

THE EXTREMES OF A TRIANGULAR ARRAY OF NORMAL RANDOM VARIABLES

BY TAILEN HSING,¹ JÜRIG HÜSLER AND ROLF-DIETER REISS

Texas A & M University, University of Bern and University of Siegen

Consider a triangular array of stationary normal random variables $\{\xi_{n,i}, i \geq 0, n \geq 1\}$ such that $\{\xi_{n,i}, i \geq 0\}$ is a stationary normal sequence for each $n \geq 1$. Let $\rho_{n,j} = \text{corr}(\xi_{n,i}, \xi_{n,i+j})$. We show that if $(1 - \rho_{n,j}) \log n \rightarrow \delta_j \in (0, \infty)$ as $n \rightarrow \infty$ for some j , then the locations where the extreme values occur cluster, and if $\rho_{n,j}$ tends to 0 fast enough as $j \rightarrow \infty$ for fixed n , then $\{\xi_{n,i}, i \geq 0\}$ satisfies a certain weak dependence condition. Under the two conditions, it is possible to speak about an index which measures the degree of clustering. In practice, this viewpoint can provide a better approximation of the distributions of the maxima of weakly dependent normal random variables than what is directly guided by the asymptotic theory of Berman.

1. Introduction. The distributional behavior of the extreme values from a normal process has been studied extensively. An important result due to Berman (1964) states that for a stationary normal sequence $\{\xi_j\}$ with $\rho_j = \text{corr}(\xi_0, \xi_j)$, the condition

$$(1.1) \quad \rho_j \log j \rightarrow 0 \quad \text{as } j \rightarrow \infty$$

implies that the extremes of ξ_j , $1 \leq j \leq n$, have the same asymptotic distributions as those from a sample of n independent random variables [cf. Leadbetter, Lindgren and Rootzén (1983)]. Mittal and Ylvisaker (1975) showed that the condition (1.1) on ρ_j is nearly necessary for that conclusion.

In practice, an unfortunate limitation of Berman's result is that the speed at which the independence model becomes appropriate may be rather slow. This can be seen in a number of ways. For instance, suppose we generate a sample path of 5000 observations, displayed in Figure 1, from an AR(1) stationary normal sequence with AR coefficient 0.7. Note that we rescaled the time interval to be $(0, 1]$ and for clarity we plotted only those values greater than 1.4; a small bar is placed at i/n on the horizontal axis if the i th value is recorded. Even though the correlations of the sequence decrease geometrically, one observes the obvious tendency that large values occur in clusters, a serious violation of the independence model.

Received March 1994; revised July 1995.

¹Research supported by Office of Naval Research Grant N00014-92-J-1007, the Swiss National Science Foundation and the Alexander von Humboldt Foundation.

AMS 1991 subject classifications. 60F05, 60G10, 60G15.

Key words and phrases. Dependence, extremal index, time series, weak convergence.

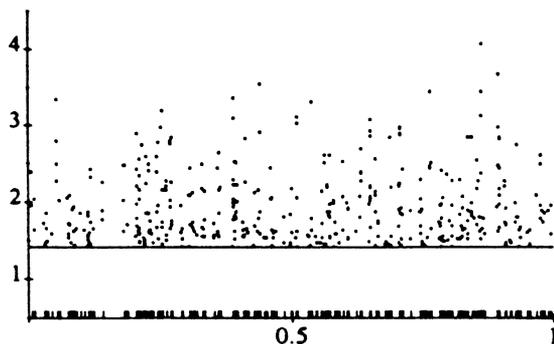


FIG. 1.

This is not surprising in view of a result in Rootzén (1983), which shows that for a stationary normal sequence $\{\xi_j\}$ satisfying (1.1), the convergence rate of

$$\sup_{u \geq v_n} \left| P \left\{ \bigvee_{i=1}^n \xi_i \leq u \right\} - \Phi^n(u) \right|$$

is roughly on the order of $n^{-(1-\rho)/(1+\rho)}$ for any v_n satisfying $n(1 - \Phi(v_n)) = O(1)$, where

$$\rho = \max(0, \rho_1, \rho_2, \dots).$$

Berman's theorem and its proof, based on the so-called normal comparison lemma, do not take clustering into account.

The purpose of this paper is then to develop an asymptotic theory for the extremes of a normal sequence which addresses the issue of clustering. The basic idea is essentially to magnify the correlations among neighboring observations so that the clustering information will be captured in the limit. Specifically, we use an embedding scheme explained as follows.

Consider a triangular array of normal random variables $\xi_{n,i}$, $i = 0, 1, 2, \dots$, $n = 1, 2, \dots$, such that for each n , $\{\xi_{n,i}, i \geq 0\}$ is a stationary normal sequence. Without loss of generality, assume that

$$E\xi_{n,i} = 0 \quad \text{and} \quad E\xi_{n,i}^2 = 1.$$

Suppose within each row a weak dependence condition such as Berman's condition holds so that, at least theoretically, the extremes of $\xi_{n,1}, \dots, \xi_{n,N}$ as $N \rightarrow \infty$ do not cluster. At the same time, however, as the row number n increases, allow some of the correlations among neighboring observations in $\{\xi_{n,i}\}_{i=0}^{\infty}$ to increase to 1 so that the local dependence cannot be ignored in the limit. Putting suitable restrictions on the correlations in this manner, one would be able to show, for example, that the point process with points i/n , $1 \leq i \leq n$, for which $\xi_{n,i} > v_n$, converges weakly as $n \rightarrow \infty$ to a compound Poisson process on the unit interval. Coming back to the sequence $\{\xi_j\}$, if the correlations of $\{\xi_j\}$ mimic those of some rows of such a triangular array, then

the clustering aspect in the asymptotic theory for the array should provide useful information about the clustering of the large values in ξ_1, \dots, ξ_n .

It is worth mentioning that the above idea is motivated by Hüsler and Reiss (1989), who considered the joint asymptotic distribution of component-wise maxima of normal random vectors. Also, related general discussions of the extremes of a triangular array of random variables can be found in Hüsler (1993), Nandagopalan (1990) and Nandagopalan, Leadbetter and Hüsler (1993).

Specifically, writing $\rho_{n,j} := E(\xi_{n,i} \xi_{n,i+j})$ and assuming that

$$(1.2) \quad (1 - \rho_{n,j}) \log n \rightarrow \text{some } \delta_j \in (0, \infty] \quad \text{for all } j \geq 1 \text{ as } n \rightarrow \infty$$

($\delta_0 = 0$) and for each n , $\rho_{n,j}$ decays fast enough as j increases, we show that one can speak about a cluster index which is similar to the notion of the extremal index (reviewed in Section 2) for stationary sequences. This covers, in particular, the case where the rows are identical and satisfy (1.1), and hence $\delta_j = \infty$ for each $j \geq 1$. As a result, Berman's theorem is covered as a special case (Remark 4 of Section 2). The main results of this paper are presented in Section 2, an example and some numerical comparisons are included in Section 3 and proofs are given in Section 4.

2. Main results. We continue to use the notation developed in Section 1. Fix $x \in (-\infty, \infty)$ and let

$$u_n(x) = x/a_n + b_n,$$

where

$$a_n = (2 \log n)^{1/2}$$

and

$$b_n = (2 \log n)^{1/2} - \frac{1}{2}(2 \log n)^{-1/2}(\log \log n + \log 4\pi).$$

It is well known that

$$\lim_{n \rightarrow \infty} n[1 - \Phi(u_n(x))] = e^{-x}, \quad -\infty < x < \infty.$$

THEOREM 2.1. *Assume that (1.2) holds. Also assume that there exist positive integers r_n, l_n such that*

$$(2.1) \quad \frac{l_n}{r_n} \rightarrow 0, \quad \frac{r_n}{n} \rightarrow 0$$

and for which

$$(2.2) \quad \lim_{n \rightarrow \infty} \frac{n^2}{r_n} \sum_{j=l_n}^n \frac{|\rho_{n,j}|}{(1 - \rho_{n,j}^2)^{1/2}} \exp\left\{-\frac{2 \log n - \log \log n}{1 + \rho_{n,j}}\right\} = 0$$

and

$$(2.3) \quad \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \sum_{j=m}^{r_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1-\rho_{n,j}^2)^{1/2}} = 0.$$

Then

$$\lim_{n \rightarrow \infty} P \left\{ \bigvee_{i=1}^n \xi_{n,i} \leq u_n(x) \right\} = \exp(-\vartheta \exp(-x)), \quad -\infty < x < \infty,$$

where

$$\vartheta = P \left\{ E/2 + \sqrt{\delta_k} W_k \leq \delta_k \text{ for all } k \geq 1 \text{ such that } \delta_k < \infty \right\}$$

($\vartheta = 1$ if $\delta_k = \infty$ for all $k \geq 1$), with E denoting a standard exponential random variable independent of the W_k and $\{W_k: \delta_k < \infty, k \geq 1\}$ being jointly normal with zero means and

$$(2.4) \quad EW_i W_j = \frac{\delta_i + \delta_j - \delta_{|i-j|}}{2(\delta_i \delta_j)^{1/2}}.$$

REMARK 1. By the triangle inequality,

$$\sqrt{E(\xi_{n,i} - \xi_{n,j})^2} \leq \sqrt{E(\xi_{n,0} - \xi_{n,i})^2} + \sqrt{E(\xi_{n,0} - \xi_{n,j})^2}.$$

Hence,

$$\sqrt{(1-\rho_{n,|i-j|}) \log n} \leq \sqrt{(1-\rho_{n,i}) \log n} + \sqrt{(1-\rho_{n,j}) \log n}$$

so that

$$\sqrt{\delta_{|i-j|}} \leq \sqrt{\delta_i} + \sqrt{\delta_j}.$$

Thus $\delta_{|i-j|}$ is finite if both δ_i and δ_j . Also we obtain the W_k in the proof through weak convergence and thus we need not worry about issues related to existence such as whether the correlations in (2.4) are nonnegative definite.

REMARK 2. It is completely straightforward to extend Theorem 2.1 to consider the weak convergence, as $n \rightarrow \infty$, of the point process containing those i/n , $1 \leq i \leq n$, for which $\xi_i > u_n(x)$. See Hsing, Hüsler and Leadbetter (1988).

REMARK 3. The conditions (2.2) and (2.3) are conditions which restrict, respectively, the global and local dependence of the random variables. They can be simplified in the direction of (1.1), as in Theorem 2.2.

THEOREM 2.2. Assume that (1.2) holds and that there exist positive integers l_n satisfying $l_n = o(n)$ and for which

$$(2.5) \quad \lim_{n \rightarrow \infty} \sup_{j \geq l_n} |\rho_{n,j}| \log n = 0$$

and

$$(2.6) \quad \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \sum_{j=m}^{l_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1-\rho_{n,j}^2)^{1/2}} = 0.$$

Then the conclusion of Theorem 2.1 holds.

REMARK 4. Suppose $\{\xi_j\}$ is a stationary sequence satisfying (1.1). Then, letting $\{\xi_{n,j}\} = \{\xi_j\}$ for every n , it is straightforward to show that (1.2) holds with $\delta_k = \infty$ for all $k \geq 1$, and (2.5) and (2.6) hold with $l_n = [n^\beta]$ for some $\beta > 0$. Thus Berman’s result is recovered in the present scheme.

There is a strong resemblance between the role of ϑ in the preceding results and that of the extremal index for a stationary sequence. Assume that $\{X_i\}$ is a strictly stationary sequence and $v_n(\tau)$ is such that

$$\lim_{n \rightarrow \infty} nP\{X_1 > v_n(\tau)\} = \tau, \quad \tau > 0,$$

The sequence $\{X_i\}$ is said to have extremal index θ if

$$\lim_{n \rightarrow \infty} P\left\{ \bigvee_{i=1}^n X_i \leq v_n(\tau) \right\} = e^{-\theta\tau}, \quad \tau > 0.$$

An extremal index is necessarily, and can, in fact, be any value, in $[0, 1]$. The closeness of θ to 0 and 1 signifies, respectively, strong and weak dependence. For details of the extremal index, see Leadbetter, Lindgren and Rootzén (1983) [cf. Hsing (1993)].

For convenience, we call the cluster index ϑ in the context of Theorems 2.1 and 2.2 the extremal index of the array $\{\xi_{n,i}\}$. The implication of the theoretical results is the following. Let $\{\xi_i\}$ be a stationary normal sequence with $\rho_j = \text{corr}(\xi_0, \xi_j)$ satisfying (1.1). Suppose for some fixed sample size n , we wish to find an approximate formula for $P\{\bigvee_{i=1}^n \xi_i \leq x\}$. The quantities

$$\bar{\delta}_k = \begin{cases} (1 - \rho_k) \log n, & \text{if } \rho_k > 0, \\ \infty, & \text{if } \rho_k \leq 0, \end{cases}$$

then play the role of the δ_k in (1.2) and our results suggest how to compute $\tilde{\vartheta}$ such that $P\{\bigvee_{i=1}^n \xi_i \leq x\}$ is better approximated by $\Phi(x)^{\tilde{\vartheta}n}$ than by $\Phi(x)^n$. However, care has to be taken in this practice since the quantities

$$\frac{\tilde{\delta}_i + \tilde{\delta}_j - \tilde{\delta}_{|i-j|}}{2(\tilde{\delta}_i \tilde{\delta}_j)^{1/2}}, \quad i, j \geq 1,$$

are no longer guaranteed to be nonnegative definite. It can be seen from the proof, in particular the derivations in (4.2), that a sensible choice of $\tilde{\vartheta}$ is

$$(2.7) \quad P\left\{ \rho_k E/2 + \sqrt{\tilde{\delta}_k} \tilde{W}_k \leq \tilde{\delta}_k \text{ for all } k \geq 1 \text{ such that } \tilde{\delta}_k < \infty \right\},$$

where E is standard exponential independent of the \tilde{W}_k and $\{\tilde{W}_k: \tilde{\delta}_k < \infty, k \geq 1\}$ is jointly normal with zero means and

$$E\tilde{W}_i\tilde{W}_j = \frac{\rho_{|i-j|} - \rho_i\rho_j}{2(1-\rho_i)^{1/2}(1-\rho_j)^{1/2}}.$$

By the monotone convergence theorem, the probability in (2.7) can be computed as a limit

$$\lim_{m \rightarrow \infty} P\left\{\rho_k E/2 + \sqrt{\tilde{\delta}_k} \tilde{W}_k \leq \tilde{\delta}_k \text{ for all } 1 \leq k \leq m \text{ such that } \tilde{\delta}_k < \infty\right\},$$

where, if $\rho_n \rightarrow 0$ fast, then one expects the speed of convergence of this probability to its limit to be fast also.

We do not evaluate such approximations theoretically. Instead, we illustrate numerically via an example in Section 3.

3. An example and numerical results. Consider a triangular array of Gaussian random variables $\{\xi_{n,j}\}$ with standard normal margins and such that for each n , $\{\xi_{n,j}\}_{j=0}^\infty$ is a stationary AR(1) process. That is,

$$\xi_{n,i} = d_n \xi_{n,i-1} + \sqrt{1-d_n^2} Z_i, \quad i \geq 0,$$

where the Z_i are iid standard normal. Assume that

$$(3.1) \quad d_n = 1 - \frac{\zeta}{\log n} \quad \text{for some } \zeta \in (0, \infty).$$

Clearly

$$\rho_{n,j} = d_n^j = \left(1 - \frac{\zeta}{\log n}\right)^j$$

and (1.2) holds with

$$\delta_j = j\zeta.$$

Now pick $l_n = (\log n)(\log(\log n))^2$. To verify (2.5), observe that

$$\sup_{j \geq l_n} |\rho_{n,j}| \log n \leq |\rho_{n,l_n}| \log n \leq \exp(-\zeta l_n / \log n + \log(\log n)) \rightarrow 0.$$

Now we verify (2.6). Fix any $\varepsilon \in (0, 2)$. If $j\zeta/\log n > \varepsilon$, then

$$\rho_{n,j} \leq e^{-\varepsilon}$$

and if $j\zeta/\log n \leq \varepsilon$, then by Taylor expansion we obtain

$$\frac{j\zeta}{\log n} \left(1 - \frac{\varepsilon}{2}\right) \leq 1 - \rho_{n,j} \leq \frac{j\zeta}{\log n}.$$

Thus,

$$\begin{aligned} & \sum_{j=m}^{l_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1-\rho_{n,j})^{1/2}} \\ &= \sum_{j=m}^{l_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{(1/2)(1-\rho_{n,j})/(1+\rho_{n,j})}}{((1-\rho_{n,j})\log n)^{1/2}} \\ &\leq \left(\sum_{j=m}^{l_n} n^{-(1-\exp(-\varepsilon))/(1+\exp(-\varepsilon))} \frac{(\log n)^{1/2}}{((1-\exp(-\varepsilon))\log n)^{1/2}} \right) \\ &\quad \vee \left(\sum_{j=m}^{l_n} \exp\left(-\frac{j\zeta}{2}\left(1-\frac{\varepsilon}{2}\right)\right) \frac{\exp[(j\zeta/2)(\log(\log n))/\log n]}{(j\zeta(1-\varepsilon/2))^{1/2}} \right). \end{aligned}$$

Hence (2.6) is readily verified and the conclusion of Theorem 2.2 holds.

Note that for this example, it is more natural to replace W_k by $k^{-1/2}\sum_{i=1}^k Z_i$ and represent ϑ by

$$(3.2) \quad \vartheta = P\left\{E/2 + \sqrt{\zeta} \sum_{i=1}^k Z_i \leq k\zeta \text{ for all } k \geq 1\right\},$$

where E denotes a standard exponential random variable independent of the Z_i and the Z_i are iid standard normal. Smith (1992) and Perfekt (1994) observed that for a class of weakly dependent stationary Markov chains, the extremal index can be expressed as the probability that a random walk process never up-crosses a threshold. From that perspective, it is intuitively clear why the representation in (3.2) arises.

We now present some numerical computations regarding the above example. We first compare the approximation of $P\{\bigvee_{1 \leq i \leq n} \xi_{n,i} \leq x\}$ by means of Φ^n and $\Phi^{\vartheta n}$, the latter being suggested by Theorem 2.2. The computation results are consistent with theory; namely, even though the ‘‘Berman-like’’ condition (2.5) holds, the increased correlation among neighboring random variables within the rows causes clustering and the extremal index ϑ has an important role to play in approximating the distribution of the maximum.

To get a better visual insight into the approximation, the approximating d.f. Φ^n will be replaced by the limiting standard Gumbel d.f. $G_3(x) = \exp(-e^{-x})$ by using the nonlinear transformation $T = G_3^{-1}(\Phi)$. Observe that

$$\sup_x \left| P\left\{\bigvee_{i=1}^n \xi_{n,i} \leq x\right\} - \Phi(x)^{\vartheta n} \right| = \sup_x |H_n(x) - G_{3, \log \vartheta}(x)|,$$

where $H_n(x) := P\{\bigvee_{i=1}^n \eta_{n,i} - \log n \leq x\}$, the marginal r.v.’s $\eta_{n,i} = T(\xi_{n,i})$ have the common d.f. G_3 and $G_{3, \mu}$ is the Gumbel d.f. with location parameter μ . Thus, if $\vartheta < 1$, the limiting d.f. is G_3 shifted to the left. Clearly, the preceding equality also holds in terms of the variational distance.

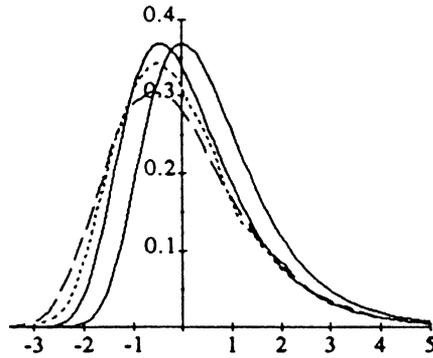


FIG. 2. Gumbel densities g_3 and $g_{3, -0.444}$ and kernel estimates of h_{100} (dotted) and $h_{10,000}$ (dashed).

For our Monte Carlo simulations, we use the random number generator of Turbo Pascal, version 4.0; normal pseudo-random numbers are obtained by applying the polar method. Moreover, the normal d.f. Φ is computed numerically by formula (26.2.19) in Abramowitz and Stegun (1965).

The following illustration gives a flavor of the approximation of H_n by means of G_3 and $G_{3, \log \vartheta}$. Taking $\zeta = 1$ we have $\vartheta = 0.641$ and $\log \vartheta = 0.444$. The density h_n of H_n is estimated by means of a kernel density based on $N = 8000$ random numbers generated according to H_n .

For increasing sample sizes n , the performance of H_n can be described in the following way: Clearly, $H_1 = G_3$. For $n \geq 2$, H_n drifts to the left with increasing variance, passing the target distribution $G_{3, -0.444}$ for $n = 10$ and attaining the maximum distance to $G_{3, -0.444}$ for n between 40 and 60. Finally, H_n approaches $G_{3, -0.444}$ in the manner as indicated in Figure 2 and Table 1.

Subsequently, the distance between distributions will be measured by the variational distance. Using the same symbols for d.f.'s and the corresponding probability measures, put

$$\Delta_B = \Delta_B(n) := \sup_A |H_n(A) - G_3(A)|$$

TABLE 1
Comparison under the extremal index $\vartheta = 0.641$

| | Δ_{HHR} | Δ_B | $\Delta_B / \Delta_{\text{HHR}}$ |
|--------------|-----------------------|------------|----------------------------------|
| $n = 20$ | 0.080 | 0.200 | 2.5 |
| $n = 100$ | 0.099 | 0.224 | 2.3 |
| $n = 1,000$ | 0.071 | 0.217 | 3.1 |
| $n = 10,000$ | 0.055 | 0.206 | 3.7 |
| $n = 50,000$ | 0.047 | 0.204 | 4.3 |

TABLE 2
Extremal indices $\tilde{\vartheta}$

| | $d = 0.5$ | $d = 0.6$ | $d = 0.7$ | $d = 0.8$ | $d = 0.9$ |
|--------------|-----------|-----------|-----------|-----------|-----------|
| $n = 20$ | 0.752 | 0.697 | 0.611 | 0.495 | 0.314 |
| $n = 100$ | 0.861 | 0.809 | 0.734 | 0.615 | 0.423 |
| $n = 1,000$ | 0.935 | 0.898 | 0.841 | 0.734 | 0.530 |
| $n = 10,000$ | 0.967 | 0.942 | 0.898 | 0.809 | 0.615 |
| $n = 50,000$ | 0.982 | 0.963 | 0.925 | 0.851 | 0.671 |

and

$$\Delta_{\text{HHR}} = \Delta_{\text{HHR}}(n, \vartheta) := \sup_A |H_n(A) - G_{3, \log \vartheta}(A)|.$$

Let again $d_n = 1 - \zeta/\log n$ with $\zeta = 1$ and hence $\vartheta = 0.641$ and $\log \vartheta = -0.444$. Keep in mind that the last digit of the computed values is uncertain due to the Monte Carlo simulations. We see that $G_{3, \log \vartheta}$ always provides a more accurate approximation to H_n than G_3 and, moreover, the accuracy improves as n increases.

Next consider a stationary AR(1) normal sequence $\{\xi_i\}$ with standard normal margins and $\text{cov}(\xi_0, \xi_j) = d^j$ for some (fixed) d in $(0, 1)$. With $\tilde{\zeta} = (1 - d)\log n$ and E and the Z_i as before, we computed

$$\tilde{\vartheta} = P\left\{E/2 + \sqrt{\tilde{\zeta}} \sum_{i=1}^k Z_i \leq k\tilde{\zeta} \text{ for all } k \geq 1\right\}$$

[cf. (3.1) and (3.2)] for various sample sizes n and recorded the results in Table 2. Clearly, for larger d , it takes a larger n for $\tilde{\vartheta}$ to approach 1.

Now for $d = 0.7$ and various values of n , we compared the approximation of $P\{\bigvee_{1 \leq i \leq n} \xi_i \leq x\}$ by means of Φ^n and $\Phi^{\tilde{\vartheta}n}$. We did so by computing the variational distances $\Delta_{\text{HHR}}(n, \tilde{\vartheta})$ and $\Delta_{\text{B}}(n)$, where Δ_{HHR} and Δ_{B} are modified from the above in an obvious way. The results again demonstrate that the approximations $\Phi^{\tilde{\vartheta}n}$ are indeed superior to Φ^n . Note that the last ratio was omitted because of the relative inaccuracy of the value $\Delta_{\text{HHR}} = 0.014$.

A referee suggested that we compare our approximation with the one described on the bottom of page 71 of Rootzén (1983). For the example considered above and in terms of the notation in this section, Rootzén's

TABLE 3
Comparison under the fixed correlation $d = 0.7$

| | $\log \tilde{\vartheta}$ | Δ_{HHR} | Δ_{B} | $\Delta_{\text{B}} / \Delta_{\text{HHR}}$ |
|--------------|--------------------------|-----------------------|---------------------|---|
| $n = 20$ | -0.493 | 0.092 | 0.230 | 2.5 |
| $n = 100$ | -0.309 | 0.066 | 0.161 | 2.4 |
| $n = 1,000$ | -0.173 | 0.035 | 0.085 | 2.4 |
| $n = 10,000$ | -0.108 | 0.014 | 0.045 | — |

method essentially suggests approximation of $P\{T(\bigvee_{1 \leq i \leq n} \xi_i) - \log n \leq x\}$ by

$$G_3(x) \left[1 + \frac{(1+d)^{3/2}}{(1-d)^{1/2}} (4\pi \log n)^{-(1-d)/(1+d)} n^{-(1-d)/(1+d)} \exp\left(-\frac{2+d}{1+d}x\right) \right],$$

as opposed to $G_3(x + \log \tilde{\theta}^{-1})$ suggests by our method. The qualities of the two approximations are very similar numerically for the AR(1) normal sequence. A vague connection between the two methods can be seen by making a Taylor expansion on $G_3(x + \log \tilde{\theta}^{-1})$, namely,

$$G_3(x + \log \tilde{\theta}^{-1}) \approx G_3(x) \left[1 + (\log \tilde{\theta}^{-1}) e^{-x} \right].$$

However, it is not possible to be more specific since Rootzén's method has an entirely different justification and does not take into account the correlations that are less than the maximum correlation.

The illustrations and the numerical computations were partially carried out with the statistical software system XTREMES [Hassman, Reiss and Thomas (1994)].

4. Proofs. In the following, assume that $x \in (-\infty, \infty)$ is fixed and write $u_n = u_n(x)$.

LEMMA 4.1. *Assume that (1.2) holds. Then for any bounded set $K \subset \{1, 2, \dots\}$.*

$$\lim_{n \rightarrow \infty} P\{\xi_{n,k} \leq u_n, k \in K \mid \xi_{n,0} > u_n\} = P\{E/2 + \sqrt{\delta_k} W_k \leq \delta_k, k \in \tilde{K}\},$$

where $\tilde{K} = \{k \in K: \delta_k < \infty\}$ (if \tilde{K} is the null set, then the probability on the right-hand side is 1), E is a standard exponential random variable independent of $(W_k, k \in \tilde{K})$ and the W_k have a jointly normal distribution with mean 0 and

$$EW_i W_j = \frac{\delta_i + \delta_j - \delta_{|i-j|}}{2(\delta_i \delta_j)^{1/2}}.$$

PROOF. Assume that $K = \tilde{K}$. The modification is obvious if this is not so. Straightforward computations give

$$\begin{aligned} & P\{\xi_{n,k} \leq u_n, k \in K \mid \xi_{n,0} > u_n\} \\ &= \int_0^\infty P\left\{\xi_{n,k} \leq u_n, k \in K \mid \xi_{n,0} = u_n + \frac{y}{u_n}\right\} \\ & \quad \times \frac{\phi(u_n + y/u_n)}{u_n(1 - \phi(u_n))} dy \\ (4.1) \quad & \sim \int_0^\infty P\left\{\xi_{n,k} \leq u_n, k \in K \mid \xi_{n,0} = u_n + \frac{y}{u_n}\right\} \\ & \quad \times \exp\left(-y - \frac{y^2}{2u_n^2}\right) dy. \end{aligned}$$

Let $(\eta_{n,k}, k \in K)$ have the conditional distribution of $(\xi_{n,k}, k \in K)$ given $\xi_{n,0} = u_n + y/u_n$. By Anderson [(1984), Theorem 2.5.1],

$$(\eta_{n,k}, k \in K)' \sim N(\Sigma_{12}(u_n + y/u_n), \Sigma_{11} - \Sigma_{12}\Sigma'_{12}),$$

where

$$\Sigma_{11} = (\rho_{n,|i-j|})_{i,j \in K} \quad \text{and} \quad \Sigma_{12} = (\rho_{n,k}, k \in K)'.$$

That is,

$$(\eta_{n,k}, k \in K)' \sim N(\mu, \Sigma),$$

where

$$\mu = (\rho_{n,k}(u_n + y/u_n), k \in K)' \quad \text{and} \quad \Sigma = (\rho_{n,|i-j|} - \rho_{n,i}\rho_{n,j})_{i,j \in K}.$$

Moreover, define

$$Z_{n,k} = \frac{\eta_{n,k} - \rho_{n,k}(u_n + y/u_n)}{(1 - \rho_{n,k}^2)^{1/2}}, \quad k \in K.$$

Hence

$$(Z_{n,k}, k \in K)' \sim N(0, \Sigma),$$

where

$$\Sigma = \left(\frac{\rho_{n,|i-j|} - \rho_{n,i}\rho_{n,j}}{(1 - \rho_{n,i}^2)^{1/2}(1 - \rho_{n,j}^2)^{1/2}} \right)_{i,j \in K}.$$

Thus,

$$\begin{aligned} & P\left\{ \xi_{n,k} \leq u_n, k \in K \mid \xi_{n,0} = u_n + \frac{y}{u_n} \right\} \\ &= P\{ \eta_{n,k} \leq u_n, k \in K \} \\ (4.2) \quad &= P\left\{ \frac{y\rho_{n,k}}{2} + \sqrt{\frac{1 + \rho_{n,k}}{2}} \sqrt{\frac{u_n^2(1 - \rho_{n,k})}{2}} Z_{n,k} \right. \\ &\quad \left. \leq \frac{u_n^2(1 - \rho_{n,k})}{2}, k \in K \right\}. \end{aligned}$$

Using the assumption (1.2) and the fact that $u_n \sim (2 \log n)^{1/2}$, it is easy to see that as $n \rightarrow \infty$,

$$\rho_{n,k} \rightarrow 1, \quad \frac{u_n^2(1 - \rho_{n,k})}{2} \sim (1 - \rho_{n,k})\log n \rightarrow \delta_k$$

and

$$EZ_{n,i}Z_{n,j} = \frac{\rho_{n,|i-j|} - \rho_{n,i}\rho_{n,j}}{(1 - \rho_{n,i}^2)^{1/2}(1 - \rho_{n,j}^2)^{1/2}} \rightarrow \frac{\delta_i + \delta_j - \delta_{|i-j|}}{2(\delta_i\delta_j)^{1/2}}, \quad i, j \in K.$$

Together with (4.2), we conclude that

$$\lim_{n \rightarrow \infty} P\{\xi_{n,k} \leq u_n, k \in K \mid \xi_{n,0} = u_n + y/u_n\} = P\{y/2 + \sqrt{\delta_k} W_k \leq \delta_k, k \in K\}.$$

The result follows readily from (4.1) using Lebesgue's dominated convergence theorem. \square

For $0 \leq l \leq n - 2$, define the mixing coefficient

$$\alpha_n(l) = \max\{|P\{\xi_{n,i} \leq u_n, i \in I \cup J\} - P\{\xi_{n,i} \leq u_n, i \in I\}P\{\xi_{n,i} \leq u_n, i \in J\}| : I, J \subset \{1, \dots, n\} \text{ and } \min\{|i - j| : i \in I, j \in J\} = l\}.$$

The following result is proved using a familiar argument, based on the so-called normal comparison lemma. See Berman (1964) and Leadbetter, Lindgren and Rootzén [(1983), pages 81–84].

LEMMA 4.2. For each $n, l \geq 1$,

$$\alpha_n(l) \leq \frac{n}{2\pi} \sum_{j=l}^n \frac{|\rho_{n,j}|}{(1 - \rho_{n,j}^2)^{1/2}} \exp\left(-\frac{u_n^2}{1 + \rho_{n,j}}\right).$$

PROOF OF THEOREM 2.1. Since

$$u_n^2 = 2 \log n - \log \log n + O(1),$$

Lemma 4.2 and (2.2) imply that

$$\lim_{n \rightarrow \infty} \frac{n}{r_n} \alpha_n(l_n) = 0.$$

Combining with the assumption (2.1), Theorem 21. of O'Brien (1987) [cf. Rootzén (1988)] implies that

$$P\left\{\bigvee_{i=1}^n X_i \leq u_n\right\} - \exp\left(-nP\{\xi_{n,0} > u_n\}P\left\{\bigcap_{j=1}^{r_n} (\xi_{n,j} \leq u_n) \mid \xi_{n,0} > u_n\right\}\right) \rightarrow 0.$$

It is clear that

$$\lim_{n \rightarrow \infty} nP\{\xi_{n,0} > u_n\} = e^{-x}$$

and it follows from Lemma 4.1 that

$$\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} P\left\{\bigcup_{j=1}^m (\xi_{n,j} \leq u_n) \mid \xi_{n,0} \geq u_n\right\} = \vartheta.$$

Therefore, the conclusion of the theorem follows if we have

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left\{\bigcup_{j=m}^{r_n} (\xi_{n,j} > u_n) \mid \xi_{n,0} > u_n\right\} = 0.$$

As in Lemma 4.1, write

$$P\left\{\bigcup_{j=m}^{r_n} (\xi_{n,j} > u_n) \mid \xi_{n,0} > u_n\right\} \\ \sim \int_{y=0}^{\infty} \exp(-y - y^2/2u_n^2) \\ \times P\left\{\bigcup_{j=m}^{r_n} \left(\rho_{n,j}(u_n + y/u_n) + Z_{n,j}\sqrt{1 - \rho_{n,j}^2} > u_n\right)\right\} dy.$$

It obviously suffices to show that for each fixed $y_0 > 0$,

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \int_{y=0}^{y_0} e^{-y} P\left\{\bigcup_{j=m}^{r_n} \left(\rho_{n,j}\left(u_n + \frac{y}{u_n}\right) + Z_{n,j}\sqrt{1 - \rho_{n,j}^2} > u_n\right)\right\} dy = 0,$$

or the stronger statement

$$(4.3) \quad \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \sup_{0 \leq y \leq j_0} \sum_{j=m}^{r_n} P\left\{\rho_{n,j}\left(u_n + \frac{y}{u_n}\right) + Z_{n,j}\sqrt{1 - \rho_{n,j}^2} > u_n\right\} = 0.$$

Next observe that

$$\sum_{j=m}^{r_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1 - \rho_{n,j}^2)^{1/2}} \\ = \sum_{j=m}^{r_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{(1/2)(1-\rho_{n,j})/(1+\rho_{n,j})}}{\sqrt{(1 + \rho_{n,j})(1 - \rho_{n,j})\log n}} \\ \geq \sum_{j=m}^{r_n} \frac{\exp(-(1 - \rho_{n,j})\log n)}{\sqrt{2(1 - \rho_{n,j})\log n}}.$$

Hence, if for some $\varepsilon > 0$,

$$\bigvee_{m \leq j \leq r_n} ((1 - \rho_{n,j})\log n)^{-1} > \varepsilon$$

then

$$\sum_{j=m}^{r_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1 - \rho_{n,j}^2)^{1/2}} \geq e^{-1/\varepsilon} \sqrt{\frac{\varepsilon}{2}}.$$

Thus, condition (2.3) implies that

$$(4.4) \quad \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \bigvee_{m \leq j \leq r_n} ((1 - \rho_{n,j})\log n)^{-1} = 0.$$

Now,

$$\begin{aligned}
 (4.5) \quad & P\left\{\rho_{n,j}\left(u_n + \frac{y}{u_n}\right) + Z_{n,j}\sqrt{1 - \rho_{n,j}^2} > u_n\right\} \\
 & = P\left\{Z > u_n\sqrt{\frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}} - \frac{y}{u_n}\frac{\rho_{n,j}}{\sqrt{1 - \rho_{n,j}^2}}\right\},
 \end{aligned}$$

where $Z \sim N(0, 1)$. By (4.4) for large n and $j \in [m, r_n]$,

$$u_n\sqrt{\frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}} - \frac{y}{u_n}\frac{\rho_{n,j}}{\sqrt{1 - \rho_{n,j}^2}} > 0.$$

By the well-known inequality,

$$1 - \Phi(x) \leq x^{-1}\phi(x), \quad x > 0,$$

we obtain

$$\begin{aligned}
 (4.6) \quad & P\left\{Z > u_n\sqrt{\frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}} - \frac{y}{u_n}\frac{\rho_{n,j}}{\sqrt{1 - \rho_{n,j}^2}}\right\} \\
 & \leq \left(u_n\sqrt{\frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}} - \frac{y}{u_n}\frac{\rho_{n,j}}{\sqrt{1 - \rho_{n,j}^2}}\right)^{-1} \\
 & \quad \times \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(u_n\sqrt{\frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}} - \frac{y}{u_n}\frac{\rho_{n,j}}{\sqrt{1 - \rho_{n,j}^2}}\right)^2\right\}.
 \end{aligned}$$

Now,

$$\begin{aligned}
 (4.7) \quad & \left(u_n\sqrt{\frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}} - \frac{y}{u_n}\frac{\rho_{n,j}}{\sqrt{1 - \rho_{n,j}^2}}\right)^2 \\
 & \geq C + \frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}u_n^2 \\
 & \geq C + \frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}(2 \log n - \log \log n)
 \end{aligned}$$

for some constant C depending only on x and y_0 . Here and hereafter, C denotes a generic constant whose value changes from line to line. By (4.6) and (4.7),

$$\begin{aligned}
 (4.8) \quad & P\left\{Z > u_n\sqrt{\frac{1 - \rho_{n,j}}{1 + \rho_{n,j}}} - \frac{y}{u_n}\frac{\rho_{n,j}}{\sqrt{1 - \rho_{n,j}^2}}\right\} \\
 & \leq Cn^{-(1 - \rho_{n,j})/(1 + \rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1 + \rho_{n,j})}}{\sqrt{1 - \rho_{n,j}^2}}
 \end{aligned}$$

for some constant C independent of n and j . By (4.5) and (4.8), (4.3) follows from (2.3). \square

PROOF OF THEOREM 2.2. Let $\varepsilon_n = \sup_{j \geq l_n} |\rho_{n,j}| \log n$ and

$$r_n = \left[n\sqrt{\varepsilon_n} \right] \vee \left[\sqrt{nl_n} \right].$$

Thus (2.1) is satisfied. For $j \geq l_n$, Taylor's theorem and condition (2.5) imply that for some constant C ,

$$\frac{2 \log n - \log \log n}{1 + \rho_{n,j}} \geq 2 \log n - \log \log n + C.$$

Thus,

$$\begin{aligned} \frac{n^2}{r_n} \sum_{j=l_n}^n \frac{|\rho_{n,j}|}{(1 - \rho_{n,j}^2)^{1/2}} \exp\left\{ -\frac{u_n^2}{1 + \rho_{n,j}} \right\} &\leq C \frac{n^3}{r_n} \frac{\varepsilon_n}{\log n} \exp(-2 \log n + \log \log n) \\ &= C \frac{n}{r_n} \varepsilon_n \rightarrow 0 \quad \text{as } n \rightarrow \infty, \end{aligned}$$

which proves (2.2). Now write

$$\sum_{j=m}^{r_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1 - \rho_{n,j})^{1/2}} = S_{n,1} + S_{n,2},$$

where

$$\begin{aligned} S_{n,1} &= \sum_{j=m}^{l_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1 - \rho_{n,j})^{1/2}}, \\ S_{n,2} &= \sum_{j=l_n+1}^{r_n} n^{-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1 - \rho_{n,j})^{1/2}}. \end{aligned}$$

By (2.5),

$$\sup_{j \geq l_n+1} n^{1-(1-\rho_{n,j})/(1+\rho_{n,j})} \frac{(\log n)^{-\rho_{n,j}/(1+\rho_{n,j})}}{(1 - \rho_{n,j})^{1/2}} \rightarrow 1,$$

and hence we have

$$\sup_{n \rightarrow \infty} S_{n,2} \leq \lim_{n \rightarrow \infty} r_n/n = 0.$$

Thus (2.3) follows from (2.6) and the result follows from Theorem 2.1. \square

Acknowledgment. The authors are grateful to the referees for their constructive comments.

REFERENCES

- ABRAMOWITZ, M. and STEGUN, I. A. (1965). *Handbook of Mathematical Functions*. Dover, New York.
- ANDERSON, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.
- BERMAN, S. (1964). Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35** 502–516.
- HASSMANN, S., REISS, R.-D. and THOMAS, M. (1994). XTREMES (statistical software and user's guide). In *DMV—Seminar on Laws of Small Numbers* (M. Falk, J. Hüslér and R.-D. Reiss, eds.), Appendix. Birkhäuser, Basel.
- HSING, T. (1993). Extremal index estimation for a weakly dependent stationary sequence. *Ann. Statist.* **21** 2043–2071.
- HSING, T., HÜSLER, J. and LEADBETTER, M. R. (1988). On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields* **78** 97–112.
- HÜSLER, J. (1993). A note on exceedances and rare events of nonstationary sequences. *J. Appl. Probab.* **30** 877–888.
- HÜSLER, J. and REISS, R.-D. (1989). Maxima of normal random vectors: between independence and complete dependence. *Statist. Probab. Lett.* **7** 283–286.
- LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, Berlin.
- MITTAL, Y. and YLVIKAKER, D. (1975). Limit distributions for the maximum of stationary Gaussian processes. *Stochastic Process. Appl.* **3** 1–18.
- NANDAGOPALAN, N. (1990). Multivariate extremes and estimation of the extremal index. Ph.D. dissertation, Dept. Statistics, Univ. North Carolina.
- NANDAGOPALAN, N., LEADBETTER, M. R. and HÜSLER, J. (1993). Limit theorems of nonstationary multi-dimensional strongly mixing random measures. Preprint.
- O'BRIEN, G. L. (1987). Extreme values for stationary and Markov sequences. *Ann. Probab.* **15** 281–289.
- PERFEKT, R. (1994). Extremal behavior of stationary Markov chains with applications. *Ann. Appl. Probab.* **4** 529–548.
- ROOTZÉN, H. (1983). The rate of convergence of extremes of stationary normal sequences. *Adv. in Appl. Probab.* **15** 54–80.
- ROOTZÉN, H. (1988). Maxima and exceedances of stationary Markov chains. *Adv. in Appl. Probab.* **20** 371–390.
- SMITH, R. L. (1992). The extremal index for a Markov chain. *J. Appl. Probab.* **29** 37–45.

T. HSING
DEPARTMENT OF STATISTICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843-3143
E-MAIL: thsing@stat.tamu.edu

J. HÜSLER
DEPARTMENT OF STATISTICS
UNIVERSITY OF BERN
BERN
SWITZERLAND
E-MAIL: huesler@math-stat.unibe.ch

R.-D. REISS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SIEGEN
SIEGEN
GERMANY
E-MAIL: reiss@xtremes.stat.math.uni-siegen.de