}) = o(N^{-1})$, uniformly for all terms in Σ'. Hence $\Sigma' EA_{ij} = o(N)$. Together with (3.27) this gives that $ET_1^2 = o(N)$.

By using Lemma 3.4 instead of Lemma 3.3, it can be similarly proved that $ET_2^2 = o(N)$. As concerns T_3, from (3.16) it follows that there exists $\varepsilon_1 < 1/10$ such that

$$ET_3^2 = O(N^{-1+4\varepsilon_1}E(\textstyle\sum_{i=1}^N |\operatorname{sign} \hat{Z}_i - \operatorname{sign} Z_i|)^2) + o(N) . \tag{3.29}$$

Performing similar steps as in (3.18) and (3.19), we find that there exist $\varepsilon_2 < \varepsilon_3 < 1/10$ such that $\sum_{i=1}^N |\operatorname{sign} \hat{Z}_i - \operatorname{sign} Z_i| \leqq 2\sum_{i=1}^N u(mN^{-\frac{1}{2}+2\varepsilon_2} - |Z_i|) \leqq 2N^{\frac{1}{2}+2\varepsilon_3}$, except on a set of probability $o(N^{-1})$. Combining this with (3.29) we arrive at $ET_3^2 = O(N^{4(\varepsilon_1+\varepsilon_3)}) + o(N) = o(N)$. □

The proof of Theorem 3.1 now follows from Lemmas 3.1 and 3.5.

REFERENCES

[1] ANDERSON, T. W. (1971). *The Statistical Analysis of Time Series.* Wiley, New York.
[2] CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[3] GASTWIRTH, J. L. (1971). On the sign test for symmetry. *J. Amer. Statist. Assoc.* **66** 821–824.

[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] HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[8] MODESTINO, J. W. (1969). Nonparametric and adaptive detecting on dependent data. Technical Report 27, Dept. of Electrical Engineering, Princeton Univ.

[9] OKAMOTO, M. (1958). Some inequalities relating to the partial sum of binomial probabilities. *Ann. Inst. Statist. Math.* **10** 29–35.

[10] SERFLING, R. J. (1968). The Wilcoxon two-sample statistic on strongly mixing processes. *Ann. Math. Statist.* **39** 1202–1209.

DEPARTMENT OF MATHEMATICS
TWENTE UNIVERSITY OF TECHNOLOGY
P.O. BOX 217
ENSCHEDE, THE NETHERLANDS