The Annals of Probability
1982, Vol. 10, No. 3, 737–744

THE MAXIMUM TERM AND FIRST PASSAGE TIMES FOR AUTOREGRESSIONS[1]

BY MARK FINSTER

Cornell University

The limiting distribution of the maximum term of the non-normal stationary sequence $\cdots X_{-1}, X_0, X_1 \cdots$ satisfying the autoregressive equation $X_n = \varepsilon_n + a_1 X_{n-1} + a_2 X_{n-2} + \cdots$ is investigated when $\sum |a_k| < 1$ and $\cdots \varepsilon_{-1}, \varepsilon_0, \varepsilon_1 \cdots$ are integrable real valued i.i.d. random variables having distributions with tails that are either Pareto or exponential in nature. Asymptotic results for the joint distribution of the first passage time $t = \inf\{n : X_n \geq c\}$ and the excess $R_t = X_t - c$ are also given as $c \to \infty$.

1. Introduction and summary. The limiting distribution of the maximum term of the non-normal stationary sequence $\cdots X_{-1}, X_0, X_1 \cdots$ satisfying the autoregression equation

$$X_n = \varepsilon_n + a_1 X_{n-1} + a_2 X_{n-2} + \cdots \tag{1.1}$$

is determined when $\sum |a_k| < 1$ and $\cdots \varepsilon_{-1}, \varepsilon_0, \varepsilon_1, \cdots$ are real valued i.i.d. random variables having distributions that are either Pareto or exponential in nature. The distribution of ε_n is Pareto if there exists an $\alpha > 1$ and a function $L(x)$ varying slowly at infinity (Feller: 1966, page 276) such that $P(|\varepsilon_n| > x) = x^{-\alpha} L(x)$ and the distribution of ε_n is exponential if there exists $\alpha > 0$ such that $P(|\varepsilon_n| > x) \sim e^{-\alpha x}$ as $x \to \infty$. Utilizing the asymptotic distribution of the extremes, the joint normalized limiting distribution of the first passage time

$$t = t_c = \inf\{n \geq 1 : X_n \geq c\} = \inf\{n \geq 1 : \max[X_1, \cdots, X_n] \geq c\}$$

and the remainder $R_t = X_t - c$ is calculated as $c \to \infty$. An application to optimal stopping is found in Finster (1982).

Extensive research has been directed at the determination of the asymptotic distribution of the properly normalized maximum term of a sequence of random variables. Gnedenko (1943) has completely characterized the possible nondegenerate limiting distributions and their respective domains of attraction of i.i.d. sequences. If the marginal distribution of one random variable is Pareto, the normalized maximum term converges in distribution to the law

$$H_1(x) = \exp(-x^{-\alpha}) I_{(0,\infty)}(x) \tag{1.2}$$

where I is the indicator function, and if exponential, the limiting distribution is of the type

$$H_2(x) = \exp(-e^{-x}). \tag{1.3}$$

Many have given conditions under which the asymptotic distribution of the extreme value is identical to that of an i.i.d. sequence having the same marginal distribution. Frequently the dependence requirement of the stochastic sequence $\{X_n\}$ is weakened by assuming that $\{X_n\}$ is strictly stationary and that the dependence between X_i and X_j decreases in some fasion as $|i - j|$ increases. Watson (1954) generalized Gnedenko's results by the simplest such restriction, that of m dependence, which requires that X_i and X_j

Received May 1980; revised July 1981.

[1] Research supported by NSF Grant No. MCS75-22481.

AMS 1970 *subject classifications.* Primary 60G10, 60G40; secondary 60G50, 60F99.

Key words and phrases. Autoregression, maximum term, Pareto, exponential, first passage time, asymptotic distribution.

actually be independent if $|i - j| > m$. Watson also required an assumption similar to Leadbetter's D' condition (Leadbetter: 1974b). Loynes (1965) relaxes the m-dependency restriction by assuming strong (or uniform) mixing and O'Brien (1974a, 1974b) adds to Loynes' results. Specifically, $\{X_n\}$ satisfies a strong mixing assumption if there exists a "mixing" function $g(k) = o(1)$ such that for all j and k

$$|P(AB) - P(A)P(B)| < g(k)$$

wherever $A \in \mathscr{F}(X_n : n = 1, 2, \cdots, j)$ and $B \in \mathscr{F}(X_n : n = j + k + 1, \cdots)$, $\mathscr{F}$ being the sigma field generated by the corresponding variables. Berman (1964) shows that for stationary normal sequences the classic theorems hold providing the correlation $\rho_n = \rho(X_1, X_{n+1})$ decreases according to

$$\rho_n \log n = o(1) \quad \text{or} \quad \sum \rho_n^2 < \infty. \tag{1.4}$$

Berman (1964) also uses a "comparison technique" introduced by Lévy (1937) and extended by Loève (1960) to obtain still different conditions for non-normal stationary sequences. Leadbetter (1974a, 1974b, 1975, 1979) assumes condition D, a weaker type of mixing that involves only sets of the form $\{X_1 \le c, \cdots, X_n \le c\}$ and that generalizes the mixing function to a mixing sequence. His results extend Loynes' and both correlation restrictions of Berman.

Berman (1962) also investigates the case of exchangeable random variables and the case where the number N_n of random variables considered in the determination of the maximum is itself a random variable and $N_n \to_P \infty$ as $n \to \infty$. Galambos (1972) applies combinatorial concepts and a generalization of the inclusion-exclusion principle (a graph sieve theorem) to obtain results for arbitrary and possible nonstationary sequences, results that encompass those for uniform mixing sequences. In addition, Galambos (1978) has aptly compiled the fundamentals of the asymptotic theory of extremes.

Rootzén (1978) has also studied the extremes of moving averages of stable processes in both discrete and continuous time. Chernick (1978, 1981) has given counter examples with first order autoregressive sequences—one with Cauchy marginals and hence of Pareto type for which the classical limit theorem does not apply, and one with uniform marginals for which the limit theorems of Loynes and Leadbetter do not apply.

Using Lévy's comparison technique we show that the limiting distribution of the modulus $\bar{M}_n = \max\{|X_1|, \cdots, |X_n|\}$ of the maximum term of the autoregression (1.1) coincides with that of the corresponding i.i.d. case. Furthermore, for a Pareto distribution this result holds even after conditioning on the past $\{X_k, k \le 0\}$. If the right tails of $\{\varepsilon_n\}$ dominate, that is if $P(\varepsilon_n < -x)/P(\varepsilon_n > x) = o(1)$ as $x \to \infty$, then Gnedenko's theorem holds for the maximum term $M_n = \max\{X_1, \cdots, X_n\}$.

Employing these results, the first passage time

$$t = t_c = \inf\{n \ge 1 : M_n \ge c\}$$

and the remainder $R_t = X_t - c$ properly normalized are shown to be asymptotically independent and their joint limiting distribution is calculated as $c \to \infty$. The normalized stopping time converges in distribution to the exponential law $[1 - e^{-x}]I_{(0,\infty)}(x)$. In the exponential case R_t also converges in distribution to this exponential law and in the Paretian case the distribution of $R_t c^{-1}$ converges to the Pareto law $[1 - (x + 1)^{-\alpha}]I_{(0,\infty)}(x)$. Under certain restrictions the limits of Et and ER_t are determined asymptotically via uniform integrability.

Of course if ε_n has dominant left tails which are either Pareto or exponential, analogous results hold for the minimum.

2. The maximum term—Pareto Case. Let $a = (a_1, a_2, \cdots)' \in \mathbb{R}^\infty$ and define $\{b_n\}$ inductively by $b_0 = 1$ and

$$b_n = a_1 b_{n-1} + \cdots + a_n b_0 \quad \text{for} \quad n > 0$$

so that $b = (b_0, b_1, \cdots)' = \sum_{k=0}^\infty S^k(a^{k*})$ where the * product is convolution, $S(a) =$

$(0, a_1, \cdots)'$ is the shift right operator, and $'$ denotes transpose. In other words, a and b are column vectors in ℓ^1, the absolutely summable sequences, the ith coordinate of $a*b$ is $\sum_1^i a_j b_{i-j+1}$, and $|b| \le (1-|a|)^{-1} < \infty$ since we always assume $|a| = \sum |a_k| < 1$. If e_k is the standard kth basis vector for $\mathbb{R}^\infty$, b can be alternatively defined by $b*(e_1 - S(a)) = e_1$.

$$X_n = \sum_0^\infty b_k \varepsilon_{n-k} \tag{2.1}$$

is then the unique stationary time series satisfying (1.1).

Define the transpose matrix $A' = (a, e_1, e_2, \cdots)$ and symbolize the past by $Z_n = (X_n, X_{n-1}, \cdots)'$ so that

$$Z_n = A^n Z_0 + \sum_0^{n-1} A^k e_{n-k}\varepsilon_{n-k}. \tag{2.2}$$

Let P_z represent the probability measure on $\{Z_n\}$ or $\{X_n\}$ when the past values $\{X_k, k \le 0\}$ are given by $Z_0 = z \in \mathbb{R}^\infty$ in (2.2). The values of Z_0 are restricted to the set $\Lambda = \Lambda(a)$. To define Λ set $\eta_n = (\varepsilon_n, \varepsilon_{n-1}, \cdots)'$ so that $X_n = b'\eta_n$ and $Z_n = B\eta_n$ where the upper triangular matrix $B = (b_{ij})$ has entries $b_{ij} = b_{j-i}$ $(j \ge i)$. Set

$$\Lambda = \{B\eta : \eta \in \mathbb{R}^\infty \quad \text{and} \quad \lim_n\sup T^n(\bar{a}*\bar{b})'\bar{\eta} < \infty\}.$$

Here $\bar{w} = (|w_1|, |w_2|, \cdots)'$ for $w = (w_1, w_2, \cdots)' \in \mathbb{R}^\infty$ and $T(w) = (w_2, w_3, \cdots)'$ is the left shift operator. Since $(a*b)\eta_0 = a'Z_{-1} + \varepsilon_0 = X_0$, it is easy to see $P(Z_0 \in \Lambda) = 1$.

LEMMA 2.1. *If $z \in \Lambda$, there exists a constant $K = K(a, z)$ so that on $\{Z_0 = z\}$*

$$|a'Z_n| \le |\sum_0^{n-1} a_{j+1}X_{n-j}| + K \qquad \forall n.$$

PROOF. Let $\eta = (\varepsilon_0, \varepsilon_{-1}, \cdots)'$ satisfy $z = B\eta$ and set $S_n = |\sum_0^{n-1} a_{j+1}X_{n-j}|$ so

$$|a'Z_n| \le S_n + |\sum_0^\infty a_{n+j+1}X_{-j}| = S_n + |T^n(a)'B\eta|$$
$$\le S_n + [T^n(\bar{a}) * \bar{b}]'\bar{\eta} \le S_n + T^n(\bar{a} * \bar{b})'\bar{\eta}.$$

For $z \in \Lambda$, $K(a, z) = \sup_n T^n(\bar{a} * \bar{b})'\eta$ is finite. □

Let F be the distribution function (d.f.) of ε_n and G the d.f. of $|\varepsilon_n|$. Call a d.f. H Pareto with exponent $\alpha > 1$ if there exists a function $L(x)$ varying slowing at infinity such that $1 - H(x) = x^{-\alpha}L(x)$. Throughout this paper L will represent the slowly varying function corresponding to the particular Pareto d.f. under discussion.

Define $M_n = \max\{X_1, \cdots, X_n\}$ and $\bar{M}_n = \max\{|X_1|, \cdots, |X_n|\}$. We now show that the classical results remain valid for $\bar{M}_n$ and, under an additional assumption for M_n, even after conditioning on Z_0. The notation $E[Y; A] = \int_A Y dP$ will be used throughout this article.

THEOREM 2.1. *If the distribution function G of $|\varepsilon_n|$ is Pareto and $\{c_n\}$ are constants satisfying $1 - G(c_n) = O(n^{-1})$ then $\forall z \in \Lambda$, $P_z(\bar{M}_n \le c_n) - G^n(c_n) = o(1)$.*

PROOF. On $\mathbb{R}^\infty \times \mathbb{R}^\infty$ define coordinates $(X, U) = (X_1, X_2, \cdots) \times (U_1, U_2, \cdots)$ and a probability P_z' so that the marginal distribution of X coincides with that of our autoregression sequence (1.1) under P_z and so that U is independent of X with i.i.d. coordinates U_j having marginal d.f. $G(x) = 1 - x^{-\alpha}L(x)$. Let A_1 and B_{n+1} be probability one events and set $A_j = \{\bar{M}_{j-1} \le c_n\}$, $B_j = \{U_i \le c_n, i = j, \cdots, n\}$ and $\mathscr{F}_j = \mathscr{F}(X_i : i \le j)$. Writing P for P_z', E for expectation under P_z', c for c_n and utilizing $\{|X_j| \le c\} = \{-c - a'Z_{j-1} \le \varepsilon_j \le c - a'Z_{j-1}\}$ we have

$$|P(\bar{M}_n \le c) - G^n(c)| = |P(A_{n+1}) - P(B_1)|$$

$$= |\sum_1^n P(A_{j+1}B_{j+1}) - P(A_jB_j)| \tag{2.3}$$

$$= |\sum_1^n P(B_{j+1})E[P(|X_j| \le c \mid \mathscr{F}_{j-1}) - P(U_j < c); A_j]$$

$$\le \sum E[|F(c - a'Z_{j-1}) - F(-c - a'Z_{j-1}) - G(c)|; A_j]$$

$$\le [1 - G(c)]E[\sum (\Omega_j^+ + \Omega_j^-)] \tag{2.4}$$

where

$$\Omega_j^{\pm} = |1 - [1 - G(c \pm a'Z_{j-1})][1 - G(c)]^{-1}| I_{A_j}.$$

By the choice of c it suffices to show n^{-1} times the integrand in (2.4) is dominated and converges to zero. Writing μ_j for $c^{-1}a'Z_{j-1}$ we have

$$\textstyle\sum_1^n \Omega_j^+ \le \sum |1 - [1 - \mu_j]^{-\alpha}| I_{A_j} + \sum |[1 - \mu_j]^{-\alpha}[1 - L(c - c\mu_j)/L(c)] I_{A_j}. \tag{2.5}$$

For K satisfying Lemma 2.1 the third summation in (2.5) is bounded uniformly by

$$n(1 - |a| - c^{-1}K)^{-\alpha} \sup\{|1 - L(cy)/L(c)| : 1 - |a| - c^{-1}K \le y \le 1 + |a| + c^{-1}K\}$$

which is $o(n)$ by Feller (1966, page 276). Since $|\mu_j|$ is bounded away from one on A_j for sufficiently large n, the mean value theorem gives a constant $J = J(a)$ such that for all j

$$|1 - [1 - \mu_j]^{-\alpha}| \le J|\mu_j|.$$

Hence, if $z = B\eta$ then the middle summation in (2.5) is bounded by J/c times

$$\textstyle\sum_{j<n} |c\mu_j| = \sum_{j<n} |(a * b)'\eta_{j-1}| \le \sum_{k\le j<n} (\bar{a} * \bar{b})'e_k |\varepsilon_{j-k}| + \sum_{j<n} T^j(\bar{a} * \bar{b})'\bar{\eta}.$$

The last quantity is $o(nc)$ for $z \in \Lambda$ *and*

$$\textstyle E[\sum_{k\le j<n} (\bar{a}*\bar{b})'e_k |\varepsilon_{j-k}|] \le n(\bar{a}*\bar{b})E|\varepsilon_1| = o(nc).$$

That $E[\sum \Omega_j^-] = o(n)$ in (2.4) may be established similarly. □

REMARK 2.1. We have actually proved more. If $\{c_n(\ell)\}$ is any class of sequences tending uniformly in ℓ to infinity as $n \to \infty$ and if $\sup_{n,\ell} n\{1 - G[c_n(\ell)]\} < \infty$ then the convergence in Theorem 2.1 is uniform in ℓ.

Theorem 2.1 can be paraphrased in distributional terms to yield a version of Gnedenko's Theorem. Let H_1 be the limiting d.f. of Gnedenko defined in (1.2).

COROLLARY 2.1. *If G is Pareto with exponent $\alpha > 1$ and $z \in \Lambda$ then the distribution function of $(\bar{M}_n - v_n)/u_n$ under P_z converges to $H_1(x)$ provided $nu_n^{-\alpha}L(u_n) \to 1$ and $v_n = o(u_n)$.*

The right tail of a d.f. F dominates if $F(-x) = o[1 - F(x)]$ as $x \to \infty$. Although Theorem 2.1 is concerned only with the magnitude of the maximum term, the following corollary extends this result to $M_n = \max\{X_1, \cdots, X_n\}$ itself.

COROLLARY 2.2. *If the distribution function F of ε_n has dominating Paretian right tails and if $\{c_n\}$ are constants satisfying $1 - F(c_n) = O(n^{-1})$, then*

$$P_z(M_n \le c_n) - P_z(\bar{M}_n \le c_n) = o(1) \qquad \forall z \in \Lambda.$$

Hence, Theorem 2.1 and Corollary 2.1 hold with $\bar{M}_n$ and G replaced by M_n and F respectively.

PROOF. If K is chosen to satisfy Lemma 2.1, then for $P = P_z$ and $c = c_n$

$$\begin{aligned} P(M_n \le c) - P(\bar{M}_n \le c) &= P(\bar{M}_n > c, M_n \le c) \\ &\le \textstyle\sum_{j=1}^n P(\bar{M}_{j-1} \le c, X_j < -c) \\ &= \textstyle\sum E[F(-c - a'Z_{j-1}); \bar{M}_{j-1} \le c] \\ &\le nF(-c[1 - |a|] + K) = o(1) \quad \square \end{aligned}$$

All the results in this section remain valid under P, of course, by dominated convergence.

3. The maximum term-exponential case. If $\cdots \varepsilon_{-1}, \varepsilon_0, \varepsilon_1, \cdots$ are i.i.d. normally distributed with mean μ and variance σ^2, then by (2.1) $\{X_n\}$ is a stationary Gaussian sequence with X_n having mean $\sum b_k\mu$ and variance $\sum b_k^2\sigma^2$. Hence, Berman's correlation condition (1.4) gives $P(M_n \le c_n) \to e^{-x}$ provided $P(\varepsilon_n \le c_n) \sim x/n$.

This section investigates a non-Gaussian exponential case using the structure on the autoregression (1.1) introduced in Section 2. The d.f. F of ε_n is exponential if the right tail of F dominates and is asymptotically exponential; that is, if there exists $\alpha > 0$ and $\delta(x) = o(e^{-\alpha|x|})$ as $|x| \to \infty$ such that

$$F(x) = [1 - e^{-\alpha x}]I_{(0,\infty)}(x) + \delta(x). \tag{3.1}$$

Define

$$H(x) = [1 - \beta e^{-\alpha x}]I_A(x) \tag{3.2}$$

where $A = (\alpha^{-1}\log \beta, \infty)$ and $\beta = E[\exp(\alpha a'Z_0)] = E[\exp \alpha(a*b)'\eta_0]$. Note that $\exp(\alpha a'Z_0)$ is integrable, for if $\ell_k = (a*b)'e_k$ then an integration by parts yields

$$E[\exp(\alpha\ell_k\varepsilon_k)] = 1 + \alpha\ell_k \int_0^\infty [1 - F(y) + F(-y)]\exp(\alpha\ell_k y)\, dy.$$

Since each coordinate of b has magnitude bounded by one, $|\ell_k| \le |a| < 1$, and the last integral is uniformly bounded. As $a*b = S(b) \in \ell^1$

$$\beta = \textstyle\prod_{k=1}^{\infty} E[\exp(\alpha\ell_k\varepsilon_k)]$$

converges.

THEOREM 3.1. *Suppose the distribution function F of ε_n has exponential form* (3.1) *and $\{c_n\}$ are constants satisfying* $1 - F(c_n) = O(n^{-1})$; *then* $P(M_n \le c_n) - H^n(c_n) = o(1)$.

PROOF. On $\mathbb{R}^\infty \times \mathbb{R}^\infty$ define coordinates $(X, U) = (\cdots X_{-1}, X_0, X_1 \cdots) \times (U_1, U_2, \cdots)$ and a probability P' so that the marginal distribution of X coincides with the distribution of the autoregression, and so that U is independent of X with i.i.d. coordinates U_k having marginal d.f. H. Set $A_j = \{M_{j-1} \le c\}$ and let $q = q(m, n)$ represent the greatest integer in n/m. Writing P for P', μ_j for $\exp(\alpha a'Z_{j-1})$ and proceeding as in the proof of Theorem 2.1

$$\begin{aligned} |P(M_n \le c) - H^n(c)| &= |\textstyle\sum_1^n H^{n-j}(c)E[\{1 - H(c)\} - \{1 - F(c - a'Z_{j-1})\}; A_j]| \\ &\le ne^{-\alpha c}(qm)^{-1}|\textstyle\sum_1^{qm} H^{n-j}(c)E[\beta - \mu_j; A_j]| + R(n, m) \end{aligned} \tag{3.3}$$

where

$$R(n, m) = |\textstyle\sum_{qm}^n H^{n-j}(c)E[H(c) - F(c - a'Z_{j-1}); A_j]| + nE\,|1 - F(c - a'Z_0) - e^{-\alpha c}\mu_0|.$$

The first summation in $R(n, m)$ is clearly $o(1)$ for fixed m. The integrability of μ_0 and the assumption $1 - F(c) \sim e^{-\alpha c} = O(n^{-1})$ imply the remaining expectation of $R(n, m)$ is $o(n^{-1})$ by dominated convergence since $e^{\alpha c}$ times the integrand is bounded on $\{c \ge a'Z_0)$ by

$$e^{\alpha c}\delta(c - a'Z_0) = o(1) \le \mu_1 \sup_{x \ge 0}[e^{\alpha x}\delta(x)]$$

and on $\{c < a'Z_0\}$ by $3\mu_1$. For $r = 0, 1, \cdots q$ define

$$H_r = m^{-1}\textstyle\sum_{j=rm+1}^{(r+1)m} H^{n-j}(c).$$

Let $N_m = \min[m, \inf\{j \ge 0 : \mathrm{M}_{j+1} > c\}]$ so that the first part of (3.3) is dominated by

$$ne^{-\alpha c}q^{-1}\textstyle\sum_{r=0}^{q-1}\{H_r D_r + m^{-1}\sum_{j=rm+1}^{(r+1)m} 2\beta B_j\}$$

where

$$B_j = H^{n-j}(c) - H_r(c) \le H^{rm+1}(c)[1 - H^m(c)] = o(1)$$

uniformly in j for fixed m and

$$D_r = E[m^{-1} \mid \textstyle\sum_{j=rm+1}^{(r+1)m} (\beta - \mu_j) I_{\{X_k \le c, rm+1 \le k \le j\}} \mid; A_{rm+1}]$$

$$\le E \mid m^{-1} \textstyle\sum_{j=1}^{m} (\beta - \mu_j) I_{A_j} \mid$$

$$\le E[\mid \beta - N_m^{-1} \textstyle\sum_{j=1}^{N_m} \mu_j \mid; N_m > 0].$$

As $n \to \infty$, $N_m \to m$ with probability one. Hence

$$\lim_n\sup \mid P(A_{n+1}) - H^n(c) \mid \le O(1) E \mid \beta - m^{-1} \textstyle\sum_1^m \mu_j \mid.$$

Since any linear combination $X_n = \sum b_k \varepsilon_{n-k} (b \in \ell^1)$ is mixing (Rosenblatt, 1962, pages 105–112) and hence ergodic, the last quantity tends to zero as $m \to \infty$ by the ergdoic theorem.

Remark 2.1 also applies to Theorem 3.1.

Let $H_2(x)$ be the limiting d.f. of Gnedenko defined in (1.3).

COROLLARY 3.1. *Suppose the distribution function F of ε_n has exponential form* (3.1); *then*

(a) *the distribution function of $\alpha M_n - \log(n\beta)$ converges to $H_2(x)$ and*

(b) *$P(M_n \le c_n) \to e^{-y}$ provided $c_n = \alpha^{-1}\log n - \alpha^{-1}\log(y/\beta) + o(1)$.*

4. First passage times-Pareto Case. Let $t = t_c = \inf\{n \ge 1 : |X_n| \ge c\}$ be the first passage of $|X_n|$ over the boundary c and set $\bar{t} = \bar{t}_c = \inf\{n \ge 1 : X_n \ge c\}$. For $\tau = t$ or $\bar{t}$ define $R_\tau = |X_\tau| - c$ and $\bar{R}_\tau = X_\tau - c$ to be the respective overshoots. In this section we calculate the asymptotic normalized joint distribution of the first passage time with its overshoot as the barrier $c \to \infty$. Denote convergence in distribution by $\to_{\mathscr{D}}$.

THEOREM 4.1. *Let X and Y be mutually independent random variables with distribution functions $1 - e^{-x}$ and $1 - (y + 1)^{-\alpha}$ respectively. If the distribution function G of $|\varepsilon_n|$ is Pareto, then as $c \to \infty$*

$$([1 - G(c)]t, c^{-1}R_t) \to_{\mathscr{D}} (X, Y).$$

PROOF. The notation in the proof of Theorem 2.1 is used. If $1 - G(x) = x^{-\alpha}L(x)$ set $d = d_c = c^{-\alpha}L(c)$ and $A_j = \{\bar{M}_{j-1} < c\}$. For positive x, y, and v

$$\begin{aligned} P(x \le td \le v, R_t > yc) &= \textstyle\sum P(\mid X_j \mid > (y+1)c, A_j) \\ &= \textstyle\sum \{P(\mid X_j \mid > c + yc, A_j) - G^{j-1}(c)[1 - G(c + yc)]\} \\ &\quad + (y+1)^{-\alpha} \textstyle\sum (1-d)^{j-1} d\{L(c + yc)/L(c)\} \end{aligned} \tag{4.1}$$

where the summation is over $x \le jd \le v$. The last summation converges as a Riemann sum to $(y + 1)^{-\alpha}(e^{-x} - e^{-v})$. A summand in the first summation of (4.1) is bounded by

$$\mid P(\mid X_j \mid > c + yc, A_j) - P(U_j > c + yc, A_j) \mid + \{1 - G(c + yc)\}\{P(A_j) - G^{j-1}(c)\}. \tag{4.2}$$

We can apply Remark 2.1 to obtain

$$\mid P(A_j) - G^{j-1}(c) \mid = o(1)$$

uniformly in j, for $xd^{-1} \le j \le vd^{-1}$. Since $1 - G(c + yc) = O(d)$, we need only consider the summation over the absolute difference in (4.2) which equals (2.3) with c replaced by $(y + 1)c$ and hence tends to zero as the proof there indicates. □

The following corollary extends the above results to $\bar{t}_c$.

COROLLARY 4.1. *If the distribution function F of ε_n is Pareto with dominant right tails then $\forall z \in \Lambda$, $([1 - F(c)]\bar{t}, c^{-1}\bar{R}_{\bar{t}})$ and $([1 - G(c)]t, c^{-1}R_t)$ have the same limiting distribution under P_z.*

PROOF. *If* $1 - F(c) = c^{-\alpha}L(c) \sim 1 - G(c)$, let Δ be the symmetric difference of $\{\bar{t}c^{-\alpha}L(c) \le x, \bar{R}_{\bar{t}}c^{-1} > y\}$ with $\{tc^{-\alpha}L(c) \le x, R_t c^{-1} > y\}$. Set $m_n = \min\{X_1, \cdots, X_n\}$ and choose K to satisfy Lemma 2.1 so that

$$P_z(\Delta) \le P_z(tc^{-\alpha}L(c) \le x, t \le \bar{t}) = P_z(m_t \le -c, M_t \le c, tc^{-\alpha}L(c) \le x)$$
$$= \sum E_z[F(-c - a'Z_{j-1}); \bar{M}_{j-1} < c]$$
$$\le [xc^{\alpha}/L(c)]F(-c + c \mid a \mid + K) = o(1)$$

where the summation is over $1 \le j \le xc^{\alpha}/L(c)$. □

REMARK 4.1. In the above corollary one can interchange t and $\bar{t}$ *or insert any of* R_t, $R_{\bar{t}}$, $\bar{R}_t$, or $\bar{R}_{\bar{t}}$ for R_t or $\bar{R}_{\bar{t}}$ without changing the limiting distribution.

The next corollary extends Theorem 4.1 to convergence in expectation.

COROLLARY 4.2. *If* $|\varepsilon_n|$ *has a Pareto distribution function G with exponent* $\alpha > 1$ *then for* $z \in \Lambda$

a. $E_z(t) = 1 - G(c) + o[1 - G(c)]$
b. $E_z(R_t) = c/(\alpha - 1) + o(c)$ *and hence* $E_z(X_t) = c\alpha/(\alpha - 1) + o(c)$.

PROOF. In the light of Theorem 4.1 it suffices to show uniform integrability (u.i.). To see $[1 - G(c)]t$ is u.i. let K satisfy Lemma 3.1, $y \ge 0$ and set $d = d_c = 1 - G(c + |a| c + K)$. Since $1 - G(c) = o(d)$, it suffices to show td is u.i. If $A_j = \{\bar{M}_j < c\}$ and $P = P_z$,

$$P(t > j) = P(A_j) \le E[G(c + | a'Z_{j-1} |); A_j] \le (1 - d)P(A_{j-1})$$

so that $P(t > j) < (1 - d)^j$. Hence, if $\gamma = y^{d^{-1}}$, $dE(t; td > y) \le yP(td > y) + (1 + d)^{\gamma} \to e^{-y}(y + 1)$ uniformly in y as $c \to \infty$.

To show $c^{-1}R_t$ is u.i. it suffices to show the u.i. of $c^{-1}\bar{M}_t$. For $y \ge 2$

$$E[\bar{M}_t; \bar{M}_t > yc] = \sum_1^{\infty} E[E(| X_n | I_{\{|X_n| > yc\}} \mid \mathscr{F}_n); t > n - 1].$$

After integration by parts, the integrand on $\{t > n - 1\}$ is bounded by

$$yc[1 - G(yc - 2 \mid a \mid c)] + \int_{yc-2|a|c}^{\infty} 1 - G(x)\, dx$$

for sufficiently large c. Hence

$$c^{-1}E(\bar{M}_t; \bar{M}_t > yc) \le [1 - G(c)]E(t)o(1)$$

where the representation theorem for slowly varying functions (Feller, 1966, page 282) indicates $o(1) \to 0$ as $y \to \infty$ uniformly in c. □

REMARK 4.2. Of course if $P_z(\varepsilon_k < 0) = 0$ and each $a_k \ge 0$ then Corollary 4.2 holds with t, R_t, and G replaced by $\bar{t}$, $\bar{R}_{\bar{t}}$, and the d.f. F of ε_n. The author conjectures the validity of this replacement when the right tail of F dominates. This conjecture follows if, after normalization, the variables $\bar{t}$ and $\bar{R}_{\bar{t}}$ are u.i.

5. First passage times—Exponential Case. In this section the asymptotic joint distribution of the first passage time $t = t_c = \inf\{n \ge 1 : X_n \ge c\}$ and the remainder $R_t = X_t - c$ is calculated when ε_k has the dominant right tail exponential d.f. (3.1). Note the change in t and R_t from the Paretian case. Conditions implying u.i. and hence convergence in mean are also given. Replacing d_c by $e^{-\alpha c}$, G by the d.f. H defined in (3.2), $\bar{M}_n$ by M_n, and adhering to the proof of Theorem 4.1 creates the following.

THEOREM 5.1. *If* ε_k *has the distribution function of* (3.1) *then* $(\beta e^{-\alpha c}t, \alpha R_t) \to_D (X, Y)$ *where X and Y are* i.i.d. *with distribution function* $1 - e^{-x}$.

The analogue of Corollary 4.2 is the following.

COROLLARY 5.1. *Suppose* $a_k \geq 0 \ \forall \ k$ *and* ε_k *has distribution function* (3.1) *with* $F(0) = 0$; *then* $e^{-\alpha c}Et \to 1/\beta$. *If, in addition, there exists N such that* $a_k = 0 \ \forall \ k > N$ *then* $\alpha ER_t \to 1$ *and* $EX_t = c + \alpha^{-1} + o(1)$.

The proof entails showing the u.i. of $e^{-\alpha c}t$ and R_t and is similar to the proof of Corollary 4.2.

REMARK 5.1. The author conjectures the validity of the corollary for arbitrary $\{a_k\}$ and $F(0)$. It would also be valuable to know that the results obtained for exponential distributions in Sections 3 and 5 hold under each P_z.

Acknowledgment. The author is thankful to Professor Michael Woodroofe for many helpful discussions and to the referee for a careful reading.

REFERENCES

[1] BERMAN, SIMEON M. (1962). Limiting distribution of the maximum term in sequences of dependent random variables. *Ann. Math. Statist.* **33** 894–908.
[2] BERMAN, SIMEON M. (1964). Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35** 502–516.
[3] CHERNICK, MICHAEL R. (1978). Unpublished Ph.D. thesis. Department of Statistics, University of Stanford.
[4] CHERNICK, MICHAEL R. (1981). A limit theorem for the maximum of autoregression processes with uniform marginal distributions. *Ann. Probability* **9** 145–149.
[5] FELLER, WILLIAM (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.
[6] FINSTER, MARK (1982). Optimal stopping on autoregressive schemes. *Ann. Probability* **10** 745–753.
[7] GALAMBOS, JANOS (1972). On the distribution of the maximum of random variables. *Ann. Math. Statist.* **43** 516–521.
[8] GALAMBOS, JANOS (1978). *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New York.
[9] GNEDENKO, B. V. (1943). Sur la distribution limite du terme maximum d'un série al éatorie. *Ann. Math.* **44** 423–453.
[10] LEADBETTER, M. R. (1974). Lectures on extreme value theory. Tech. Report Dept. Math. Stat., Lund University, Sweden.
[11] LEADBETTER, M. R. (1974). On extreme values in stationary sequences. *Z. Wahrsch. verw. Gebiete* **28** 289–303.
[12] LEADBETTER, M. R. (1975). Aspects of extreme value theory for stationary processes—a survey. In *Stochastic Processes and Related Topics, Vol 1.* Academic, New York.
[13] LEADBETTER, M. R., LINDGREN, G., and ROOTZÉN, H. (1979). Extremal and related properties of stationary processes. Statistical research report No. 1979-2. University of Umea, Sweden.
[14] LÉVY, P. (1937). *Théorie de l'Addition des Variables Altéatories.* Gauthier-Villars, Paris.
[15] LOÈVE, MICHAEL (1960). *Probability Theory.* Van Nostrand, Princeton.
[16] LOYNES, R. M. (1965). Extreme values in uniformly mixing stationary stochastic process. *Ann. Math. Statist.* **36** 993–999.
[17] O'BRIEN, G. L. (1974a). Limit theorems for the maximum term of a stationary process. *Ann. Probability* **2** 540–545.
[18] O'BRIEN, G. L. (1974b). The maximum term of uniformly mixing stationary processes. *Z. Wahrsch. Verw. Gebiete* **30** 57–63.
[19] ROSENBLATT, MURRAY (1962). *Random Processes.* Oxford, New York.
[20] ROOTZÉN, HOLGAR (1978). Extremes of moving averages of stable processes. *Ann. Probability* **6** 847–869.
[21] WATSON, G. S. (1954). Extreme values in samples from m-dependent stationary stochastic processes. *Ann. Math. Statist.* **25** 798–800.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
1210 W. DAYTON ST.
MADISON, WISCONSIN 53706