

*Research sponsored by Office of Naval Research under Contract N00014-67-A-0321-002 TSK NR-042-214.

ASPECTS OF EXTREME VALUE THEORY
FOR STATIONARY PROCESSES - A SURVEY

by
M. R. Leadbetter

*Department of Statistics
University of North Carolina at Chapel Hill*

Institute of Statistics Mimeo Series No. 951
September, 1974

ASPECTS OF EXTREME VALUE THEORY
FOR STATIONARY PROCESSES - A SURVEY

by

M. R. Leadbetter

*Department of Statistics
University of North Carolina at Chapel Hill*

SUMMARY

The primary concern in this paper is with the distributional results of classical extreme value theory, and their development to apply to stationary processes. The main emphasis is on stationary sequences, where the theory is well developed. Results available for continuous parameter processes are also described, but with particular reference to stationary normal processes.

Research sponsored by Office of Naval Research under Contract N00014-67-A-0321-002 TSK NR-042-214.

1. INTRODUCTION

This paper is concerned with certain aspects of extreme value theory of independent and identically distributed (i.i.d.) random variables, and their generalization to apply to stationary sequences, and continuous parameter processes. If $\{\xi_n\}$ is any sequence of random variables we shall write

$$M_n = \max(\xi_1, \xi_2, \dots, \xi_n)$$

for the maximum of the first n terms. Specifically the following topics from classical extreme value theory will be considered:

- (i) Gnedenko's Theorem - giving the possible types of asymptotic distribution of M_n , i.e. the possible distributions G such that
$$P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x)$$
- (ii) The relation $P\{M_n \leq u_n\} \rightarrow e^{-\tau}$ when $P(\xi_i > u_n) \sim \tau/n$
- (iii) The form of the limiting distribution in the *normal* case
- (iv) The Poisson limiting distribution for the number of exceedances of u_n by $\xi_1 \dots \xi_n$ where u_n is chosen as in (ii)
- (v) The asymptotic distribution of the r^{th} largest of $\xi_1 \dots \xi_n$ for fixed r .

In Section 2 these results will be stated (together with indications of methods of proof) for i.i.d. random variables. Section 3 contains a brief discussion of dependence conditions which have been used in the past, and

those which are used here, in order that the asymptotic results for the maximum may apply to stationary sequences, as will be seen in Section 4. These results will be compared with those of Berman [1] for stationary normal sequences.

In Section 5, we look at the "exceedances" of high levels by stationary sequences, and indicate their weak convergence as *point processes*, to a Poisson process, by means of a simple and useful general point process convergence theorem of Kallenberg [13]. The desired properties of the r^{th} largest value, $M_n^{(r)}$, will follow as immediate consequences of this convergence.

Certain of the results are known to apply to the supremum $M(T) = \sup\{\xi(t) : 0 \leq t \leq T\}$ of a *continuous* parameter stationary process $\xi(t)$ as $T \rightarrow \infty$. These are described in Section 6. The most detailed properties presently known concern the case when $\{\xi(t)\}$ is a *normal* process, and these are also discussed in Section 6.

Finally we remark that other questions besides distributional ones, are of interest in extreme value theory. For example convergence questions akin to laws of large numbers are clearly important. We do not consider such matters here, but concentrate on the distributional results indicated.

2. EXTREME VALUE THEORY FOR I.I.D. RANDOM VARIABLES

In this section we shall look briefly at each of the topics mentioned in the Introduction, for i.i.d. random variables.

(i) Gnedenko's Theorem.

This theorem is central to classical extreme value theory and is stated precisely as follows:

THEOREM 2.1 (GNEDENKO): Let $\xi_1, \xi_2 \dots$ be i.i.d. random variables, and let $M_n = \max(\xi_1, \xi_2 \dots \xi_n)$. Suppose that for some sequences of constants $\{a_n > 0\}$, $\{b_n\}$,

$$P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x) \quad (\text{at continuity points of } G)$$

where G is a non-degenerate d.f.. Then G belongs to one of the three "extreme value types":

$$\text{Type I } G(x) = \exp(-e^{-x}) \quad -\infty < x < \infty$$

$$\begin{aligned} \text{Type II } G(x) &= 0 \quad x < 0 \\ &= \exp(-x^{-\alpha}) \quad x \geq 0 \quad (\alpha > 0) \end{aligned}$$

$$\begin{aligned} \text{Type III } G(x) &= \exp(-(-x)^\alpha) \quad x < 0 \quad (\alpha > 0) \\ &= 1 \quad x \geq 0 \end{aligned}$$

(In saying that G is one of these "types" it is to be understood that x may be replaced by $ax + b$ for any fixed $a > 0$, b in the functional form.)

While we shall not prove Gnedenko's Theorem here (see [10] or [5] for detailed proof) it does seem worthwhile to point out the main ideas of the proof, since these are also used later. The main part of the proof given by Gnedenko may be displayed as the three lemmas given below.

First note that if the common distribution function (d.f.) of the $\{\xi_i\}$ is F , then M_n has the d.f. $P\{\xi_1 \leq x, \xi_2 \leq x \dots \xi_n \leq x\} = F^n(x)$ and

$$P\{a_n(M_n - b_n) \leq x\} = F^n\left(\frac{x}{a_n} + b_n\right).$$

If the left-hand side converges to some $G(x)$ then for any $k = 1, 2, 3, \dots$,

$$P\{a_{nk}(M_n - b_{nk}) \leq x\} = F^k\left(\frac{x}{a_{nk}} + b_{nk}\right) \rightarrow G^{1/k}(x).$$

Stating this formally we have

LEMMA 2.2 Let ξ_1, ξ_2, \dots be i.i.d. . If

$$(2.1) \quad P\{a_{nk}(M_n - b_{nk}) \leq x\} \rightarrow G^{1/k}(x) \text{ as } n \rightarrow \infty$$

holds for $k = 1$, then it holds for all $k = 1, 2, 3, \dots$.

In particular if the assumptions of Theorem 2.1 are satisfied then (2.1) holds for all k . On the other hand if (2.1) holds for all k , G must have certain "stability properties" indicated in the following lemma.

LEMMA 2.3 If (2.1) holds for all k and G is non-degenerate then for each k , there exist constants $\alpha_k > 0$, β_k such that

$$(2.2) \quad G^k(\alpha_k x + \beta_k) = G(x) \text{ for all } x$$

(Sometimes such a d.f. is called *stable* (e.g. [5]) though this is, of course, not the usual definition of stability.)

This result is due to Khintchine and is given in [11]. A very general result of this type (from which this lemma follows as a special case) is given

by Feller [7, p. 246]. This result states that if $\{F_n\}$ is any sequence of d.f.'s such that

$$F_n(\lambda_n x + \mu_n) \rightarrow G(x) , F_n(\lambda_n^* x + \mu_n^*) \rightarrow G^*(x) ,$$

for some sequences $\lambda_n > 0$, μ_n , $\lambda_n^* > 0$, μ_n^* , where G , G^* are assumed non-degenerate, then $\lambda_n/\lambda_n^* \rightarrow 1$, $(\mu_n - \mu_n^*)/\lambda_n \rightarrow 0$ and $G^*(x) = G(ax + b)$ for some $a > 0$, b . This theorem may be applied to (2.1) with $\lambda_n = 1/a_n$, $\mu_n = b_n$, $\lambda_n^* = 1/a_{nk}$, $\mu_n^* = 1/b_{nk}$, $G^* = G^{1/k}$, to give the lemma.

The following lemma now completes the proof of Gnedenko's Theorem for i.i.d. random variables.

LEMMA 2.4 If G is a non-degenerate d.f. satisfying (2.2) for each k (and some constants $\alpha_k > 0$, β_k) then it is an extreme value d.f. of Type I, II, or III.

The proof of this lemma is the major part of Gnedenko's derivation in [10] and we do not give it here.

To summarize, if the ξ_n are i.i.d. it follows that

- (a) If (2.1) holds for $k = 1$ it holds for all $k = 1, 2, \dots$. In particular (2.1) holds for all k under the assumptions of Gnedenko's Theorem.
- (b) If (2.1) holds for all k then G (if non-degenerate) has the stability properties (2.2),
- (c) If (2.2) holds for each k , G is a Type I, II, or III extreme value d.f. and hence

(d) under the assumptions of Gnedenko's Theorem, (2.1) holds and hence (2.2) does and hence G is an extreme value d.f. .

Looking ahead, we note that the same proof will apply under any circumstances in which the truth of (2.1) for $k = 1$ implies its truth for all k . This will be used in the next section to obtain the desired generalization of Gnedenko's Theorem.

(ii) Convergence of $P\{M_n \leq u_n\}$ to $e^{-\tau}$ when $1 - F(u_n) \sim \tau/n$.

Suppose that a sequence $\{u_n\}$ may be chosen so that

$$(2.3) \quad 1 - F(u_n) \quad (= P\{\xi_1 > u_n\}) \sim \tau/n$$

for some fixed $\tau > 0$. (Note that this is not always possible - e.g. if F_1 increases only by jumps at $1, 2, 3, \dots$ with $F(j) = 1 - \tau/(2^j - 1)$, as may be checked.) Then

$$(2.4) \quad P\{M_n \leq u_n\} = [1 - (1 - F(u_n))]^n = [1 - \tau/n + o(1/n)]^n \rightarrow e^{-\tau} .$$

Conversely it is easily seen that (2.3) is also *necessary* for (2.4). (If (2.4) holds then

$$n \log[1 - (1 - F(u_n))] \rightarrow -\tau$$

from which (2.3) follows.) Thus the following result holds.

THEOREM 2.5. For the i.i.d. case $P\{M_n \leq u_n\} \rightarrow e^{-\tau}$

($\tau > 0$) if and only if $1 - F(u_n) \sim \tau/n$.

(iii) The limiting distribution in the normal case.

If ξ_i are i.i.d. standard normal r.v.'s we may choose u_n so that $1 - \Phi(u_n) = \tau/n$ (for a given $\tau > 0$) where Φ is the standard normal d.f. . Now $1 - \Phi(x) \sim \phi(x)/x$ as $x \rightarrow \infty$ where ϕ is the standard normal density so that $n\phi(u_n)/u_n \rightarrow \tau$. It follows by taking logs and rearranging that

$$(2.5) \quad u_n^2 = 2 \log n - \log 2\pi - 2 \log u_n - 2 \log \tau$$

and dividing by $\log u_n$ shows that $\log u_n = o(\log n)$. It then follows from (2.5) that $u_n^2 \sim 2 \log n$ or $u_n^2/(2 \log n) \rightarrow 1$, whence by taking logs we have

$$(2.6) \quad \log u_n = \frac{1}{2} \log 2 + \frac{1}{2} \log \log n + o(1).$$

By substituting (2.6) in (2.5) and taking the square root we may obtain

$$\begin{aligned} u_n &= (2 \log n)^{\frac{1}{2}} \left[1 - \frac{1}{2} \frac{\log \log n + \log 4\pi}{2 \log n} + \frac{\log \tau}{2 \log n} + o\left(\frac{\log \log n}{\log n}\right)^2 \right] \\ &= b_n - \frac{\log \tau}{a_n} + o\left(\frac{1}{a_n}\right) \end{aligned}$$

where

$$(2.7) \quad \begin{cases} a_n = (2 \log n)^{\frac{1}{2}} \\ b_n = (2 \log n)^{\frac{1}{2}} - \frac{1}{2} (2 \log n)^{-\frac{1}{2}} (\log \log n + \log 4\pi) . \end{cases}$$

Hence, writing $\tau = e^{-x}$ in Lemma 2.5, we have

$$P\{a_n(M_n - b_n) + o(1) \leq x\} \rightarrow \exp(-e^{-x})$$

so that also $P\{a_n(M_n - b_n) \leq x\} \rightarrow \exp(-e^{-x})$.

Thus in the normal case the asymptotic distribution for M is of Type I with the normalizing constants a_n, b_n given by (2.7).

(iv) Limiting Poisson distribution of Exceedances of u_n .

Let ξ_1, ξ_2, \dots be i.i.d. and suppose $\{u_n\}$ may be chosen to satisfy (2.3). Any j for which $\xi_j > u_n$ will be called an *exceedance* of u_n by $\{\xi_k\}$. The number N_n of exceedances of u_n by $\{\xi_k\}$ for $1 \leq k \leq n$ is binomial with parameters $n, 1 - F(u_n)$. Since $n(1 - F(u_n)) \rightarrow \tau$ it follows that N_n is asymptotically Poisson, i.e. the following result holds.

THEOREM 2.6. If N_n is the number of exceedances of u_n (satisfying (2.3)) by the i.i.d. random variables $\xi_1, \xi_2, \dots, \xi_n$, then

$$(2.8) \quad P\{N_n = r\} \rightarrow e^{-\tau} \tau^r / r!$$

for any fixed $r = 1, 2, 3, \dots$.

Thus the exceedances have this "asymptotic Poisson character" whatever the form of F (provided (2.3) can be satisfied). This notion will be developed much more in Section 5 where the exceedances will be regarded as a point process.

(v) Asymptotic distribution of $M_n^{(r)}$, the r^{th} largest of $\xi_1 \dots \xi_n$.

Let $\{u_n\}$ be a sequence such that (2.3) holds, and hence by Theorem 2.6, (2.8) holds. But the event $\{M_n^{(r)} \leq u_n\}$ is precisely the event $\{N_n < r\}$ so that $P\{M_n^{(r)} \leq u_n\} = \sum_{s=0}^{r-1} P(N_n = s) \rightarrow e^{-\tau} \sum_{s=0}^{r-1} \tau^s/s!$. Summarizing this formally we have the following result.

THEOREM 2.7. If $\{u_n\}$ satisfies (2.3) and $M_n^{(r)}$ is the r^{th} largest of the i.i.d. random variables $\xi_1, \xi_2 \dots \xi_n$ then

$$(2.9) \quad P\{M_n^{(r)} \leq u_n\} \rightarrow e^{-\tau} \sum_{s=0}^{r-1} \tau^s/s! \quad \text{as } n \rightarrow \infty,$$

for all $r = 1, 2, 3, \dots$

COROLLARY: If $\{u_n\}$ is a sequence such that $P\{M_n \leq u_n\} \rightarrow e^{-\tau}$ then (2.9) holds. (For then by Theorem 2.5, (2.3) holds.)

We may develop this line a little further - to obtain the asymptotic distribution of $M_n^{(r)}$ from that of M_n . Specifically if $P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x)$ where G is non-degenerate (and hence of Type I, II, or III) then, if $0 < G(x) < 1$, by applying the above corollary with $u_n = x/a_n + b_n$, $\tau = -\log G(x)$ we see that

$$(2.10) \quad P\{a_n(M_n^{(r)} - b_n) \leq x\} \rightarrow G(x) \sum_{s=0}^{r-1} \left(-\log G(x)\right)^s/s!.$$

It is also easily seen (using continuity of the extreme value distribution G) that the limit is zero if $G(x) = 0$ and 1 if $G(x) = 1$. Thus the following result holds.

THEOREM 2.8. Suppose that $P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x)$ is $n \rightarrow \infty$ where G is a non-degenerate d.f. . Then the asymptotic distribution of $M_n^{(r)}$ the r^{th} largest of $\xi_1, \xi_2, \dots, \xi_n$ is given by (2.10). (If $G(x) = 0$ or 1 the right-hand side of (2.10) is to be taken as 0 or 1 respectively.)

3. DEPENDENCE RESTRICTIONS.

We turn now to consider *dependent* stochastic sequences ξ_1, ξ_2, \dots . Although it is not always necessary to do so (cf. [8]), we shall concern ourselves with strictly *stationary* sequences, i.e. such that the finite dimensional distributions $F_{i_1 \dots i_n}(x_1 \dots x_n)$ of $\xi_{i_1} \dots \xi_{i_n}$ have the property that $F_{i_1+p, \dots, i_n+p}(x_1 \dots x_n) = F_{i_1 \dots i_n}(x_1 \dots x_n)$ for any choice of $n, i_1 \dots i_n, x_1 \dots x_n$.

Certain of the properties of Section 2 were generalized by G. S. Watson ([25]) to apply to "m-dependent" stationary sequences, and by P. M. Loynes ([19]) to "strongly mixing" sequences. The strong-mixing assumption is the usual one, i.e. requiring that

$$(3.1) \quad |P(A \cap B) - P(A)P(B)| < g(k)$$

for any events $A \in \sigma(\xi_1 \dots \xi_m)$ $B \in \sigma(\xi_{m+k+1}, \xi_{m+k+2}, \dots)$ ($\sigma(\xi_1 \dots \xi_m)$ denoting the σ -field generated by $\xi_1 \dots \xi_m$ etc.) for any m , where $g(k)$ (the "mixing function") tends to zero as $k \rightarrow \infty$. (m-dependence requires $g(k) = 0$ for $k \geq m$.)

In particular Loynes obtained Gnedenko's Theorem under strong mixing, and showed that (2.4) holds under natural further conditions. In addition he

showed that ~ under such conditions - the asymptotic distribution of $M_n = \max(\xi_1 \dots \xi_n)$ is the same as it would be if the ξ_i were i.i.d. with the same marginal d.f. as when dependent.

When the sequence is strongly mixing but does not necessarily satisfy further conditions, it is possible to obtain an asymptotic distribution of M_n which is different from that of the i.i.d. sequence with the same marginal d.f. Examples of this - and other related matters complementing the work of Loynes, are given in papers by O'Brien ([20], [21]).

For stationary *normal* sequences, conditions on the covariance function $r_n = E\xi_i \xi_{i+n}$ (taking $E\xi_i = 0$, $E\xi_i^2 = 1$) are more natural than mixing conditions. It has been shown by S. M. Berman ([1]) that the same asymptotic limit holds for M_n in this case as for an i.i.d. normal sequence (i.e. a Type I limit with normalizing constants given by (2.7)) if one of the following conditions holds.

$$(3.2) \quad r_n \log n \rightarrow 0 \text{ as } n \rightarrow \infty \text{ or } \sum_{n=1}^{\infty} r_n^2 < \infty .$$

(Mittal and Ylvisaker ([27]) have recently shown that if the conditions (3.2) do *not* hold a variety of limiting distributions are possible.

It is possible to weaken the strong mixing condition and still obtain Gnedenko's Theorem and the other results indicated above. It will further turn out that this weakening provides results which apply to normal sequences under the conditions (3.2). To see how the strong mixing condition should be weakened, note first that not all events A, B in (3.1) are of interest. In

fact we are clearly primarily interested in events of the form

$\{\xi_1 \leq u, \xi_2 \leq u, \dots, \xi_n \leq u\}$ which suggests requiring a condition of the type (2.1) to hold for all $A = \{\xi_{i_1} \leq u \dots \xi_{i_p} \leq u\}$ $B = \{\xi_{j_1} \leq u \dots \xi_{j_q} \leq u\}$

where $i_1 < i_2 \dots < i_p < j_1 < j_2 \dots < j_q$ and $j_1 - i_p \geq k$. Put in terms of the finite dimensional distribution functions (and writing $F_{i_1 \dots i_n}(u)$

for $F_{i_1 \dots i_n}(u, u \dots u)$) this would require

$$(D): \quad |F_{i_1 \dots i_p j_1 \dots j_q}(u) - F_{i_1 \dots i_p}(u)F_{j_1 \dots j_q}(u)| < g(k)$$

whenever $i_1 < i_2 \dots < i_p < j_1 \dots < j_q$, $j_1 - i_p \geq k$, $g(k) \rightarrow 0$ as $k \rightarrow \infty$.

However even (D) is not sufficiently weak to deal with *normal* sequences satisfying the weak covariance conditions (3.2). We may obtain a suitable condition

by modifying (D) so that it is essentially required to hold for a single

sequence of interest, rather than for every fixed value of u . More precisely

if (u_n) denotes a given fixed sequence we shall say that $D(u_n)$ holds if

$$(3.3) \quad (D(u_n)) \quad |F_{i_1 \dots i_p j_1 \dots j_q}(u_n) - F_{i_1 \dots i_p}(u_n)F_{j_1 \dots j_q}(u_n)| < \alpha_{n,k}$$

whenever $i_1 < i_2 \dots < i_p < j_1 \dots < j_q$, $j_1 - i_p \geq k$, where $\alpha_{n,k}$ is non-increasing in k and where

$$(3.4) \quad \lim_{n \rightarrow \infty} \alpha_{n, k_n} = 0$$

for some sequence $k_n \rightarrow \infty$ with $k_n/n \rightarrow 0$.

It is apparent that (D) implies $(D(u_n))$ for any sequence u_n (any sequence $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ may be chosen). However it will turn out that $D(u_n)$ is implied by either condition (3.2) in the normal case, for suitably

chosen u_n .

We shall see that Gnedenko's Theorem holds under conditions of the type $D(u_n)$. However (as for strong mixing) an additional condition is required to obtain results such as (2.4) and the limit law for normal sequences. The following condition $D'(u_n)$ is a variant of one used by Watson [25] and Loynes [19] and will be appropriate here. Again let $[u_n]$ be a given real sequence. Then $D'(u_n)$ holds if

$$(3.5) (D'(u_n)) : \limsup_{n \rightarrow \infty} n \sum_{j=2}^n P\{\xi_1 > u_{nk}, \xi_j > u_{nk}\} = o(1/k) \text{ as } k \rightarrow \infty .$$

This condition involves the *bivariate* distribution of the $\{\xi_j\}$ and is easily seen to hold for example if the ξ_j are i.i.d. and u_n satisfies (2.3). It also holds (Section 6) if the ξ_j form a stationary normal sequence satisfying either condition (3.2), again if $\{u_n\}$ satisfies (2.3).

4. EXTREME VALUE THEORY FOR STATIONARY SEQUENCES

(i) Gnedenko's Theorem under $D(u_n)$.

The following generalization of Gnedenko's Theorem holds.

THEOREM 4.1. If $\{\xi_n\}$ is a stationary sequence, $M_n = \max(\xi_1 \dots \xi_n)$ having non-degenerate asymptotic distribution,

$$P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x) \quad (a_n > 0)$$

and if $D(u_n)$ holds for $u_n = x/a_n + b_n$ (for each real x), then $G(x)$ is a Type I, II, or III extreme value distribution function.

The proof may be found in detail in [15]. Its main pattern is that used by Loynes [19] for the mixing case with, however, some essential differences. It will be useful to indicate the general lines of proof here, without detail.

As for the i.i.d. case, it is only necessary to show that if (2.1) holds for $k = 1$, then it holds for all $k = 1, 2, 3, \dots$. The following basic lemma uses $D(u_n)$ to "approximate by independence". In this $M(E)$ will denote $\max(\xi_j : j \in E)$ for any set E of integers, and a set $(i, i + 1 \dots j)$ will be called an *interval*.

LEMMA 4.2. Let N, r, k be fixed and $D(u_n)$ hold for a given sequence (u_n) . Let E_j be intervals, $j = 1 \dots r$, separated by at least k from each other. Then

$$\left| P\left\{ \bigcap_{j=1}^r M(E_j) \leq u_N \right\} - \prod_{j=1}^r P\{M(E_j) \leq u_N\} \right| \leq (r - 1)\alpha_{N,k} :$$

Lemma 4.2 is proved by a simple induction and is used in obtaining the following result:

LEMMA 4.3. Let n, m, k be positive integers. If $D(u_n)$ holds then writing $N = nk$, $I_1 = (1, 2 \dots (n - m))$, $I_1^* = (n - m + 1, \dots, n)$,

$$\left| P\{M_N \leq u_N\} - P^k\{M_n \leq u_N\} \right| \leq (k + K)P\{M(I_1) \leq u_N \leq M(I_1^*)\}$$

for some constant K .

This is proved (along the lines used by Loynes) by dividing the first $N = nk$ integers into intervals $I_1, I_1^*, I_2, I_2^*, \dots, I_k, I_k^*$ of "lengths" $n - m, m, n - m, m \dots$. $P\{M_N \leq u_N\}$ is then approximated by $P\left\{ \bigcap_{j=1}^k M(I_j) \leq u_N \right\}$

which in turn is (by Lemma 4.2) approximated by $\prod_{j=1}^k P\{M(I_j) \leq u_N\}$. Each term is finally approximated by $P\{M_n \leq u_N\}$.

From this lemma it can be shown that $P\{M_N \leq u_N\} - P^k\{M_n \leq u_N\} \rightarrow 0$ for each k . If $D(u_n)$ holds with $u_n = \frac{x}{a_n} + b_n$ and $P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x)$ (i.e. if (2.1) holds for $k = 1$) then $P\{M_N \leq u_N\} \rightarrow G(x)$ and hence $P^k\{M_n \leq u_{nk}\} \rightarrow G(x)$ or $P\{a_{nk}(M_n - b_{nk}) \leq x\} \rightarrow G^{1/k}(x)$ for all $k = 1, 2, \dots$ (i.e. 2.1 holds for all $k = 1, 2, \dots$) as required to finish the proof of Theorem 4.1.

(ii) Convergence of $P\{M_n \leq u_n\}$ under (2.3), $D(u_n)$, $D'(u_n)$.

It is shown in [19] that under strong mixing if $P\{M_n \leq u_n(\tau)\} \rightarrow \psi(\tau)$ when $u_n(\tau)$ satisfies (2.3) for each τ then $\psi(\tau) = e^{-\alpha\tau}$ for some α with $0 \leq \alpha \leq 1$. It may similarly be shown that this applies if, instead of strong mixing, the conditions $D(u_n(\tau))$ hold for each τ . (This is seen simply by using the fact that $[1 - F(u_{2n}(\tau))] \sim (\tau/2)/n$ so that $P\{M_n \leq u_{2n}(\tau)\} \rightarrow \psi(\tau/2)$ and thus, by Lemma 2.5 of [15], $\psi(\tau) = \psi^2(\tau/2)$ from which the exponential limit may be deduced as in [19]).

O'Brien ([20]) gives examples in which $\alpha < 1$. We are here particularly interested in the "usual" case $\alpha = 1$ which may be guaranteed by the condition $D'(u_n)$ of the last section. Specifically we obtain the following result, again generalizing one given in [19] under strong mixing. It should be noted that the limit is shown to exist under the stronger conditions (rather than *assumed* to exist and then shown to have a given form as above, when just $D(u_n)$ holds).

THEOREM 4.4. Let $\{\xi_n\}$ be a stationary sequence and let $D(u_n)$, $D'(u_n)$ hold, where $\{u_n\}$ satisfies (2.3) for some fixed $\tau > 0$. (i.e. $1 - F(u_n) = P\{\xi_1 > u_n\} \sim \tau/n$.) Then

$$P\{M_n \leq u_n\} \rightarrow e^{-\tau} \text{ as } n \rightarrow \infty .$$

The proof of this result is given in [15] and will not be repeated here. It is based on the standard inequalities

$$\begin{aligned} \sum_i P\{\xi_i > u\} - \sum_{i < j} P\{\xi_i > u, \xi_j > u\} &\leq P\{U(\xi_i > u)\} = P(M_n > u) \\ &\leq \sum_i P\{\xi_i > u\} . \end{aligned}$$

(iii) Relations with the i.i.d. case.

We write \hat{M}_n to denote the maximum of n i.i.d. r.v.'s with the same distribution function F as the terms of the stationary sequence $\{\xi_i\}$.

THEOREM 4.5. Suppose that $P\{\hat{M}_n \leq u_n\} \rightarrow \theta$ for some sequence u_n and some θ , ($0 < \theta < 1$), and that $D(u_n)$, $D'(u_n)$ hold for the stationary sequence $\{\xi_j\}$. Then $P\{M_n \leq u_n\} \rightarrow \theta$.

PROOF: By Theorem (2.5) $1 - F(u_n) \sim \tau/n$ with $\tau = -\log \theta$, i.e. (2.3) holds with this $\{u_n\}$. Hence by Theorem 4.4

$$P\{M_n \leq u_n\} \rightarrow e^{-\tau} = \theta .$$

Thus the same limit holds for the maximum in the dependent sequence as would apply if the random variables were i.i.d. with the same marginal distribution function. As a corollary we may consider the asymptotic distribution

of M_n under the standard normalization.

THEOREM 4.6. If $P\{a_n(\hat{M}_n - b_n) \leq x\} \rightarrow G(x)$ (non-degenerate) then
 $P\{a_n(M_n - b_n) \leq x\} \rightarrow G(x)$ if $D(u_n)$, $D'(u_n)$ hold for $u_n = x/a_n + b_n$.

This follows at once from Theorem 4.5 for $0 < G(x) < 1$ and by continuity of G , at points where $G(x) = 0$ or 1 .

(iv) Stationary normal sequences.

If $\{\xi_j\}$ is a (zero mean, unit variance) stationary normal sequence whose covariances $\{r_n\}$ satisfy either condition (3.2), then it is known (Berman, [1]) that

$$P\{a_n(M_n - b_n) \leq x\} \rightarrow \exp(-e^{-x})$$

where a_n and b_n are given by (2.7). This result also follows as a consequence of Theorem 4.6. For this limit applies to i.i.d. random variables (as shown in Section 2) and it follows by Lemma 4.3 of [15] that either condition (3.2) implies both $D(u_n)$, $D'(u_n)$ when $u_n = x/a_n + b_n$, in this particular case.

The ultimate amount of work in this route to the asymptotic distribution in the normal case is not less than that of the original proof in [1]. However it does indicate the satisfactory nature of the D , D' conditions, since they are implied by the very weak conditions (3.2) when the sequence is normal.

5. EXCEEDANCES OF HIGH LEVELS BY A STATIONARY SEQUENCE, AND THE DISTRIBUTION OF r^{TH} LARGEST VALUES.

(i) Exceedances of high levels.

As previously noted, any point j where ξ_j exceeds u will be called an "exceedance of the level u " by the sequence $\{\xi_j\}$. The exceedances form a *point process* i.e. a series of events occurring in "time" according to some probabilistic law. Suppose now that $\{u_n\}$ is a sequence of constants satisfying (2.3). As n increases, the probability of an exceedance becomes smaller and the exceedances consequently tend to become rarer. An obvious question is whether one may thus obtain a convenient "limiting point process". We have seen in the i.i.d. case that the number of exceedances in $(1,2\dots n)$ has an asymptotic Poisson distribution and one may expect that this will be true for stationary sequences under appropriate assumptions. We shall indicate how this may be shown - and indeed that the sequence of point processes formed from the exceedances of u_n converges weakly - as random elements of a natural metric space - to a Poisson process.

First we make a time scale change to avoid degeneracy. Let η_n be a discrete-parameter process defined on the points $\frac{j}{n}$, $j = 0,1,2\dots$, by $\eta_n(j/n) = \xi_j$. Consider the point process consisting of exceedances of u_n by η_n and let $N_n(B)$ denote the number of such exceedances in the Borel set B . (i.e. the number of exceedances of u_n by ξ_j for $j \in nB$.) Thus while we "lose exceedances" by increasing u_n we gain them in a compensating way by this change of time scale.

N_n is a point process for each n - i.e. a random element in the space N of non-negative integer-valued Borel measures on the real line R . N becomes a metric space with the "vague topology" (cf. [13]) and we may therefore consider convergence in distributions of these random elements. (That is a sequence $\{\zeta_n\}$ of point processes converges in distribution to a point process ζ , $\zeta_n \xrightarrow{D} \zeta$ if $Ef(\zeta_n) \rightarrow Ef(\zeta)$ for every bounded, vaguely continuous real f on N . The following is a useful criterion (due to Kallenberg [13] and modified according to a remark of Kurtz [14]) for such convergence.

THEOREM 5.1. (Kallenberg)

Let ζ_n , $n = 1, 2, \dots$ be point processes on the positive real line and let ζ be a point process without multiple points and such that $\zeta(\{a\}) = 0$ a.s. for every fixed real $a \geq 0$. If

(i) $P\{\zeta_n(B) = 0\} \rightarrow P\{\zeta(B) = 0\}$ for all sets B of the form

$$\bigcup_1^r (a_i, b_i] \quad (a_1 < b_1 < a_2 < \dots < a_r < b_r)$$

(ii) $\limsup_n E\zeta_n(a, b] \leq E\zeta(a, b]$ for all finite $a < b$ then $\zeta_n \xrightarrow{d} \zeta$.

In our case $\zeta_n = N_n$, and ζ is a Poisson process with parameter τ , so that e.g.

$$P\{\zeta(B) = r\} = e^{-\tau m(B)} [\tau m(B)]^r / r!$$

where B is any Borel set and m denotes Lebesgue measure.

We shall suppose that $D(u_n)$, $D'(u_n)$ both hold and thus by Theorem 4.4 $P\{M_n \leq u_n\} \rightarrow e^{-\tau}$. But $\{M_n \leq u_n\}$ is the same as $\{N_n((0, u_n]) = 0\}$ so that

$$P\{N_n((0,1]) = 0\} \rightarrow e^{-\tau} = P\{\zeta((0,1]) = 0\}$$

i.e. Condition (i) holds when $B = (0,1]$.

It may be shown from this that Condition (i) holds for intervals $(0,\alpha]$ and then for intervals $(a,b]$ and finally for finite disjoint unions $U(a_i, b_i]$. (The details of this development may be found in [17].)

Thus Condition (i) holds. Condition (ii) is easy to verify since (writing $[x]$ for the integer part of x)

$$\begin{aligned} EN_n(a,b] &= ([nb] - [na])(1 - F(u_n)) \\ &\sim n(b - a)\tau/n = \tau(b - a) \\ &= E\zeta((a,b]) . \end{aligned}$$

Thus Theorem 5.1 applies and we obtain the following result.

THEOREM 5.2. Let $\{\xi_j\}$ be a stationary sequence such that $D(u_n)$, $D'(u_n)$ hold for $\{u_n\}$ satisfying (3.2). Let the point process N_n consist of those points j/n for which $\xi_j > u_n$ ($N_n(B)$ being the number of such points in the Borel set B .) Then $N_n \xrightarrow{D} \zeta$ where ζ is a Poisson process with parameter τ .

COROLLARY. Under the conditions of the theorem, if B is any Borel set whose boundary has Lebesgue measure zero ($m(\partial B) = 0$) then for any $r = 0,1,2,\dots$,

$$P\{N_n(B) = r\} \rightarrow e^{-\tau m(B)} [\tau m(B)]^r / r!$$

(with similar convergence for the joint distribution of any $N_n(B_1) \dots N_n(B_k)$).

This follows at once since the random variables $N_n(B)$ converge in distribution to $\zeta(B)$ when $N_n \xrightarrow{D} \zeta$ (cf. [13]), and similarly for joint distributions.

(ii) Asymptotic distribution of $M_n^{(r)}$

THEOREM 5.3. Let $\{\xi_j\}$ be a stationary sequence such that $D(u_n)$, $D'(u_n)$ hold for $\{u_n\}$ satisfying (2.3). Then

$$P\{M_n^{(r)} \leq u_n\} \rightarrow \sum_{s=0}^{r-1} e^{-\tau} \tau^s / s!$$

($M_n^{(r)}$ being the r^{th} largest among $\xi_1 \dots \xi_n$).

PROOF: This result follows simply from Theorem 5.2 by noting that

$$P\{M_n^{(r)} \leq u_n\} = P\{N_n((0,1]) < r\} \quad (\text{cf. Theorem 2.7}).$$

Finally we may generalize Theorem 4.6 to obtain (2.10) for our stationary sequences. \hat{M}_n will again denote the maximum of n i.i.d. random variables with the same distributions as ξ_1 .

THEOREM 5.4. Let $\{\xi_i\}$ be a stationary sequence. Suppose that

$P\{a_n(\hat{M}_n - b_n) \leq x\} \rightarrow G(x)$ (non-degenerate) and that $D(u_n)$, $D'(u_n)$ hold for $u_n = x/a_n + b_n$. Then (2.10) holds, i.e.

$$P\{a_n(M_n^{(r)} - b_n) \leq x\} \rightarrow G(x) \sum_{s=0}^{r-1} (-\log G(x))^s / s! .$$

PROOF: $\{u_n\} = \{x/a_n + b_n\}$ satisfies (2.3) by Theorem 2.5. The result now follows by Theorem 5.3 with $\tau = -\log G(x)$ (where $0 < G(x) < 1$) and by

continuity arguments where $G = 0$ or 1 .

6. CONTINUOUS PARAMETER STATIONARY PROCESSES

In this section we consider a continuous parameter stationary process $\{\xi(t): t \geq 0\}$ (assumed to have continuous sample functions), and write $M(T) = \sup\{\xi(t): 0 \leq t \leq T\}$.

(i) Strongly mixing processes.

The definition of strong mixing given is easily adapted to the continuous parameter case requiring, for any t ,

$$|P(A \cap B) - P(A)P(B)| < g(\tau)$$

for $A \in \sigma\{\xi_s: s \leq t\}$, $B \in \sigma\{\xi_s: s \geq t + \tau\}$ where $g(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$.

Gnedenko's Theorem may be extended as follows.

THEOREM 6.1. Suppose the stationary process $\{\xi(t)\}$ is strongly mixing and $P\{a_T(M(T) - b_T) \leq x\} \rightarrow G(x)$, non-degenerate, for some families $a_T > 0$, b_T of constants. Then $G(x)$ has one of the three extreme value forms.

PROOF: Write $Z_i = \sup\{\xi(t): i - 1 \leq t \leq i\}$. Then $\{Z_i: i = 1, 2, \dots\}$ may be seen to be a stationary and strongly mixing sequence. Further, putting $T = n$, an integer, $M(n) = \max\{Z_i: 1 \leq i \leq n\}$, and

$$P\{a_n(M(n) - b_n) \leq x\} \rightarrow G(x)$$

so that by Gnedenko's Theorem for strongly mixing sequences $G(x)$ has one of the three extreme value forms.

Other partial results may be shown under strong mixing or D, D' type conditions. For example we may see under appropriate conditions, that

$$P\{M(n) \leq u_n\} \rightarrow e^{-\tau} \quad \text{if} \quad P\{Z_1 > u_n\} \sim \tau/n$$

($Z_1 = \sup\{\xi(t): 0 \leq t \leq 1\}$) and thence

$$(6.1) \quad P\{M(T) \leq u_T\} \rightarrow e^{-\tau}$$

if

$$(6.2) \quad P\{Z_1 > u_T\} \sim \tau/T .$$

This may be made rigorous under sufficiently strong conditions. However natural, simple conditions need to be found to make the result satisfying and useful in the general context. For the *normal* case there are such conditions, as will be seen below.

(ii) Stationary normal processes.

Extreme values of stationary normal processes have been considered in detail by a number of authors, including S. M. Berman, J. Pickands, C. Qualls and H. Watanabe. Let $\{\xi(t)\}$ then be such a process with (for convenience) zero mean, unit variance and covariance function $r(\tau) = E\xi(t)\xi(t + \tau)$. It is usually assumed that $r(\tau)$ has the following expansion as $\tau \rightarrow 0$.

$$(6.3) \quad r(\tau) = 1 - C|\tau|^\alpha + o(|\tau|^\alpha)$$

for some α with $0 < \alpha \leq 2$, though the theory can be extended to apply to the case where a function $G(\tau)$ of slow "growth" is included as a multiplicative

factor together with $|\tau|^\alpha$ (cf. [24]).

$\alpha = 2$ gives the "regular" case - where $\xi(t)$ has a quadratic mean derivative and where the mean number of "upcrossings" of any level per unit time, is finite. If $0 < \alpha < 2$, the sample functions are continuous, but the mean number of upcrossings of a level u in any interval is infinite, and there is no q.m. derivative. The case $\alpha = 1$ corresponds to the Ornstein-Uhlenbeck Process.

The development indicated under (i) above may be applied to stationary normal processes satisfying (6.3) and in so doing we may replace strong mixing by either of the conditions

$$(6.4) \quad r(t) \log t \rightarrow 0 \quad \int_0^\infty r^2(t) dt < \infty$$

(i.e. the continuous analogues of (3.2)). Indeed the following result holds.

LEMMA 6.1. Suppose (6.3) and (6.4) are satisfied, for the stationary normal process $\xi(t)$ considered above. Then

$$P\{M(T) \leq u_T\} \rightarrow e^{-\tau} \quad \text{as } T \rightarrow \infty$$

if u_T is chosen so that (6.2) holds, i.e.

$$P\{\sup\{\xi(t): 0 \leq t \leq 1\} > u_T\} \sim \tau/T .$$

This lemma may be obtained e.g. from the discussion in [22] or [3]. From these references we may also obtain the asymptotic form of $P\{\sup\{\xi(t): 0 \leq t \leq 1\} > u\}$ as follows.

LEMMA 6.2. Under the conditions of Lemma 6.1, as $u \rightarrow \infty$,

$$P\{\sup(\xi(t): 0 \leq t \leq 1) > u\} \sim C^{1/\alpha} H_\alpha u^{2/\alpha} \phi(u)/u$$

where H_α is a certain constant and ϕ is the standard normal probability density function. Thus the conclusion $P(M(T) \leq u_T) \rightarrow e^{-\tau}$ of Lemma 6.1 holds if u_T is chosen to satisfy

$$(6.5) \quad C^{1/\alpha} u_T^{2/\alpha} \phi(u_T)/u_T \sim \tau/T \quad \text{as } T \rightarrow \infty.$$

By taking logarithms in (6.5) we may obtain, after some manipulation

$$(6.6) \quad u_T = (2 \log T)^{1/2} - (2 \log T)^{-1/2} \left[\log \tau + \left(\frac{1}{2} - \frac{1}{\alpha} \right) \log \log \tau \right. \\ \left. + \log [(2\pi)^{1/2} C^{1/\alpha} 2^{(\alpha-2)/2\alpha} H_\alpha^{-1}] \right] + o(\log T)^{-1/2}$$

and Lemma 6.2 at once gives the following theorem.

THEOREM 6.3. Let the stationary normal process $[\xi(t)]$ satisfy (6.2) and (6.3). Then

$$P\{a_T(M(T) - b_T) \leq x\} \rightarrow \exp(-e^{-x}) \quad \text{as } T \rightarrow \infty$$

where

$$a_T = (2 \log T)^{1/2} \\ b_T = (2 \log T)^{1/2} - (2 \log T)^{-1/2} \left[\left(\frac{1}{2} - \frac{1}{\alpha} \right) \log \log T \right. \\ \left. + \log [(2\pi)^{1/2} C^{1/\alpha} 2^{(\alpha-2)/2\alpha} H_\alpha^{-1}] \right]$$

PROOF: Put $\tau = e^{-x}$ in (6.6) to give $u_{\tau} = x/a_{\tau} + b_{\tau}$ and apply Lemma 6.2.

The value of the constant H_{α} is known only for $\alpha = 1, 2$ ($H_1 = 1$ $H_2 = \pi^{-1/2}$). Its form is given in [22], and depends on α , but not otherwise on the process $\xi(t)$.

The derivation of Lemmas 6.1 and 6.2 naturally involves considerable calculation, and different approaches may be used (cf. [22], [3]). In [22] use is made of the " ϵ -upcrossings" of a level. $\xi(t)$ is said to have an *upcrossing* of u at t_0 if $\xi(t) \leq u$ for $t_0 - \delta < t \leq t_0$ and $\xi(t) \geq u$ for $t_0 \leq t < t_0 + \delta$, for some $\delta > 0$, and this is an ϵ -*upcrossing* if there are no other upcrossings in $(t_0 - \epsilon, t_0)$. As noted before if $\alpha = 2$ in (6.3) the number of upcrossings in $0 \leq t \leq 1$ has a finite mean, whereas this is not so for $\alpha < 2$. However the number of ϵ -upcrossings (for any fixed $\epsilon > 0$) in any finite interval is a bounded random variable, which turns out to be useful in the case $\alpha < 2$.

The (ϵ -)upcrossings in the continuous case naturally replace the exceedances in the discrete case. It is also possible to show that their limiting distribution (for suitably chosen high levels) is Poisson (cf. [4], [2], [23]). In fact it has recently been shown by Lindgren et al. ([18]) that these (ϵ -)upcrossings converge weakly to a Poisson process. The proof is similar to that described here in the sequence case (Sec. 5) for exceedances. Further questions concerning the asymptotic distribution of r^{th} largest maxima and closely related matters are also discussed in [18].

Finally we note that it has also been recently shown by Mittal and Ylvisaker ([27]), that if the conditions (6.4) do not hold, a variety of other limit laws are possible for the maximum.

REFERENCES

(Including some related works not referred to in text)

- [1] Berman, S.M. Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.*, 35 (1964) pp. 502-516.
- [2] Berman, S.M. Asymptotic independence of the numbers of high and low level crossings of stationary Gaussian processes. *Ann. Math. Statist.*, 42 (1971) pp. 927-945.
- [3] Berman, S.M. Maxima and high level excursions of stationary Gaussian processes. *Trans. Amer. Math. Soc.*, 160 (1971) pp. 67-85.
- [4] Cramér, H. & Leadbetter, M.R. Stationary and Related Stochastic Processes, John Wiley & Sons, New York, (1967).
- [5] De Haan, L. On regular variation and its application to the weak convergence of sample extremes. *Math. Centre Tract*, 32 (1970) Amsterdam.
- [6] Deo, C.M. A note on strong mixing Gaussian sequences. *Ann. Prob.*, 1 (1973) pp. 186-187.
- [7] Feller, W. An Introduction to Probability Theory and its Applications 2. John Wiley & Sons, New York, (1966).
- [8] Galambos, J. On the distribution of the maximum of random variables. *Ann. Math. Statist.*, 43 (1972) pp. 516-521.
- [9] Galambos, J. A general Poisson limit theorem of probability theory. *Duke Math. J.*, 40 (1973) pp. 581-586.
- [10] Gnedenko, B.V. Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, 44 (1943) pp. 423-453.
- [11] Gnedenko, B.V. & Kolmogorov, A.N. Limit Distributions for Sums of Independent Random Variables. Addison Wesley, New York, (1954).
- [12] Gumbel, E.J. Statistics of Extremes. Columbia Univ. Press, New York, (1958).
- [13] Kallenberg, O. Characterization and convergence of random measures and point processes. *Z. Wahr. verw. Geb.*, 27 (1973) pp. 9-21.
- [14] Kurtz, T. Point processes and completely monotone set functions. Tech. Rept. Dept. Math, Univ. of Wisconsin (1973).

REFERENCES (cont.)

- [15] Leadbetter, M.R. On extreme values in stationary sequences. *Z. Wahr. verw. Geb.*, 28 (1974) pp. 289-303.
- [16] Leadbetter, M.R. Lectures on extreme value theory. Tech. Rept. Dept. Math. Stat., Lund Univ. Sweden, (1974).
- [17] Leadbetter, M.R. Weak convergence of high level exceedances by a stationary sequence. *Inst. of Stat. Mimeo Series No. 933*, Univ. of N.C., (1974).
- [18] Lindgren, G., de Maré, J., & Rootzén, H. Weak convergence of high level crossings and maxima for one or more Gaussian processes. Tech. Rept. 4 Dept. Math. Stat., Lund Univ., Sweden, (1974).
- [19] Loynes, R.M. Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.*, 36 (1965) pp. 993-999.
- [20] O'Brien, G.L. Limit theorems for the maximum term of a stationary process. *Ann. Prob.*, (1974) pp. 540-543.
- [21] O'Brien, G.L. The maximum term of uniformly mixing stationary processes. To appear in *Z. Wahr. verw. Geb.*
- [22] Pickands, J. Asymptotic properties of the maximum in a stationary Gaussian process. *Trans. Amer. Math. Soc.*, 145 (1969) pp. 75-86.
- [23] Pickands, J. Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.*, 145 (1969) pp. 51-73.
- [24] Qualls, C., & Watanabe, H. Asymptotic properties of Gaussian processes. *Ann. Math. Statist.*, 43 (1972) pp. 529-596.
- [25] Watson, G.S. Extreme values in samples from m-dependent stationary stochastic processes. *Ann. Math. Statist.*, 25 (1954) pp. 798-800.
- [26] Welsch, R.E. A weak convergence theorem for order statistics from strong-mixing processes. *Ann. Math. Statist.*, 42 (1971) pp. 1737-1946.
- [27] Mittal, Y., & Ylvisaker, D. Limit distributions for the maxima of stationary Gaussian processes. To appear in *J. Stoch. Proc.*