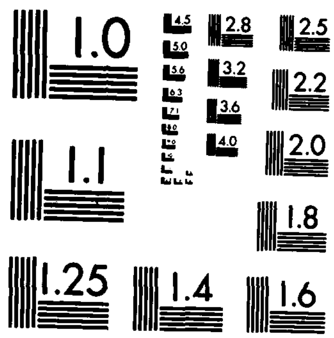
AD-A125 606 EXTREME VALUES OF NON-STATIONARY SEQUENCES AND THE EXTREMAL INDEX(U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PROC.. J HUSLER OCT 82 TR-20 1/1

UNCLASSIFIED AFOSR-TR-83-0074 F49620-82-C-0009 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

EXTREME VALUES OF NON-STATIONARY SEQUENCES AND THE EXTREMAL INDEX

by

Jürg Hüsler

TECHNICAL REPORT #20

October 1982

unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0074	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EXTREME VALUES OF NON-STATIONARY SEQUENCES AND THE EXTREMAL INDEX		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER TR 20
7. AUTHOR(s) Jürg Hüsler		8. CONTRACT OR GRANT NUMBER(s) F49620-82-C-0009
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics, University of North Carolina, 315 Phillips Hall 039 A Chapel Hill, NC 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR/NM Bolling AFB Washington, DC 20332		12. REPORT DATE October 1982
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release-- distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Extremes, non-stationary processes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The conditions used to generalize the extreme value theory for stationary random sequences to non-stationary sequences are studied with respect to their necessity. We find that the extremal index, defined in the stationary case, plays a similar role in the non-stationary case. The details show that this index describes not only the behavior of exceedances above a high level constant boundary, but also above a non-constant high level boundary.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

EXTREME VALUES OF NON-STATIONARY SEQUENCES AND THE EXTREMAL INDEX

by

Jurg Hüsler*

University of Bern
and
University of North Carolina at Chapel Hill

Summary: The conditions used to generalize the extreme value theory for stationary random sequences to non-stationary sequences are studied with respect to their necessity. ~~We~~ The authors find that the extremal index, defined in the stationary case, plays a similar role in the non-stationary case. The details show that this index describes not only the behavior of exceedances above a high level constant boundary, but also above a non-constant high level boundary.

Keywords: Extremes, non-stationary processes.

*Research supported by the Air Force Office of Scientific Research Grant No. AFOSR F49620-82C0009.

This research was done while the author was at the Center for Stochastic Processes at Chapel Hill. The author wishes to express his gratitude to Chapel Hill for the financial support and hospitality.

1. Introduction.

Let $\{X_i, i \geq 1\}$ be a random sequence with identical marginal distribution $F(x) = P\{X_i \leq x\}$ for all i. We deal with the approximation of probabilities of the type:

$$P_n = P\{X_i \leq u_{ni} \ , \ i \leq n\}$$

as $n \to \infty$, where $\{u_{ni} \ , \ i \leq n, n \geq 1\}$ is considered as the real-valued boundary.

In the case $u_{ni} \equiv u_n$ for all $i \leq n$, this probability gives the distribution of the partial maxima $M_n = \max\{X_1, \ldots, X_n\}$. The classical extreme value theory discusses the possible asymptotic distribution of M_n as $n \to \infty$, where X_i are i.i.d. r.v., i.e.,

$$P\{a_n(M_n - b_n) \leq x\} = [F(u_n(x))]^n \to G(x)$$

where $G(x)$ is one of the three known extreme value type distributions and a_n, b_n norming values, $u_n(x) = x/a_n + b_n$.

It was shown that the same result remains true even if X_i is a stationary sequence satisfying weak dependence restrictions (see e.g. Leadbetter [3] or Leadbetter, Lindgren and Rootzén [5]). To prove this result it was shown that

$$P\{M_n \leq x/a_n + b_n\} = [F(u_n(x))]^n + o(1) \ .$$

The same argument was used for the general case of a non-stationary random sequence, i.e. it was shown in [2] that

$$(1.1) \qquad P\{X_i \leq u_{ni}, i \leq n\} = \prod_{i=1}^{n} F(u_{ni}) + o(1) \quad \text{as } n \to \infty$$

under suitable conditions.

We remark that the studied probabilities covers also the extreme value case for non-stationary sequences, by transforming $P\{M_n \leq u_n\}$ into $P\{X_i \leq u_{ni}\}$. More detailed, e.g. if $\tilde{X}_i$ is any normal non-stationary sequence, with $\mu_i = EX_i$, $\sigma_i^2 = \text{Var } X_i$, $\tilde{M}_n = \max\{\tilde{X}_1, \ldots, \tilde{X}_n\}$, then

$$P\{\tilde{M}_n \le u_n\} = P\{\tilde{X}_i \le u_n, i \le n\} = P\{(\tilde{X}_i - \mu_i)/\sigma_i \le (u_n - \mu_i)/\sigma_i, i \le n\}$$
$$= P\{X_i \le u_{ni}, i \le n\}$$

where $u_{ni} = (u_n - \mu_i)/\sigma_i$, $F(x) = \Phi(x)$ the standard normal law and X_i a standardized normal non-stationary sequence.

Define $x_0 = \sup\{x: \ F(x)<1\} \le \infty$ and let $F(x_0-) = 1$. We suppose throughout the paper that $u_{min} = u_{min}(n) = \min\{u_{ni}, i \le n\} \to x_0$ as $n \to \infty$. Furthermore we restrict our attention to the interesting case where $\Pi F(u_{ni})$ tends to a value different from 0 or 1.

The sufficient conditions used in proving (1.1) are as follows:

Condition A: Let $F_n = F_n(u_{ni}) = \sum_{i=1}^{n} \bar{F}(u_{ni})$ with $\bar{F}(x) = 1 - F(x)$. Then assume

$$\limsup_{n\to\infty} F_n < \infty \quad \text{and} \quad \liminf_{n\to\infty} F_n > 0 \tag{1.2}$$

The dependence restrictions are

Condition D $(=D(u_{ni}))$: For any integers $1 \le i_1 < i_2 < \ldots < i_p < j_1 < j_2 < \ldots < j_q \le n$ for which $j_1 - i_p \ge m$, let $I = \{i_\ell, \ell=1,\ldots, p\}$, $J = \{j_\ell, \ell=1,\ldots,q\}$, $B(I) = \{X_i \le u_{ni}, i \in I\}$ and similar $B(J)$. Then we assume $\sup_{I,J} |P(B(I \cap J)) - P(B(I)) \cdot P(B(J))| \le \alpha_{n,m}$ where $\alpha_{n,m_n^*} \to 0$ as $n \to \infty$ for some sequence m_n^* such that

$$m_n^* \bar{F}(u_{min}) \to 0 \quad \text{as} \quad n \to \infty . \tag{1.3}$$

Condition D' $(=D'(u_{ni}))$: Let n,r be integers and I a subset of $\{1,\ldots, n\}$ of the form $\{i_1 \le i \le i_2\}$ such that $\sum_{i\in I} \bar{F}(u_{ni}) \le F_n/r$. Then assume that

$$\max_I \min_{I^* \subset I} \sum_{i<j\in I^*}\sum P\{X_i > u_{ni}, X_j > u_{nj}\} \le \alpha^*_{n,r} \tag{1.4}$$

such that

$$\lim_{r\to\infty} \limsup_{n\to\infty} r\, \alpha^*_{n,r} = 0 \tag{1.5}$$

where the min in (1.4) is considered on subsets I* of I with

(1.6) $$\sum_{i\in I-I^*} \bar{F}(u_{ni}) \le g(r)/r \quad \text{for all} \quad n \ge n_0(r)$$

and $$g(r) \to 0 \quad \text{as} \quad r \to \infty .$$

These conditions are sufficient for (1.1) (see Theorem 2.2 of [2]) and in addition we got that if

(1.7) $$F_n \to \tau \quad \text{as} \quad n \to \infty$$

then

(1.8) $$P_n \to e^{-\tau} \quad \text{as } n \to \infty , \text{ for } \tau > 0 .$$

The purpose of this paper is to discuss the necessity of the three conditions. In Section 2 we show mainly that the conditions D and D' with (1.8) imply (1.7). In Section 3 we assume only the Condition D and (1.7) and find that the possible limits of P_n may still be described. This gives us the relation to the stationary case of the extreme value theory and to the extremal index, defined in this context by Leadbetter [4].

2. Necessity of Condition A.

We consider in this section the equivalence of (1.7) and (1.8) if the conditions D and D' hold for a given random sequence $\{X_i\}$ and a boundary $\{u_{ni}\}$. Since we have shown in [2] that (1.7) implies (1.8), it remains to prove the converse. It suffices to prove only that liminf $F_n > 0$ and limsup $F_n < \infty$, since by the first part of Theorem 2.2 [2]: $P_n - e^{-F_n} = o(1)$, which implies (1.7) by (1.8). This proof uses the same technique as in [2]. Since the same technique is used also in Section 3, we mention some of the results of [2] in detail.

Lemma 2.1. Let n,r be fixed integers and $I_1, \ldots, I_r$ intervals of $\{1, \ldots, n\}$ such that I_i and I_j are separated by at least m for $i \ne j$. Suppose Condition D holds for a given boundary $\{u_{ni}\}$. Then

$$\left|P\left(\bigcap_{i=1}^{r} B(I_i)\right) - \prod_{i=1}^{r} P(B(I_i))\right| \le r\,\alpha_{n,m} .$$

We use the following construction. Split the set $\{1,\dots, n\}$ into intervals I_ℓ, $\ell = 1,\dots, r$, such that $I_1 = \{1,\dots, i_1\}$ with

$$F_{n,1} = \sum_{i=1}^{i_1} \overline{F}(u_{ni}) \le F_n/r$$

and

$$F_{n,1} + \overline{F}(u_{n,i_1+1}) > F_n/r$$

(i.e. i_1 is chosen as large as possible). Let $I_2 = \{i_1+1,\dots, i_2\}$ such that

$$F_{n,2} = \sum_{i=i_1+1}^{i_2} \overline{F}(u_{ni}) \le F_n/r$$

with i_2 maximally chosen. By repeating this procedure r times, we find intervals I_ℓ with $i_r \le n$,

(2.1) $$F_{n,\ell} = \sum_{i\in I_\ell} \overline{F}(u_{ni}) \le F_n/r$$

and

(2.2) $$\sum_{\ell=1}^{r} F_{n,\ell} = \sum_{i=1}^{i_r} \overline{F}(u_{ni}) \le F_n \ .$$

Furthermore let $0 < \varepsilon < 1$. Split each interval I_ℓ into two subintervals $I_{\ell,1}$ and $I_{\ell,2}$ where

$$I_{\ell,2} = \{i_\ell - m_{\ell+1},\dots, i_\ell\}$$

contains the last m_ℓ points of I_ℓ, $I_{\ell,1}$ the remaining points such that

$$\sum_{i\in I_{\ell,2}} \overline{F}(u_{ni}) \le F_n\ \varepsilon/r$$

and m_ℓ is maximally chosen.

We proved in [2] that since $\overline{F}(u_{min}) \to 0$ as $n \to \infty$

(2.3) $$m_\ell + 1 \ge F_n\ \varepsilon/r\cdot\overline{F}(u_{min}) \ .$$

Lemma 2.2. If $\varepsilon = \varepsilon(n)$ with

(2.4) $$\varepsilon(n) \cdot F_n \to 0 \quad \text{as} \quad n \to \infty ,$$

then for any integer r

$$P_n - P\{X_i \le u_{ni}, i \in \bigcup_{\ell=1}^{r} I_{\ell,1}\} \to 0 \quad \text{as } n \to \infty .$$

Proof: We have

$$0 \le P\{X_i \le u_{ni}, i \in \cup I_{\ell,1}\} - P_n \le \sum_{\ell=1}^{r} \sum_{i \in I_{\ell,2}} \overline{F}(u_{ni}) + \sum_{i=i_r+1}^{n} \overline{F}(u_{ni}) .$$

The first term is bounded by $r \cdot \varepsilon(n)F_n/r = \varepsilon(n) \cdot F_n \to 0$, using the construction of $I_{\ell,2}$ and (2.4). A simple argument showed in [2] that the second term is bounded by $r \cdot \overline{F}(u_{min}) \to 0$ as $n \to \infty$ for all r. □

Lemma 2.3. i) (1.8) implies $\liminf_{n\to\infty} F_n > 0$

ii) $\varepsilon(n) = (m_n^*+1)\overline{F}(u_{min})r/F_n \to 0$ and satisfies (2.4), where m_n^* is given by Condition D.

Proof: i) Since $P_n = 1 - P\{\exists i: X_i > u_{ni}\} \ge 1 - F_n$ we have $F_n > 1 - P_n$, but $P_n \to e^{-\tau} < 1$. Thus $\liminf F_n > 0$.

ii) The given $\varepsilon(n)$ satisfies

$$\varepsilon(n) \cdot F_n = (m_n^*+1)\overline{F}(u_{min}) \cdot r \to 0 \quad \text{for any } r \text{ by Condition D.}$$

By i) we have $\varepsilon(n) \le K \cdot (m_n^*+1)\overline{F}(u_{min}) \cdot r \to 0$ for a suitable constant K. □

Lemma 2.4. If (1.8), Condition D and D' hold, then $\limsup_{n\to\infty} F_n < \infty$.

Proof: Lemma 2.1 and 2.2 with the chosen $\varepsilon(n)$ imply that

$$\limsup_{n\to\infty} |P_n - \Pi P(B_{\ell,1})| \to 0 \quad \text{as } r \to \infty ,$$

with $B_{\ell,1} = B(I_{\ell,1})$. Thus by using (1.8) and $\log(1-x) \le -x$, it implies

(2.4) $$\limsup_{n\to\infty} \sum_{\ell=1}^{r} (1-P(B_{\ell,1})) < K \quad \text{for any } r.$$

We proved in [2] the inequality

$$(2.5)\qquad S_{n,r} := \sum_{\ell=1}^{r} \sum_{\ell\in I_{\ell,1}} \overline{F}(u_{ni}) \geq \sum_{\ell=1}^{r} (1-P(B_{\ell,1})) \geq S_{n,r} - g(r) - r\,\alpha^*_{n,r}$$

by using Condition D'.

As in Lemma 2.2 we find

$$(2.6)\qquad F_n \leq S_{n,r} + \varepsilon(n)\cdot \overline{F}(u_{min}) + r\overline{F}(u_{min}) .$$

Finally, combining (2.5) and (2.6) and using (2.4) gives the desired result. □

Thus we proved

Theorem 2.5. Let $\{X_i, i\geq 1\}$ be a random sequence with identical marginal distribution $F(x)$ and $\{u_{ni}, i\leq n, n\geq 1\}$ a real-valued boundary. Assume that the conditions D and D' hold together with $u_{min} \to x_0$ as $n\to\infty$. Then for $\tau > 0$

$$F_n \to \tau \quad \text{as } n\to\infty$$

is equivalent to

$$P_n = P\{X_i \leq u_{ni}, i\leq n\} \to e^{-\tau} \quad \text{as } n\to\infty .$$

Next we state an easy consequence of Theorem 2.5 for the case where we consider only a subset of $\{1,\ldots, n\}$ in the probability P_n.

Corollary 2.6. If in addition to Theorem 2.5, $I_n \subset \{1,\ldots, n\}$ such that

$$(2.7)\qquad F_n(I_n) := \sum_{i\in I_n} \overline{F}(u_{ni}) \to \tau' \quad \text{as } n\to\infty,\ \tau'\leq\tau,$$

then

$$P_n(I_n) := P\{X_i \leq u_{ni}, i\in I_n\} \to e^{-\tau'} \quad \text{as } n\to\infty.$$

Proof: It remains to prove that the Condition D and D' hold with respect to the "new" boundary

$$\tilde{u}_{ni} = \begin{cases} u_{ni} & i\in I_n \\ x_0 & i\notin I_n \end{cases}$$

Condition D holds for $\tilde{u}_{ni}$ since $\tilde{B}(I) = \{X_i \leq \tilde{u}_{ni}, i\in I\} = B(I\cap I_n)$ for any $I\subset\{1,\ldots, n\}$ and $\overline{F}(\tilde{u}_{min}) \leq \overline{F}(u_{min})$. Condition D' holds in an analogous way: Let I be a subinterval of $\{1,\ldots, n\}$ with $\sum_{i\in I} \overline{F}(\tilde{u}_{ni}) \leq F_n(I_n)/r$. Then also $\sum_{i\in I} \overline{F}(\tilde{u}_{ni}) = \sum_{i\in I\cap I_n} \overline{F}(u_{ni}) \leq F_n/r$. Thus there exists a subset I* satisfying

$$(2.8)\qquad \sum_{i<j\in I^*}\sum P\{X_i > u_{ni}, X_j > u_{nj}\} \leq \alpha^*_{n,r}$$

by $D'(u_{ni})$. But the l.h.s. of (2.8) is larger than

$$\sum_{i<j\in I^*\cap I_n}\sum P\{X_i>u_{ni}, X_j>u_{nj}\} = \sum_{i<j\in I^*}\sum P\{X_i>\tilde{u}_{ni}, X_j>\tilde{u}_{nj}\}$$

Since also $\sum_{i\in I-I^*} \overline{F}(\tilde{u}_{ni}) \le \sum_{i\in I-I^*} \overline{F}(u_{ni})$, the condition $D'(\tilde{u}_{ni})$ holds with the same values $\alpha^*_{n,r}$ and $g(r)$. □

From this it is obvious that the Poisson limit result in [2] for the number of exceedances $N_n(I) = \#\{i\in I: X_i>u_{ni}\}$, with $I = \{1,\ldots, n\}$, generalizes for any sequence of subsets I_n, i.e.

$N_n(I_n)$ has an asymptotic Poisson distribution with parameter τ', if (2.7) holds in addition to the Condition D, D' and (1.7).

3. Results under Condition D.

With the construction and results of Section 2 we discuss now the asymptotic behavior of P_n without assuming Condition D'. In the stationary extreme value case it was shown by Leadbetter [4] that if $u_n(\tau)$ is such that (1.7) holds and the Condition D is satisfied for a particular $u_n(\tau_o)$, $\tau_o > 0$, then there exist constants θ, θ', $0 \le \theta \le \theta' \le 1$ such that

$$\limsup_{n\to\infty} P\{X_i \le u_n(\tau), i\le n\} = e^{-\theta\tau}$$

(3.1)

$$\liminf_{n\to\infty} P\{X_i \le u_n(\tau), i\le n\} = e^{-\theta'\tau}$$

for all $0 < \tau \le \tau_0$. The notation $u_n(\tau) = u_n$ indicates the value τ used in (1.7). We remark that in the constant boundary case $u_n(\tau)$ in (3.1) is defined by e.g. $u_n(\tau) = u_{[n\tau_0/\tau]}(\tau_0)$; then $u_n(\tau)$ satisfies (1.7) for any $\tau>0$. Analogously we define now $u_{ni}(\tau)$ for any $0<\tau\le\tau_0$, if $u_{ni}(\tau_0)$ satisfies (1.7) for a $\tau_0>0$ as follows:

$$u_{ni}(\tau) = \begin{cases} u_{ni}(\tau_0) & i\le s \\ x_\cap & s<i\le n \end{cases} \tag{3.2}$$

where s is maximally chosen such that $\sum_{i\le s} \overline{F}(u_{ni}(\tau_0)) \le F_n \tau/\tau_0$. Naturally, $\sum_{i\le n} \overline{F}(u_{ni}(\tau)) \to \tau$, as $n \to \infty$, since $u_{min} \to x_0$. If P_n converges, then the value θ in (3.1) is called the extremal index, thus denoted generally

$$\theta = -\log \lim_{n\to\infty} P\{X_i \le u_{ni}(\tau_0), i \le u\}/\tau_0 \ .$$

In the stationary case with a constant boundary (3.1) shows that θ does not depend on τ_0. Since we do assume neither the stationarity of the random sequence nor the constancy of the boundary, we expect a greater variety of properties of θ as a simple example indicates.

Let $Y_1, Y_2, \ldots$ be an i.i.d. sequence with continuous marginal distribution F and normalization $u_n(\tau) = \overline{F}^{-1}(\tau/n), \tau>0$. Let

$$X_i = Y_{[(i+1)/2]}, i\ge 1 \ .$$

Then it is easily checked that $P\{X_i \le u_n(\tau), i\le n\} \to e^{-\tau/2}$ as $n \to \infty$, $\tau > 0$. Thus $\theta = 1/2$ for the fixed level boundary. Take now e.g.

$$u_{ni}(\tau) = \begin{cases} u_n(2\tau) & \text{for i odd} \\ x_0 & \text{for i even} \end{cases}$$

where x_0 is again the endpoint of F. Naturally $\sum_{i\le n} \overline{F}(u_{ni}(\tau)) \to \tau$ and

$$P\{X_i \le u_{ni}(\tau), i\le n\} = P\{Y_i \le u_n(2\tau), i \le [\tfrac{n+1}{2}]\} \to e^{-\tau} \ .$$

Thus $\theta = 1$ for this particular boundary. By defining other boundaries u^*_{ni} in a similar way, fixed for some i's and equal to x_0 for the remaining i's, we find other values $\theta<1$. The same fact holds even if we define X_i to be stationary by $P\{X_i = Y_{[(i+1)/2]}, i\ge 1\} = P\{X_i = Y_{[i/2]+1}, i\ge 1\} = 1/2$. For the same random sequence we show that θ may even depend on the given value τ for a non-smooth boundary. Let

$$u_{ni}(\tau_0) = \begin{cases} u_n(2\tau_0) & \text{i odd}, i\le n/2 \\ x_0 & \text{i even}, i\le n/2 \\ u_n(\tau_0) & n/2<i\le n \ . \end{cases}$$

Then the obvious calculations show that (1.7) holds with $\tau_0>0$ and $\theta = \theta(u_{ni}(\tau_0))=3/4$. But for $\tau\le\tau_0/2$ we have $s\le n/2$ in the definition (3.2) and thus $\theta = \theta(u_{ni}(\tau)) = 1$.

This shows that for particular stationary and non-stationary sequences the possible parameter θ depends strongly on the given boundary; i.e. $\theta = \theta(u_{ni})$. This dependency is not restricted to a particular extremal index $\theta = \theta(u_n(\tau)) = 1/2$ as in our example, for we may replace in the above example the i.i.d. sequence Y_i by sequences Y_i given in Chesnick [1], Rootzén [7] or de Haan in Leadbetter [4], where $\theta = \theta(u_n(\tau))$ may be any value ≤ 1.

On the other hand it is obvious that for i.i.d. sequences, we find the same parameter $\theta = 1$ for any boundary values satisfying (1.7), for any value $\tau > 0$. The following result shows that this is not only true for i.i.d. sequences.

Theorem 3.1. Let $\{X_i,i\ge1\}$ be a Gaussian sequence with identical marginal distribution $\Phi(x)$, the unit normal law. Assume that the correlation function $r(i,j)$ satisfies

$$\max_{|i-j|\ge n} |r(i,j)|\log n \to 0 \quad \text{as } n\to\infty , \tag{3.3}$$

then
$$P\{X_i\le u_{ni}(\tau), i\le n\} \to e^{-\tau} \quad \text{as } n\to\infty,\ \tau>0 ,$$

where $\{u_{ni}(\tau)\}$ is any boundary satisfying (1.7) for any value $\tau > 0$. Thus $\theta = 1$ for any boundary.

The proof of this result is mainly given in Hüsler [2]. This is a particular case of the more general statement in Corollary 2.6, since (3.3) implies D and D' for any boundary. From this one might argue that if $\theta = 1$ for a certain boundary, then $\theta = 1$ for any boundary as long as D holds. But the above example indicates that this is not true in general, and also that the stronger condition $D'(u_{ni})$ does not imply $D'(u^*_{ni})$ to hold for any other boundary u^*_{ni}.

In the following we discuss some properties of the extremal index for general cases. The first result shows that θ cannot be larger than 1.

Lemma 3.2: Let $\{X_i, i\geq 1\}$ be a random sequence with identical marginal distribution. If (u_{ni}) satisfies (1.7) for some value $\tau>0$ and D, then

$$\liminf_{n\to\infty} P\{X_i \leq u_{ni}, i\leq n\} \geq e^{-\tau}$$

Proof: We use the technique of Section 2 to define intervals I_ℓ, $\ell=1,\dots, r$ for r,n fixed. Then we know that

$$P\{X_i \leq u_{ni}, i\leq n\} - \prod_{\ell=1}^{r} P(B(I_\ell)) \to 0 \quad \text{as } n\to\infty$$

But $P(B(I_\ell)) \geq 1 - \sum_{i\in I_\ell} \overline{F}(u_{ni}) \to 1 - \tau/r$ as $n \to \infty$ by the construction of the I_ℓ's.

Thus
$$\liminf_{n\to\infty} P\{X_i \leq u_{ni}, i\leq n\} \geq (1-\tau/r)^r \to e^{-\tau} \quad \text{as } r \to \infty .$$ □

Now we prove that the extremal index, if existing, is equal for boundaries, which differ only slightly from each other.

Lemma 3.3: Let $\{X_i, i\geq 1\}$ and (u_{ni}) be as in Lemma 3.2 with $\tau > 0$. Let $\{u^*_{ni}\}$ be another boundary satisfying (1.7) with the same value τ. If for each n either

$$(3.4)\qquad u_{ni} \leq u^*_{ni} \quad \forall i\leq n \qquad \text{or} \qquad u_{ni} \geq u^*_{ni} \quad \forall i\leq n$$

then i) Condition D holds also with respect to u^*_{ni} ,

ii) If $\theta = \theta(u_{ni})$ exists, then $\theta(u^*_{ni}) = \theta$

Proof: i) Similar to the proof of Corollary 2.6 it is sufficient to show that $P(B(I)) - P(B^*(I)) \to 0$ as $n \to \infty$, for any $I\subset\{1,\dots, n\}$ where $B^*(I) = \{X_i \leq u^*_{ni}, i\in I\}$. But using (3.4) we have

$$0 \leq P(B^*(I))-P(B(I)) \leq \sum_{i\in I} (\overline{F}(u_{ni})-\overline{F}(u^*_{ni})) \leq \sum_{i=1}^{n} (\overline{F}(u_{ni}) -\overline{F}(u^*_{ni}))$$

which tends to 0 by (1.7), if $u_{ni} \leq u^*_{ni}$. The converse case holds in the same way.

ii) This follows as in i) by setting $I = \{1,\dots, n\}$, without use of Condition D. □

This indicates that there are classes of boundaries having the same extremal index, in case of existence. We give a description of such a class, more general than (3.4).

Let $\{u_{ni}\}$ be a given boundary. Then for another boundary $\{u^*_{ni}\}$ define $I_n = \{i \le n: \ u_{ni} \le u^*_{ni}\}$ and assume that either

$$\sum_{i \in I_n} \overline{F}(u_{ni}) = o(1) \quad \text{or} \quad \sum_{i \notin I_n} \overline{F}(u^*_{ni}) = o(1) \tag{3.5}$$

Theorem 3.4. Let $\{X_i, i \ge 1\}$ be a random sequence and $\{u_{ni}(\tau)\}$ a boundary satisfying (1.7) and $\tau > 0$. Let $\{u^*_{ni}(\tau)\}$ be another boundary satisfying (3.5) and (1.7) for the same value τ. Then

i) If $D(u_{ni})$ holds, then also $D(u^*_{ni})$.

ii) If $\theta = \theta(u_{ni})$ exists, then $\theta(u^*_{ni}) = \theta$.

Proof: Define in the case $\sum_{i \notin I_n} \overline{F}(u_{ni}) = o(1)$

$$\tilde{u}_{ni} = \begin{cases} u^*_{ni} & i \in I_n \\ x_0 & i \notin I_n \end{cases}$$

By the assumption (3.5)

$$0 \le \sum_{i \le n} \overline{F}(u^*_{ni}) - \sum_{i \le n} \overline{F}(\tilde{u}_{ni}) = \sum_{i \notin I_n} \overline{F}(u^*_{ni}) = o(1) \ .$$

Thus $(\tilde{u}_{ni})$ satisfies (1.7) with τ. By Lemma 3.3, Condition $D(\tilde{u}_{ni})$ holds and $\theta(\tilde{u}_{ni}) = \theta$, since $\tilde{u}_{ni} \ge u_{ni}$, for all $i \le n$. But also $\tilde{u}_{ni} \ge u^*_{ni}$ for all $i \le n$, thus the two statements of the lemma follow by using Lemma 3.3 again with $\tilde{u}_{ni}$ in place of u_{ni}. The proof for the case $\sum_{i \in I_n} \overline{F}(u_{ni}) = o(1)$ is similar, by defining

$$\tilde{u}_{ni} = \begin{cases} u_{ni} & i \notin I_n \\ x_0 & i \in I_n \end{cases} \ . \qquad \square$$

We now give a sufficient condition for the existence of the extremal index with respect to a smooth boundary. This generalizes the condition D'.

Let $$S_n^{(k)}(I) = \sum_{i_1<i_2<\ldots<i_k\in I} \cdots \sum P\{X_{i_1} > u_{ni_1}, X_{i_2} > u_{ni_2}, \ldots, X_{i_k} > u_{ni_k}\}, \quad k\geq 1$$

Then assume that for a value θ, $0\leq\theta\leq 1$,

(3.6) $$\limsup_{n\to\infty} \max_{I} \min_{I^*\subset I} |rS_n^{(2)}(I^*) - \tau_0(1-\theta)| \to 0 \text{ as } r\to\infty$$

and

$$\limsup_{n\to\infty} \max_{I} \min_{I^*} rS_n^{(3)}(I^*) \to 0 \text{ as } r\to\infty$$

where the max on I is taken over intervals $I = \{i_1 \leq i \leq i_2\} \subset \{1,\ldots, n\}$ with $F_n/r - \overline{F}(u_{min}) \leq \sum_{i\in I} \overline{F}(u_{ni}) \leq F_n/r$ and where the min on I^* is taken over subset $I^*\subset I$ such that $S_n^{(1)}(I-I^*) = \sum_{i\in I-I^*} \overline{F}(u_{ni}) \leq g(r)/r$ for all $n \geq n_0(r)$, $g(r)\to 0$ as $r\to\infty$.

<u>Theorem 3.5</u>. Let $\{X_i, i\geq 1\}$ be a random sequence with identical marginal distribution and $\{u_{ni}\}$ a boundary satisfying condition D, (1.7) and (3.6) for a θ and a $\tau_0>0$. Then $\theta(u_{ni}(\tau)) = \theta$ for all $0<\tau\leq\tau_0$, where $u_{ni}(\tau)$ is defined in (3.2).

<u>Proof</u>: We prove first $\theta(u_{ni}(\tau_0)) = \theta$. Define as in Section 2 for n,r fixed the intervals I_ℓ, $\ell=1,\ldots, r$. Condition D implies again that

$$P_n - \prod_{\ell=1}^{r} P(B(I_\ell)) \to 0 \quad \text{as } n\to\infty .$$

Now for each ℓ, there exists a I_ℓ^* such that

$$0 \leq P(B(I_\ell^*)) - P(B(I_\ell)) \leq S_n^{(1)}(I_\ell - I_\ell^*) \leq g(r)/r$$

for all $n\geq n_0(r)$ and by Bonferroni's inequality

$$P(B(I_\ell^*)) \leq 1 - S_n^{(1)}(I_\ell^*) + S_n^{(2)}(I_\ell^*)$$

$$P(B(I_\ell^*)) \geq 1 - S_n^{(1)}(I_\ell^*) + S_n^{(2)}(I_\ell^*) - S_n^{(3)}(I_\ell^*) .$$

Thus

$$1 - \frac{\tau_0\theta}{r} - \limsup_{n\to\infty} \max_{I} \min_{I^*} \left|S_n^{(2)}(I^*) - \frac{\tau_0(1-\theta)}{r}\right| - \limsup_{n\to\infty} \max_{I} \min_{I^*} S_n^{(3)}(I^*) - \frac{g(r)}{r}$$

$$\le \liminf_{n\to\infty} P(B(I_\ell)) \le \limsup_{n} P(B(I_\ell))$$

$$\le 1 - \frac{\tau_0\theta}{r} + \limsup_{n\to\infty} \max_{I} \min_{I^*} \left|S_n^{(2)}(I^*) - \frac{\tau_0(1-\theta)}{r}\right| + g(r)/r .$$

By the assumption (3.6) we have that

$$\left(1 - \frac{\theta\tau_0 + o(1)}{r}\right)^r \le \liminf_{n\to\infty} \prod_{\ell=1}^{r} P(B(I_\ell)) \le \limsup_{n\to\infty} \prod_{\ell=1}^{r} P(B(I_\ell))$$

$$\le \left(1 - \frac{\theta\tau_0 + o(1)}{r}\right)^r$$

where $o(1) \to 0$ as $r \to \infty$. Thus by letting $r \to \infty$ we have $\lim_{n\to\infty} P_n = e^{-\theta\tau_0}$.

ii) Let the sets I_ℓ be as in i) depending on τ_0; denote by $r' = [\tau r/\tau_0]$. Then by the definition of s in (3.2)

$$\bigcup_{\ell=1}^{r'} I_\ell \subset \{1, \ldots, s\} = J \subset \bigcup_{\ell=1}^{r'+1} I_\ell$$

But $\quad 0 \le P(B(\bigcup_{\ell=1}^{r} I_\ell)) - P(B(J)) \le S_n^{(1)}(I_{r'+1}) \le F_n/r \to 0 \quad$ as $r \to \infty$.

The proof in i) shows that

$$\left(1 - \frac{\theta\tau_0 + o(1)}{r}\right)^{r'} \le \liminf_{n\to\infty} P(B(\bigcup_{\ell=1}^{r'} I_\ell)) \le \limsup_{n\to\infty} P(B(\bigcup_{\ell=1}^{r'} I_\ell))$$

$$\le \left(1 - \frac{\theta\tau_0 + o(1)}{r}\right)^{r'} .$$

Thus for $r \to \infty$ we find by combining the above facts that

$$\lim_{n\to\infty} P\{X_i \le u_{ni}(\tau), i \le n\} = \lim_{n\to\infty} P(B(J)) = e^{-\theta\tau} . \qquad \square$$

We remark that we might use instead of $u_{ni}(\tau)$ defined in (3.2) any other boundary $u^*_{ni}(\tau)$, which is equal to $u_{ni}(\tau_0)$ on a certain interval J' and equal to x_0 on the complement of J', where J' such that $S_n^{(1)}(J') \sim F_n\tau/\tau_0$.

This theorem generalizes the result (3.1) known for stationary random sequence with a constant boundary to non-stationary sequences with respect to a non-constant, but smooth boundary. Together with Theorem 3.4 we know now that the extremal index $\theta = \theta(u_n)$ defined in the case of stationarity with a constant boundary holds to be the same value for a class of non-constant boundaries, which differ slightly (depending naturally on the finite-dimensional distributions) from the constant boundary.

References

[1] Chernick M.R. "A limit theorem for the maximum of autoregressive processes with uniform marginal distribution" Ann. Probab. 9, 145-149, 1981.

[2] Hüsler J. "Asymptotic approximation of crossing probabilities of random sequences 1981. To appear in Z. Wahrsch. verw. Geb.

[3] Leadbetter M.R. "On extreme values in stationary sequences" Z. Wahrsch. verw. Geb. 28, 289-303 1974.

[4] Leadbetter M.R. "Extremes and local dependence in stationary sequences" Tech. Rept. No. 12, 1982, Center for Stoch. Processes, Dept. of Stat., Univ. of North Carolina, Chapel Hill.

[5] Leadbetter M.R., Lindgren G., Rootzén H. "Extremes and related properties of random sequences and processes" to appear in Springer Statistics Series 1982.

[6] O'Brien G.L. "The maximum term of uniformly mixing stationary processes" Z. Wahrsch. verw. Geb. 30, 57-63 1974.

[7] Rootzén, H. "Extremes of moving averages of stable processes. Ann. Probab. 6, 847-869 1978.

4-

DT