Holger Rootzén

The Rate of Convergence of Extremes of Stationary Normal Sequences

Preprint
June
1981

8

Institute of Mathematical Statistics
University of Copenhagen

Holger Rootzén

THE RATE OF CONVERGENCE OF EXTREMES OF STATIONARY NORMAL SEQUENCES

Preprint 1981 No. 8

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

June 1981

Abstract

Let $\{\xi_t\}$ be a stationary normal sequence with zero means, unit variances, and covariances $r_t = E\,\xi_s\xi_{s+t}$, let $\{\hat{\xi}_t\}$ be independent and standard normal, and write $M_n = \max_{1 \leqq t \leqq n} \xi_t$, $\hat{M}_n = \max_{1 \leqq t \leqq n} \hat{\xi}_t$. In this paper we find bounds on $|P(M_n \leqq u) - P(\hat{M}_n \leqq u)|$ which are roughly of the order

$$(\frac{1}{n})^{(1-\rho)/(1+\rho)},$$

where ρ is the maximal correlation, $\rho = \sup\{0, r_1, r_2, \ldots\}$, and it is shown that, at least for m-dependent sequences, the bounds are of the right order. Further, bounds of the same order on the rate of convergence of the point processes of exceedances of one or several levels are obtained using a "representation" approach (which seems to be of rather wide applicability). As corollaries we obtain rates of convergence of several functionals of the point processes, including the joint distribution function of the k largest values amongst $\xi_1, \ldots, \xi_n$.

AMS 1980 subject classification: Primary 60G15, secondary 60F05, 60G10, 60F17.

Key words and phrases: Maxima of normal sequences, extreme values, rate of convergence.

Research supported in part by the Office of Naval Research under contract N 0014-75-C-0809.

1. Introduction and discussion of the results

The asymptotic theory of extremes of independent and of stationary normal sequences has found many applications as testified e.g. by the books by Gumbel (1958) and Leadbetter, Lindgren & Rootzén (1982) and the references therein. However, for practical use of asymptotic theory, it is important to know the rate of convergence. The aim of this paper is to study in some detail the rate of convergence in extremal results for *dependent* stationary normal sequences. For the independent case, the reader is referred to the papers by Hall (1979) and Nair (1981).

Let $\xi_1, \xi_2, \ldots$ be a stationary normal sequence, which for convenience will be assumed to have zero means and unit variances, and let $r_t = E\,\xi_s \xi_{s+t}$ be its covariance function. Further, let $\hat{\xi}_1, \hat{\xi}_2, \ldots$ be an "associated independent sequence", i.e. a sequence of independent standard normal variables, write $M_n = \max_{1 \leqq t \leqq n} \xi_t$, $\hat{M}_n = \max_{1 \leqq t \leqq n} \hat{\xi}_t$, and let Φ be the standard normal distribution function. The first main result of this paper (Theorem 3.1) is that, for u_n given by $n(1-\Phi(u_n)) = K$,

$$\sup_{u \geqq u_n} |P(M_n \leqq u) - P(\hat{M}_n \leqq u)| \leqq 4R_n \quad . \tag{1.1}$$

Here R_n depends on K and the covariances $\{r_t\}$ in a rather complicated way (given by (2.3) below). The leading term of R_n is determined by the largest covariance $\rho = \sup\{0, r_1, r_2, \ldots\}$ and for $\rho > 0$ it is of the order

$$\left(\frac{1}{n}\right)^{(1-\rho)/(1+\rho)} (\log n)^{-\rho/(1+\rho)} \quad ,$$

while for $\rho = 0$ the order is

$$\frac{1}{n}\log n \quad ,$$

which is improved to the order $1/n$ in the case when in addition only finitely many of the r_t's are non-zero. Next, it is shown that if only finitely many r_t's are non-zero and if $u_n \to \infty$ in such a way that $P(M_n \leqq u_n)$ converges to a non-trivial limit (or equivalently if $n(1-\Phi(u_n)) \to K$, for some constant $K > 0$), then

$$P(M_n \leqq u_n) - P(\hat{M}_n \leqq u_n) \sim e^{-K} R_n$$

if $\rho > 0$, and

$$P(M_n \leqq u_n) - P(\hat{M}_n \leqq u_n) \sim -e^{-K} R_n$$

if $\rho = 0$. (Here $A \sim B$ has the standard meaning that $A = B(1+o(1))$.) In particular this shows that the bound in (1.1) is of the right order, at least in these cases.

It is often instructive to consider u as a "level" and to study the exceedances of the level by $\{\xi_t\}$ - the connection with the maximum of course being that M_n is less than u if and only if there are no exceedances of u by $\xi_1,\ldots,\xi_n$. More generally, we will consider the time-normalized point processes of exceedances of r levels $u^{(1)} \geqq \ldots \geqq u^{(r)}$ by $\xi_1,\ldots,\xi_n$, defined for $j = 1,\ldots,r$ as $N_n^{(j)}(B) = \#\{t \geqq 1;\ t/n \in B,\ \xi_t > u^{(j)}\}$ for Borel sets $B \subset [0,1]$, where $\#\{\ldots\}$ is the cardinality of the set within brackets. The reader is referred to e.g. Kallenberg (1976) or to the appendix of Leadbetter, Lindgren & Rootzén (1979) for definitions and information about point processes. Further, we will write $N_n = (N_n^{(1)},\ldots,N_n^{(r)})$ and will consider it as a random variable in the appropriate product space. The second main result (Theorem 4.1) is a representation theorem for N_n. Let $\hat{N}_n$ be defined from

$\{\hat{\xi}_t\}$ in the same way as N_n is defined from $\{\xi_t\}$. Then we show, using an idea of Serfling (1976), that there exist versions of N_n and $\hat{N}_n$ such that

(1.2) $$P(N_n \neq \hat{N}_n) \leqq 16r^2 R_n ,$$

and similarly N_n is approximated by a vector of "successively more severely thinned Poisson processes". (It may deserve mention that this approach seems potentially useful also in connection with other problems than the one studied here.) One easy corollary concerns $M_n^{(k)}$, the k-th largest among $\xi_1,\ldots,\xi_n$ and is that

$$\sup_{u^{(1)} \geqq \ldots \geqq u^{(r)} \geqq u_n} |P(M_n^{(1)} \leqq u^{(1)},\ldots,M_n^{(r)} \leqq u^{(r)})$$
$$- P(\hat{M}_n^{(1)} \leqq u^{(1)},\ldots,\hat{M}_n^{(r)} \leqq u^{(r)})| \leqq 16r^2 R_n ,$$

where of course $\hat{M}_n^{(k)}$ is the k-th largest among $\hat{\xi}_1,\ldots,\hat{\xi}_n$.

Much of the interest in extremes of normal sequences has been centered on the double exponential limit of the distribution of M_n, i.e. that, for $a_n = (2 \log n)^{\frac{1}{2}}$, $b_n = a_n - \{\log \log n + \log 4\pi\}/(2a_n)$,

(1.3) $$P(a_n(M_n - b_n) \leqq x) \to \exp\{-e^{-x}\} , \quad n \to \infty ,$$

which is known to hold if $r_n \log n \to 0$, or in even more general circumstances, see Leadbetter, Lindgren & Rootzén (1978). The relation (1.3) can, to emphasize the connection with (1.1), be written as

$$P(M_n \leqq u_n) \to \exp\{-e^{-x}\} , \quad n \to \infty ,$$

for $u_n = u_n(x) = x/a_n + b_n$, and furthermore the same result holds also if a_n and b_n are replaced by different constants a_n', b_n', provided $a_n'/a_n \to 1$ and $a_n(b_n' - b_n) \to 0$. As was noted already by Fisher & Tippet (1928) the convergence in (1.3) is extremely slow in the independent case. This was made precise in Hall (1979) in the following way: if a_n and b_n are as above, then

$$(1.4) \quad P(a_n(\hat{M}_n - b_n) \leqq x) - \exp\{-e^{-x}\} \sim \frac{1}{16} e^{-x} \exp\{-e^{-x}\} \frac{(\log\log n)^2}{\log n},$$

while if $a_n' = 1/b_n'$ and b_n' is chosen to be the solution of $2\pi\, b_n' \exp(b_n'^2) = n^2$ then

$$(1.5) \quad C_1/\log n \leqq \sup_x |P(a_n'(\hat{M}_n - b_n') \leqq x) - \exp\{e^{-x}\}| \leqq C_2/\log n \quad ,$$

for some constants $0 < C_1 < C_2 \leqq 3$ and $n \geqq 3$. Further, Hall shows that the rate $1/\log n$ cannot be improved on by choosing a_n', b_n' differently. This rate of convergence seems unfortunate if one e.g. in a statistical analysis wants to approximate the distribution of M_n by the limiting double exponential distribution. However, from a computational point of view it does not pose any problems, since of course $P(M_n \leqq u) = \Phi(u)^n$ is quite simple to evaluate directly.

It is easily seen, by combining (1.1) with Hall's results (1.4), (1.5) that for dependent sequences the rate of convergence in (1.3), under appropriate conditions, is of the order $(\log\log n)^2/\log n$ or $1/\log n$, i.e. equally slow as for independent sequences. For dependent sequences the quantity $P(M_n \leqq u)$, however, is more difficult to evaluate, and perhaps the most interesting consequence of (1.1) is that it demonstrates that the approximation of $P(M_n \leqq u)$ by $P(\hat{M}_n \leqq u) = \Phi(u)^n$ is reasonably accurate, at least when the maximal covariance ρ is not too close to one. Similarly, (1.2) measures how well quite complicated probabilities, concerning the point processes, can be approximated by assuming independence.

The organization of the paper is as follows. Section 2 contains some notation and three "technical" lemmas in which most of the necessary estimates are proved. In Section 3 the elementary case, the speed of convergence of the distribution of the maximum, is treated in a fairly complete way. In the next section, Section 4, the representation theorem for the point processes of exceedances is established together with some corollaries, and finally, Section 5 contains a short discussion of possible avenues for finding improved approximations of the probabilities of interest.

2. Technical preliminaries

The estimates of this section contain some fairly involved constants, which for easy reference we will collect here. With notation as in the introduction, let

$$\rho = \sup\{0, r_1, r_2, \ldots\}$$

and, in case $\rho > 0$, let ν be the number of t's such that $r_t = \rho$. We will throughout, without further comment, assume that the supremum is attained so that $\nu \geqq 1$. In particular this is the case if $r_t \to 0$ as $t \to \infty$, and then also $\nu < \infty$. If $\rho = 0$ let $\nu \leqq \infty$ be the number of non-zero r_t's. For the second order terms, define ρ' to be the supremum for $t \geqq 1$ of the r_t's which satisfy $r_t \neq \rho$, if this quantity is positive, and zero otherwise, and let

(2.1) $$\varepsilon = 2\left(\frac{1}{1+\rho'} - \frac{1}{1+\rho}\right) = \frac{2(\rho-\rho')}{(1+\rho)(1+\rho')} .$$

Next, for $\rho \neq 0$, define

(2.2) $$c'(\rho) = \frac{(1+\rho)^{3/2}}{(1-\rho)^{1/2}}, \quad c''(\rho) = \frac{(2-\rho)(1+\rho)}{1-\rho}, \quad c(\rho) = c'(\rho)(4\pi)^{-\rho/(1+\rho)}$$

and put $\delta = \sup\{|r_1|, |r_2|, \ldots\}$. The main factor, R_n, in the bounds has a slightly different appearance in the three cases (i) $\rho > 0$, (ii) $\rho = 0$, $\nu = \infty$, and (iii) $\rho = 0$, $\nu < \infty$, and in addition depends on a constant K, which will be introduced below,

(2.3)

(i) $$R_n = c(\rho) K^{2/(1+\rho)} \left(\frac{1}{n}\right)^{(1-\rho)/(1+\rho)} (\log n/K)^{-\rho/(1+\rho)} \{\nu + r_n\}$$

where $r_n = 16K^{\varepsilon}(1-\delta^2)^{-1/2} \sum_{t=1}^{n} |r_t| \left(\frac{1}{n}\right)^{\varepsilon} (\log n/K)^{1+\varepsilon/2}\}$, if $\rho > 0$,

(ii) $R_n = 4K^2(1-\delta^2)^{-1/2} \frac{1}{n} \log n \sum_{t=0}^{n} |r_t|$, if $\rho = 0$, $\nu = \infty$

(iii) $R_n = K^2 \frac{1}{n} \nu$, if $\rho = 0$, $\nu < \infty$.

As a starting point for the estimates we will use the important identity

(2.4) $$P(M_n \leqq u) - P(\hat{M}_n \leqq u)$$
$$= \sum_{1 \leqq s < t \leqq n} r_{s-t} \int_{h=0}^{1} \underline{\int_{-\infty}^{u'} \ldots \int} f_h(x_s = x_t = u) \underline{dx}' dh \quad ,$$

where $f_h(x_s = x_t = u)$ is the function of $n-2$ variables which is obtained by putting $x_s = x_t = u$ in the density function of n stationary normal random variables with zero means, unit variances, and covariances $h\, r_t$, and where the "primes" signify that x_s and x_t are deleted from the integrations, the intervals of integration each being $(-\infty, u]$. The equation (2.4) is due, in various ways, to Slepian (1962), Berman (1964), and Cramér, and for a derivation of it see e.g. Leadbetter, Lindgren & Rootzén (1979), p. 45-47. It is useful to write it in a slightly different form. Let

$$\psi_r(u) = \frac{1}{2\pi(1-r^2)^{1/2}} \exp\{-\frac{1}{2(1-r^2)}(u^2 - 2ru^2 + u^2)\}$$
$$= \frac{1}{2\pi(1-r^2)^{1/2}} \exp\{-\frac{u^2}{1+r}\}$$

be the joint density of two standard normal variables with correlation r, evaluated at the point (u,u), and let

$$f_h(x' | x_s = x_t = u) = f_h(x_s = x_t = u)/\psi_{hr_{s-t}}(u)$$

be the conditional density, given that the s-th and t-th variables equal u, in a n-dimensional normal distribution with zero means, unit variances and covariances $h\, r_t$. The identity (2.4) can then be written as

(2.5) $P(M_n \leqq u) - P(\hat{M}_n \leqq u)$

$$= \sum_{1 \leq s < t \leq n} r_{s-t} \int_{h=0}^{1} \psi_{hr_{s-t}}(u) \int_{-\infty}^{\underline{u}'} \cdots \int f_h(x' | x_s = x_t = u)\, d\underline{x}'\, dh \;.$$

Since $f_h(x' | x_s = x_t = u)$ is a density function it is clear that

(2.6) $$0 \leqq \int_{-\infty}^{\underline{u}'} \cdots \int f_h(x' | x_s = x_t = u)\, dx' \leqq 1 \;.$$

The main proofs use the right-hand inequality in (2.6) to estimate the expression in (2.5). However, we will also see that often not much is lost by this.

<u>Lemma 2.1</u> Let $u > 0$, suppose $r \neq 0$, $|r| < 1$, write $\rho = \max\{0, r\}$ and let c, c' be given by (2.2). Then

(i) $$\frac{1}{2\pi u^2 r} \{c'(r) e^{-u^2/(1+r)} - e^{-u^2}\} / \{1 + \frac{c''(\rho)}{u^2}\} \leqq \int_0^1 \psi_{hr}(u)\, dh$$

$$\leqq \frac{1}{2\pi u^2 r} \{c'(r) e^{-u^2/(1+r)} - e^{-u^2}\} \;.$$

Suppose that furthermore $u \geqq 1$. Then

(ii) $$0 \leqq \int_0^1 \psi_{hr}(u)\, dh$$

$$\leqq 2^{(2+\rho)/(1+\rho)} c(\rho) |r|^{-1} \{(1 - \Phi(u))/u^{\rho}\}^{2/(1+\rho)}$$

and, if $r \leqq \rho'$ for some constant $0 \leqq \rho' < 1$, then

(iii)

$$0 \leqq \int_0^1 \psi_{hr}(u)\, dh$$

$$\leqq 2^{(2+\rho')/(1+\rho')} (4\pi)^{-\rho'/(1+\rho')} (1-r^2)^{-1/2} \{(1 - \Phi(u))u\}^{2/(1+p')} \;.$$

Proof By partial integration

$$(2.7)\qquad 2\pi\int_0^1 \psi_{hr}(u)\,dh = \int_0^1 (1-h^2r^2)^{-1/2} e^{-u^2/(1+hr)}\,dh$$

$$= \frac{1}{u^2r}\{c'(r)e^{-u^2/(1+r)} - e^{-u^2}\} - \frac{1}{u^2}\int_0^1 \frac{(2-hr)(1+hr)^{1/2}}{(1-hr)^{3/2}} e^{-u^2/(1+hr)}\,dh ,$$

and the second inequality in (i) follows at once, since the last integral in (2.7) is positive. Moreover, $(2-hr)(1+hr)^{1/2}(1-hr)^{-3/2} \leqq c''(\rho)(1-h^2r^2)^{-1/2}$, as is easily checked, and hence

$$(2.8)\qquad \int_0^1 \frac{(2-hr)(1+hr)^{1/2}}{(1-hr)^{3/2}} e^{-u^2/(1+hr)}\,dh \leqq 2\pi c''(\rho)\int_0^1 \psi_{hr}(u)\,dh .$$

Inserting (2.8) into (2.7) we obtain that

$$\int_0^1 \psi_{hr}(u)\,dh\{1 + \frac{c''(\rho)}{u^2}\} \geqq \frac{1}{2\pi u^2 r}\{c'(r)e^{-u^2/(1+r)} - e^{-u^2}\} ,$$

which proves the first inequality in (i).

To prove (ii) we will use that

$$(2.9)\qquad \sqrt{2\pi}(1-\Phi(u)) > \frac{\exp\{-u^2/2\}}{u}\,\frac{u^2}{1+u^2} \geqq \frac{\exp\{-u^2/2\}}{2u}$$

for $u \geqq 1$. Thus, if $r = \rho > 0$, by part (i),

$$\int_0^1 \psi_{hr}(u)\,dh \leqq \frac{c'(r)}{2\pi u^2 r} e^{-u^2/(1+r)}$$

$$\leqq 2^{(2+\rho)/(1+\rho)}(4\pi)^{-\rho/(1+\rho)}|r|^{-1}c'(\rho)\{(1-\Phi(u))/u^{\rho}\}^{2/(1+\rho)} ,$$

and similarly, for $r < 0$, $\rho = 0$,

$$\int_0^1 \psi_{hr}(u)\,dh \leqq \frac{1}{2\pi u^2|r|} e^{-u^2} \leqq \frac{4c'(0)}{|r|}(1-\Phi(u))^2 ,$$

and hence (ii) holds in either case. Finally, it is immediate that, for $u \geqq 1$,

$$\int_0^1 \psi_{hr}(u)\,dh \leqq \frac{1}{2\pi(1-r^2)^{1/2}}\, e^{-u^2/(1+\rho')}$$

$$\leqq 2^{(2+\rho')/(1+\rho')}(4\pi)^{-\rho'/(1+\rho')}(1-r^2)^{-1/2}\{(1-\Phi(u))u\}^{2/(1+\rho')} ,$$

by (2.9), which proves (iii). □

The main lemma now follows easily. In it we will only consider a restricted range of u-values (which may even be empty for small n). The remaining range of u's of interest to us is easier to treat, as shown in the proof of Theorem 3.1 below.

Lemma 2.2 Suppose that for some constant $K > 0$,

$$n(1-\Phi(u)) \leqq K \tag{2.10}$$

and that $1 \leqq u \leqq 2(1+\rho)^{-1/2}(\log n/K)^{1/2}$. Then

$$n\sum_{t=1}^{n} |r_t| \int_0^1 \psi_{hr_t}(u)\,dh \leqq 4R_n , \tag{2.11}$$

for R_n given by (2.3) (i), (ii), and (iii), respectively, for $\rho > 0$, for $\rho = 0$, $\nu = \infty$, and for $\rho = 0$, $\nu < \infty$.

Proof First, by (2.9) and (2.10),

$$\frac{n}{\sqrt{2\pi}}\,\frac{e^{-u^2/2}}{2u} \leqq K ,$$

i.e.

$$\log n/K \leqq \frac{u^2}{2}\{1+\frac{1}{u^2}\log 8\pi u^2\} \leqq 2u^2 , \tag{2.12}$$

for $u \geqq 1$.

Now, suppose that $\rho > 0$. Using Lemma 2.1 (ii) to bound summands with $r_t = \rho$ and Lemma 2.1 (iii) to bound the remaining summands we have that

(2.13)

$$n \sum_{t=1}^{n} |r_t| \int_0^1 \psi_{hr_t}(u)\,dh \leqq n2^{(2+\rho)/(1+\rho)} c(\rho)\{(1-\Phi(u))/u^{\rho}\}^{2/(1+\rho)}$$

$$+ n2^{(2+\rho')/(1+\rho')}(4\pi)^{-\rho'/(1+\rho')} \Sigma' \frac{|r_t|}{(1-r_t^2)^{1/2}} \{(1-\Phi(u))u\}^{2/(1+\rho')} ,$$

where Σ' denotes summation over all $t \in \{1,\ldots,n\}$ such that $r_t \neq \rho$. Since $n(1-\Phi(u)) \leqq K$ and $(1/2 \log n/K) \leqq u \leqq 2(1+\rho)^{-1/2}(\log n/K)^{1/2}$ by assumption and (2.12), we have that

$$\{(1-\Phi(u))/u^{\rho}\}^{2/(1+\rho)} \leqq 2^{\rho/(1+\rho)} \left(\frac{K}{n}\right)^{2/(1+\rho)} (\log n/K)^{-\rho/(1+\rho)}$$

and that

$$\{(1-\Phi(u))u\}^{2/(1+\rho')}$$

$$\leqq 2^{2/(1+\rho')}(1+\rho)^{-1/(1+\rho')}\left(\frac{K}{n}\right)^{2/(1+\rho')} (\log n/K)^{1/(1+\rho')} .$$

Inserting this into (2.13) we obtain, with $\delta = \sup\{|r_1|, |r_2|, \ldots\}$ and $\varepsilon = 2(\rho-\rho')(1+\rho)^{-1}(1+\rho')^{-1}$,

$$n \sum_{t=1}^{n} |r_t| \int_0^1 \psi_{hr_t}(u)\,dh \leqq 4c(\rho)K^{2/(1+\rho)}\left(\frac{1}{n}\right)^{(1-\rho)/(1+\rho)} (\log n/K)^{-\rho/(1+\rho)}$$

$$\times\{\nu + 16K^{\varepsilon}(1-\delta^2)^{-1/2} \sum_{t=1}^{n} |r_t| \left(\frac{1}{n}\right)^{\varepsilon} (\log n/K)^{1+\varepsilon/2}\} ,$$

and comparing with (2.3) (i), this proves (2.11) for the case $\rho > 0$.

Next, suppose $\rho = 0$. Then, using Lemma 2.1 (iii), similar calculations show that

$$n \sum_{t=1}^{n} |r_t| \int_0^1 \psi_{hr_t}(u)\,dh \leqq 4n \sum_{t=1}^{n} \frac{|r_t|}{(1-r_t^2)^{1/2}} \{(1-\Phi(u))u\}^2$$

$$\leqq 16K^2(1-\delta^2)^{-1/2} \sum_{t=1}^{n} |r_t| \frac{1}{n} \log n/K ,$$

proving (2.11) for the case $\rho = 0$, $\nu = \infty$.

Finally, suppose $\rho = 0$, $\nu < \infty$, so that by Lemma 2.1 (ii)

$$n \sum_{t=1}^{n} |r_t| \int_0^1 \psi_{hr_t}(u)\,dh \leqq 4n \sum_{\substack{t=1 \\ r_t \neq 0}}^{n} \frac{|r_t|}{|r_t|} (1-\Phi(u))^2 \leqq 4K^2 \nu \frac{1}{n},$$

which shows that (2.11) holds also in this case. □

Clearly (2.5) and Lemma 2.2 together will provide a bound for $|P(M_n \leqq u) - P(\hat{M}_n \leqq u)|$. However, for the point processes of exceedances, some further estimates are needed. Let, as in the introduction, $u^{(1)} \geqq \ldots \geqq u^{(r)}$ be r levels and define $I_t^{(i)} = I\{\xi_t > u^{(i)}\}$ where I is the indicator function, i.e. $I_t^{(i)}$ is one if $\xi_t > u^{(i)}$ and zero otherwise. Further, let $\mathcal{B}_0$ be the trivial σ-algebra, and for $t \geqq 1$ let $\mathcal{B}_t = \sigma\{I_s^{(i)};\ 1 \leqq i \leqq r,\ 1 \leqq s \leqq t\}$ be the σ-algebra generated by the exceedances up to time t.

Lemma 2.3 (i)

$$\sup_{B \in \mathcal{B}_{t-1}} |P(\{I_t^{(i)} = 0\}B) - P(I_t^{(i)} = 0)P(B)| \leqq$$

$$\leqq 2r \sum_{1 \leqq s < t} |r_{s-t}| \int_0^1 \psi_{hr_{s-t}}(u^{(r)})\,dh .$$

(ii) Suppose that $u = u^{(r)}$ satisfies the requirements of Lemma 2.2. Then

$$\sum_{t=1}^{n} \sum_{i=1}^{r} E|P(I_t^{(i)} = 0 \| \mathcal{B}_{t-1}) - P(I_t^{(i)} = 0)| \leqq 16r^2 R_n ,$$

with R_n given by (2.3).

Proof (i) This follows from an extension of the proof of Lemma 3.2 of Watts, Rootzén & Leadbetter (1980). In fact, let in that proof $\ell_n' = 1$, $\ell_n = 0$, let $B \in \mathcal{B}_{t-1}$ and write

(2.14)

$$B = B_{10}\{I_s^{(1)} = 0\} \cup B_{11}\{I_s^{(1)} = 1\} \cup \ldots \cup B_{r0}\{I_s^{(r)} = 0\} \cup B_{r1}\{I_s^{(r)} = 1\}$$

(instead of $B = B_0\{I_{n,j} = 0\} \cup B_1\{I_{n,j} = 1\}$ in the cited proof), where each of $B_{10}, \ldots, B_{r1}$ is a disjoint union of sets of the form $\cap\{I_\ell^{(j)} = x_{\ell j}\}$ where each $x_{\ell j}$ is zero or one and the intersection is over $j = 1, \ldots, r$ and $\ell = 1, \ldots, s-1, s+1, \ldots, t-1$. Proceeding as in the cited reference, each term $B_{jk}\{I_s^{(j)} = k\}$ leads to a term

(2.15)
$$\int_{-\infty}^{\infty} \cdots \int f_h(x_s = u^{(j)},\ x_t = u^{(i)})\, d\underline{x}'$$

in the estimation of the quantity $F'(h)$ defined there, where $f_h(x_s = u^{(j)}, x_t = u^{(i)})$ is the function of $t-2$ variables which is obtained by putting $x_s = u^{(j)}$, $x_t = u^{(i)}$ in the density function of t stationary normal random variables with zero means, unit variances and covariances hr_t. Now, (2.15) is just the density of two standard normal variables, with correlation hr_{s-t} evaluated at $(u^{(j)}, u^{(i)})$, and may easily be shown to be bounded by $\psi_{hr_{s-t}}(u^{(r)})$. Part (i) then follows at once, since there are 2r terms in (2.14) and since by construction $\int_0^1 F'(h)dh$ is equal to $P(\{I_t^{(i)} = 0\}B) - P(I_t^{(i)} = 0)P(B)$.

(ii) This follows easily if we show that

$$E|P(I_t^{(i)} = 0 \| \mathcal{B}_{t-1}) - P(I_t^{(i)} = 0)|$$
$$\leqq 2 \sup_{B \in \mathcal{B}_{t-1}} |P(\{I_t^{(i)} = 0\}B) - P(I_t^{(i)} = 0)P(B)|$$

since then, by part (i) and Lemma 2.2,

$$\sum_{t=1}^{n} \sum_{i=1}^{r} E|P(I_t^{(i)} = 0 \| \mathcal{B}_{t-1}) - P(I_t^{(i)} = 0)|$$

$$\leqq 2r \sum_{t=1}^{n} 2r \sum_{1 \leqq s<t} |r_{s-t}| \int_0^1 \psi_{hr_{s-t}}(u^{(r)})dh$$

$$\leqq 4r^2 n \sum_{t=1}^{n} |r_t| \int_0^1 \psi_{hr_t}(u^{(r)})dh$$

$$\leqq 16r^2 R_n \ .$$

However, for $B = \{P(I_t^{(i)} = 0 \| \mathcal{B}_{t-1}) > P(I_t^{(i)} = 0)\} \in \mathcal{B}_{t-1}$, by standard calculations

$$E|P(I_t^{(i)} = 0 \| \mathcal{B}_{t-1}) - P(I_t^{(i)} = 0)| = \int_B \{P(I_t^{(i)} = 0 \| \mathcal{B}_{t-1}) - P(I_t^{(i)} = 0)\}dP$$

$$- \int_{B^c} \{P(I_t^{(i)} = 0 \| \mathcal{B}_{t-1}) - P(I_t^{(i)} = 0)\}dP$$

$$= \{P(\{I_t^{(i)} = 0\}B) - P(I_t^{(i)} = 0)P(B)\} - \{P(\{I_t^{(i)} = 0\}B^c) - P(I_t^{(i)} = 0)P(B)\}$$

$$\leqq 2 \sup_{B \in \mathcal{B}_{t-1}} |P(\{I_t^{(i)} = 0\}B) - P(I_t^{(i)} = 0)P(B)| \ ,$$

which completes the proof of part (ii). □

3. The rate of convergence of the maximum

The rate of convergence to zero of $P(M_n \leq u) - P(\hat{M}_n \leq u)$ now follows easily. To obtain efficient bounds we will, as in Lemma 2.2, restrict the domain of variation of u by requiring that $n(1-\Phi(u)) \leq K$, for some fixed $K > 0$, or equivalently that $u \geq u_n$, where u_n is the solution to the equation $n(1-\Phi(u_n)) = K$. Since

$$P(\hat{M}_n \leq u_n) = \Phi(u_n)^n = \{1 - \tfrac{1}{n}(n(1-\Phi(u_n))\}^n ,$$

this clearly implies that

$$(3.1) \qquad P(\hat{M}_n \leq u_n) \to e^{-K} , \quad n \to \infty ,$$

and conversely, if (3.1) holds, then $n(1-\Phi(u_n)) \to K$, as is easily seen. Moreover, if $P(M_n \leq u_n) - P(\hat{M}_n \leq u_n) \to 0$, then of course the same equivalence holds for $\hat{M}_n$ replaced by M_n.

Thus, since the bounds for the rate of convergence will be proved for $u \geq u_n$, they will apply to the upper part of the range of variation of $P(M_n \leq u)$, and by taking K large an arbitrarily large part of this range is covered, but at the cost of a poorer bound.

Theorem 3.1 Let $\{\xi_t\}$ be stationary normal, with zero means, unit variances and covariances $r_t = E\xi_s\xi_{s+t}$. Let $\rho = \sup\{0, r_1, r_2, \ldots\}$ and let ν be the number of r_t's, $t \geq 1$, with $r_t = \rho$, in case $\rho > 0$, and let ν be the number of non-zero r_t's, for $t \geq 1$, otherwise. Further, let $\delta = \sup\{|r_1|, |r_2|, \ldots\}$ and let $\varepsilon > 0$, c, c', and R_n be as defined in (2.1) - (2.3). Suppose that $u \geq 1$, and that

$$(3.2) \qquad n(1-\Phi(u)) \leq K ,$$

for some constant K, with $n/K \geq e$. Then

(3.3) $$|P(M_n \leqq u) - P(\hat{M}_n \leqq u)| \leqq 4R_n ,$$

or, more explicitly, writing $\Delta_n = |P(M_n \leqq u) - P(\hat{M}_n \leqq u)|$, if $\rho > 0$ then

$$\Delta_n \leqq 4c''(\rho)K^{2/(1+\rho)} (\tfrac{1}{n})^{(1-\rho)/(1+\rho)} (\log n/K)^{-\rho/(1+\rho)} \{\nu + r_n\},$$

with r_n given by (2.3,i), if $\rho = 0$ then

$$\Delta_n \leqq 16K^2(1-\delta^2)^{-1/2} \tfrac{1}{n} \log n/K \sum_{t=0}^{n} |r_t|$$

and if in addition $\nu < \infty$ then

$$\Delta_n \leqq 4K^2 \tfrac{1}{n} \nu .$$

<u>Proof</u> By (2.5) and (2.6)

$$|P(M_n \leqq u) - P(\hat{M}_n \leqq u)| \leqq \sum_{1 \leqq s < t \leqq n} |r_{s-t}| \int_0^1 \psi_{hr_{s-t}}(u)\,dh$$

$$\leqq n \sum_{t=1}^{n} |r_t| \int_0^1 \psi_{hr_t}(u)\,dh ,$$

and it follows from Lemma 2.2 that (3.3) holds for u satisfying (3.2) and $1 \leqq u \leqq 2(1+\rho)^{-1/2} (\log n/K)^{1/2}$.

To complete the proof we will show that (3.3), rather trivially, is satisfied also for $u > 2(1+\rho)^{-1/2} (\log n/K)^{1/2}$. In fact

(3.4) $$|P(M_n \leqq u) - P(\hat{M}_n \leqq u)| = |P(M_n > u) - P(\hat{M}_n > u)|$$

$$\leqq P(M_n > u) + P(\hat{M}_n > u)$$

$$\leqq 2n(1 - \Phi(u)) ,$$

by Boole's inequality. Since $1 - \Phi(u) \leqq (2\pi)^{-1/2} e^{u^2/2}/u$, we have for $u \geqq 2(1+\rho)^{-1/2}(\log n/K) \geqq 1$ that

$$1 - \Phi(u) \leqq (2\pi)^{-1/2} \exp\{-\tfrac{1}{2}(2(1+\rho)^{-1/2}(\log n/K)^{1/2})^2\} (\log n/K)^{-1/2}$$
$$\leqq (2\pi)^{-1/2} (\tfrac{K}{n})^{2/(1+\rho)} (\log n/K)^{-1/2} \quad ,$$

and hence, by (3.4),

$$|P(M_n \leqq u) - P(\hat{M}_n \leqq u)| \leqq 2(2\pi)^{-1/2} n (\tfrac{K}{n})^{2/(1+\rho)} (\log n/K)^{-1/2}$$
$$\leqq 4R_n \quad ,$$

by straightforward calculation. □

As an easy corollary to the theorem we shall prove that an analogue of Hall's result (1.5) holds also for dependent sequences, under appropriate conditions.

<u>Corollary 3.2</u> Suppose that $\{\xi_t\}$ is stationary normal, with zero means, unit variances, and covariances $\{r_t\}$ such that

(3.5) $$(\log n)^2 (\log\log n)^2 (\tfrac{1}{n})^{(1-\rho')/(1+\rho')} \sum_{t=1}^{n} |r_t| \to 0 \quad ,$$

as $n \to \infty$, where ρ' is defined on P. 6

Then for a_n', b_n', C_1, and C_2 satisfying (1.5)

$$0 < C_1 \leqq \liminf_{n\to\infty}\{\sup_x \log n\, |P(a_n'(M_n - b_n') \leqq x) - \exp\{-e^{-x}\}|\}$$
$$\leqq \limsup_{n\to\infty}\{\sup_x \log n\, |P(a_n'(M_n - b_n') \leqq x) - \exp\{-e^{-x}\}|\}$$
$$\leqq C_2$$

and the order $1/\log n$ of convergence cannot be improved by choosing other norming constants than a_n', b_n'. In particular, for $a_n = (2\log n)^{1/2}$, $b_n = a_n - \{\log\log n - \log 4\pi\}/(2a_n)$,

$$P(a_n(M_n - b_n) \leq x) - \exp\{-e^{-x}\} \sim \frac{1}{16} e^{-x} \exp\{-e^{-x}\} \frac{(\log \log n)^2}{\log n} .$$

<u>Proof</u> By (1.5) and (1.4) it is sufficient to prove that

$$\text{(3.6)} \qquad \sup_{-\infty<u<\infty} |P(M_n \leq u) - P(\hat{M}_n \leq u)| = o(1/\log n) .$$

We first note that (3.5) implies that $\delta = \sup\{|r_1|, |r_2|, \ldots\} < 1$ (since otherwise r_t would be periodic, which contradicts (3.5)). Now, it is straightforward to check that if the constant K in the bound R_n is chosen as $K = K_n = 2 \log\log n$, and if (3.5) holds, then $R_n = o(1/\log n)$, so that

$$\text{(3.7)} \qquad \sup_{u \geq u_n} |P(M_n \leq u) - P(\hat{M}_n \leq u)| = o(1/\log n) ,$$

for u_n given by $n(1 - \Phi(u_n)) = K_n = 2 \log\log n$.

Furthermore, for $u \leq u_n$,

$$\begin{aligned} |P(M_n \leq u) - P(\hat{M}_n \leq u)| &\leq 2P(\hat{M}_n \leq u_n) + |P(M_n \leq u_n) - P(\hat{M}_n \leq u_n)| \\ &= 2\Phi(u_n)^n + o(1/\log n) , \end{aligned}$$

and since

$$\begin{aligned} \Phi(u_n) &= (1 - (1 - \Phi(u_n)))^n \\ &\leq e^{-n(1-\Phi(u_n))} \\ &= (\log n)^{-2} , \end{aligned}$$

it follows that

$$\sup_{u \leq u_n} |P(M_n \leq u) - P(\hat{M}_n \leq u)| = o(1/\log n) ,$$

which together with (3.7) proves (3.6). □

From Theorem 3.1 follows that, supposing $\Sigma_{t=1}^n |r_t|$ does not grow too rapidly, if $\rho > 0$ then the rate of convergence is at least of the order

$$\left(\frac{1}{n}\right)^{(1-\rho)/(1+\rho)} (\log n)^{-\rho/(1+\rho)} ,$$

and that if $\rho = 0$, $\nu < \infty$, then the rate is at least of the order

$$\frac{1}{n} .$$

We will now find a precise asymptotic expression for $P(M_n \leqq u) - P(\hat{M}_n \leqq u)$ in the case when $r_t = 0$ if $|t| > m$, for some constant $m < \infty$, i.e. when the sequence is m-dependent. This will show that, at least for such sequences, these rates are of the right order.

If $\rho = 0$, $\nu = \infty$, and $\Sigma_{t=0}^{\infty} |r_t| < \infty$, the bound given by Theorem 3.1 is of the order

$$\frac{1}{n} \log n .$$

It seems unlikely that this is the correct order, but the loss does not seem important, since clearly the rate of convergence cannot be better than 1/n, in general.

<u>Theorem 3.3</u> Suppose $\{\xi_t\}$ is stationary normal, with zero means, unit variances and covariances $\{r_t\}$ such that $r_t = 0$ for $|t| > m$, for some constant $m < \infty$, and that

$$n(1 - \Phi(u_n)) \to K > 0 , \quad \text{as } n \to \infty .$$

Then, if $\rho > 0$

(3.8)

$$P(M_n \leqq u_n) - P(\hat{M}_n \leqq u_n) \sim e^{-K} c(\rho) K^{2/(1+\rho)} \left(\frac{1}{n}\right)^{(1+\rho)/(1-\rho)} (\log n)^{-\rho/(1+\rho)} \nu$$

$$\sim e^{-K} R_n ,$$

and if $\rho = 0$, then

$$(3.9) \qquad P(M_n \leqq u_n) - P(\hat{M}_n \leqq u_n) \sim - e^{-K} K^2 \frac{1}{n} \nu \sim e^{-K} R_n .$$

Proof The essential part of the proof consists of a closer evaluation of the quantity

$$\int_{-\underline{\infty}}^{\underline{u}'} \!\!\cdots\! \int f_h(x' | x_s = x_t = u) d\underline{x}' = P , \quad \text{say} ,$$

which was estimated by one in (2.6), and in the proof of Theorem 3.1. For this it is convenient to introduce a further stationary normal sequence, $\{\tilde{\xi}_t\}$ say, with means zero, variances one and covariance function hr_t. Let

$$\tilde{M}_n = \max_{1 \leqq t \leqq n} \tilde{\xi}_t , \quad \tilde{M}_I = \max_{t \in I} \tilde{\xi}_t$$

so that, for $u = u_n$,

$$P = P(\tilde{M}_n \leqq u_n || \tilde{\xi}_s = \tilde{\xi}_t = u_n) ,$$

and let $I = \{k \in [1,n];\ |k-s| \leqq m \text{ or } |k-t| \leqq m\}$ and $J = \{1,\dots,n\} \cap I^c$. By Boole's inequality

$$(3.10) \qquad P(\tilde{M}_J \leqq u_n || \tilde{\xi}_s = \tilde{\xi}_t = u_n) - \sum_{k \in I} P(\tilde{\xi}_k > u_n || \tilde{\xi}_s = \tilde{\xi}_t = u_n) \leqq P$$
$$\leqq P(\tilde{M}_J \leqq u_n || \tilde{\xi}_s = \tilde{\xi}_t = u_n) .$$

Since $\{\tilde{\xi}_t\}$ is m-dependent, $P(\tilde{M}_J \leqq u_n || \tilde{\xi}_s = \tilde{\xi}_t = u_n) = P(\tilde{M}_J \leqq u_n)$ and thus by a similar calculation

$$P(\tilde{M}_n \leqq u_n) \leqq P(\tilde{M}_J \leqq u_n || \tilde{\xi}_s = \tilde{\xi}_t = u_n)$$
$$\leqq P(\tilde{M}_n \leqq u_n) + \sum_{k \in I} P(\tilde{\xi}_k > u_n) .$$

Now, since $\Sigma_{k \in I} P(\tilde{\xi}_k > u_n) \leqq 4m(1 - \Phi(u_n)) \to 0$, and since, by Theorem 3.1, $P(\tilde{M}_n \leqq u_n) \sim P(\hat{M}_n \leqq u_n) = \Phi(u_n)^n \to e^{-K}$, $n \to \infty$, it follows that

$$(3.11)\qquad P(\tilde{M}_J \leqq u_n \,\|\, \tilde{\xi}_s = \tilde{\xi}_t = u_n) \to e^{-K}\,, \quad \text{as } n \to \infty\,,$$

uniformly in s,t, and h.

Next, given that $\tilde{\xi}_s = \tilde{\xi}_t = u_n$, $\tilde{\xi}_k$ is normal with variance not exceeding one and mean $u_n(hr_{k-s} + hr_{k-t})/(1 + hr_{s-t})$. We will temporarily assume that

$$(3.12)\qquad 0 < \varepsilon \leqq 1 - \max_{k \neq s,t} h(r_{k-s} + r_{k-t})/(1 + hr_{s-t})\,,$$

for some constant ε which does not depend on k,s,t, or h. Then

$$\sum_{k \in I} P(\tilde{\xi}_k > u_n \,\|\, \tilde{\xi}_s = \tilde{\xi}_t = u_n) \leqq 4m(1 - \Phi(\varepsilon u_n)) \to 0, \quad \text{as } n \to \infty\,,$$

and by (3.10) and (3.11),

$$(3.13)\qquad P = \int_{-\underline{\infty}}^{\underline{u}_n} \cdots \int f_h(x' | x_s = x_t = u_n)\, dx' \to e^{-K}\,,$$

as $n \to \infty$, uniformly in s,t, and h.

Now, if $\rho = 0$ then (3.12) is satisfied, with $\varepsilon = 1$, and hence, by (2.5) and (3.13),

(3.14)

$$P(M_n \leqq u_n) - P(\hat{M}_n \leqq u_n) = \sum_{1 \leqq s < t \leqq n} r_{s-t} \int_0^1 \psi_{hr_{s-t}}(u) \int_{-\underline{\infty}}^{\underline{u}'_n} \cdots \int f_h(x' | x_s = x_t = u_n)\, dx'\, dh$$

$$\sim \sum_{1 \leqq s < t \leqq n} r_{s-t} \int_0^1 \psi_{hr_{s-t}}(u)\, e^{-K} dh$$

$$\sim e^{-K}\, n \sum_{t=1}^{n} r_t \int_0^1 \psi_{hr_t}(u)\, dh$$

$$\sim e^{-K}\, n \frac{1}{2\pi u_n^2} e^{-u_n^2} (-\nu)$$

$$\sim - e^{-K} K^2 \frac{1}{n} \nu\,,$$

where we have used Lemma 2.1 (i) in the fourth step and that $(2\pi)^{-1/2} e^{-u_n^2/2}/u_n \sim 1-\Phi(u_n) \sim K/n$ in the last step.

This proves (3.9), and we next suppose that $\rho > 0$ and let Σ'' denote the sum over s,t such that $1 \leq s < t \leq n$ and $r_{s-t} = \rho$. In the same way as in (3.14) we then have that

$$\Sigma'' r_{s-t} \int_0^1 \psi_{hr_{s-t}}(u_n) \int_{-\underline{\infty}}^{\underline{u}_n} \cdots \int f_h(x' \mid x_s = x_t = u_n)\, d\underline{x}'\, dh$$

$$\sim e^{-K} c(\rho) \left(\frac{1}{n}\right)^{(1-\rho)/(1+\rho)} (\log n)^{-\rho/(1+\rho)} \nu ,$$

since $u_n \sim \sqrt{2 \log n}$, and since for s,t such that $r_{s-t}=\rho$ the condition (3.12) is clearly satisfied for some suitable $\varepsilon > 0$. Since the sum of the remaining terms is $\sigma(n^{-(1-\rho)/(1+\rho)} (\log n)^{-\rho/(1+\rho)})$, as was seen in the proof of Lemma 2.2, this shows that

$$\Sigma r_{s-t} \int_0^1 \psi_{hr_{s-t}}(u_n) \int_{-\underline{\infty}}^{\underline{u}'_n} \cdots \int f_h(x' \mid x_s = x_t = u_n)\, dx'\, dh$$

$$\sim e^{-K} c(\rho) \left(\frac{1}{n}\right)^{(1-\rho)/(1+\rho)} (\log n)^{-\rho/(1+\rho)} \nu ,$$

and hence by (2.5) that (3.8) holds. □

Comparing the asymptotic expressions for $P(M_n \leq u_n) - P(\hat{M}_n \leq u_n)$ with the bounds of Theorem 3.1 we see that the bounds asymptotically are too large by a factor $4e^K$. Here the factor 4 is due to inaccuracies in the estimates (2.9) and (2.12) and could easily be reduced by restricting the range of u further. The factor e^K is due to the estimate (2.6), as was seen in the proof of Theorem 3.3, and could conceivably be reduced along similar lines as in that proof, but perhaps at the expense of a considerable increase in the complexity of proofs.

4. A representation for the point processes of exceedances

Let $N_n = (N_n^{(1)},\ldots,N_n^{(r)})$ be the vector of time-normalized exceedances of the levels $u^{(1)} \geqq \ldots \geqq u^{(r)}$, i.e. $N_n^{(i)}(B) = \#\{t; \xi_t > u^{(i)}, t/n \in B\}$, for any Borel set $B \subseteq [0,1]$, and let $\hat{N}_n = (\hat{N}_n^{(1)},\ldots,\hat{N}_n^{(r)})$ be defined similarly, with $\{\xi_t\}$ replaced by the associated independent sequence $\{\hat{\xi}_t\}$. It is known, see Leadbetter, Lindgren & Rootzén (1979), that under weak conditions (the same as those commonly used to establish (1.3)) N_n converges in distribution to a certain successively more severely thinned Poisson process (which will be described below). To formulate results about the rate of convergence of the distribution of N_n, and, more generally, to find useful ways to measure the distance between the distributions of two point processes seems to be an interesting and non-trivial question, but here we will partly circumvent this issue by using a "representation" approach. More precisely, we will construct two processes which have the same distribution as N_n and $\hat{N}_n$, respectively, and whose realisations are identical with high probability. Following common usage, we will refer to these processes as *versions* of N_n and $\hat{N}_n$ and, since it does not lead to any confusion, we will use the same letter to denote processes which are versions of one another.

The limiting process $N = (N^{(1)},\ldots,N^{(r)})$ can be described in the following way. Let $0 < \tau^{(1)} \leqq \ldots \leqq \tau^{(r)}$ be given parameters, and let $N^{(r)}$ be a Poisson process in $[0,1]$, with parameter $\tau^{(r)}$ and points $\{\sigma_k\}$. Let $\{\beta_k\}$ be independent random variables, independent also of $N^{(r)}$ and taking values in $1,\ldots,r$ with probabilities

$$
\begin{aligned}
P(\beta_k = i) &= (\tau^{(r-i+1)} - \tau^{(r-i)})/\tau^{(r)}, \qquad & i = 1,\ldots,r-1 \\
&= \tau^{(1)}/\tau^{(r)}, & i = r\ .
\end{aligned}
$$

For each k such that $\beta_k > 1$ let $N^{(r-1)},\ldots, N^{(r-\beta_k+1)}$ have points at σ_k, to complete the definition of $N=(N^{(1)},\ldots,N^{(k)})$. Thus, in particular, each $N^{(i)}$ is a Poisson process with intensity $\tau^{(i)}$, but the dependence between the component processes does not have a Poisson character.

Since, for each i, $N_n^{(i)}$ is concentrated on the set $\{1/n,\ldots,n/n\}$ while the probability is zero that $N^{(i)}$ has a point in $\{1/n,\ldots,n/n\}$, it is not possible to construct versions of N_n and N with realizations which are identical with a probability tending to one. However, such a construction is possible if N is first discretized as follows; for each i, $1 \leqq i \leqq r$, let $\tilde{N}_n^{(i)}$ be concentrated on $\{1/n,\ldots,n/n\}$ with $\tilde{N}_n^{(i)}(\{t/n\}) = N^{(i)}(((t-1)/n,t/n])$, $t=1,\ldots,n$. Thus $\tilde{N}_n$ is obtained from N by "discretizing" by placing all the points of N in the intervals $((t-1)/n,t/n]$ at the endpoints t/n of the intervals.

<u>Theorem 4.1</u> Let $\{\xi_t\}$ be stationary normal with zero means, unit variances and covariances $r_t = E\xi_s\xi_{s+t}$, let R_n be given by (2.3) and let $u^{(1)} \geqq \ldots \geqq u^{(r)} \geqq 1$.

(i) If

(4.1) $$n(1-\Phi(u^{(r)})) \leqq K ,$$

for some constant K with $n/K \geqq e$, then there exist versions N_n and $\hat{N}_n$ of the vectors of time-normalized point-processes of exceedances of $u^{(1)} \geqq \ldots \geqq u^{(r)}$ by $\{\xi_t\}$ and $\{\hat{\xi}_t\}$ respectively,

(4.2) $$P(N_n \neq \hat{N}_n) \leqq 16r^2 R_n \quad .$$

(ii) Let $N=(N^{(1)},\ldots,N^{(r)})$, be the thinned Poisson process described above, with parameters $\tau^{(1)} \leqq \ldots \leqq \tau^{(r)}$ and let $\tilde{N}_n$ be obtained from N by "discretizing". Write $\tau_n^{(i)} = n(1-\Phi(u^{(i)}))$,

$i = 1, \ldots, r$, and suppose that (4.1) is satisfied. Then there exist versions of $N_n, \tilde{N}_n$ such that

$$(4.3) \qquad P(N_n \neq \tilde{N}_n) \leq 16r^2 R_n + \sum_{i=1}^{r} |\tau_n^{(i)} - \tau^{(i)}| + \frac{1}{n} \sum_{i=1}^{r} (\tau^{(i)})^2 .$$

Proof (i) We will use the main idea of Serfling (1976) in the proof. Let $(\Omega, \mathcal{F}, P)$ be a probability space which supports independent variables $\eta_1, \ldots, \eta_n$ which are uniformly distributed on $[0,1]$, write $p^{(i)} = P(\xi_1 \leq u^{(i)}) = \Phi(u^{(i)})$, $i = 1, \ldots, r$, and recall the notations $I_t^{(i)} = I\{\xi_t > u^{(i)}\}$, $\hat{I}_t^{(i)} = I\{\hat{\xi}_t > u^{(i)}\}$, and $\mathcal{B}_t = \sigma\{I_k^{(i)};\ 1 \leq i \leq r,\ 1 \leq k \leq t\}$. Thus $P_t^{(i)} = P(I_t = 0 \,\|\, \mathcal{B}_{t-1})$ is a function, $P_t^{(i)} = P_t^{(i)}(I_1^{(1)}, \ldots, I_{t-1}^{(r)})$, of $I_1^{(1)}, \ldots, I_{t-1}^{(r)}$. We will first use Serfling's construction to define suitable versions of the processes $\{I_t^{(i)}\}$ and $\{\hat{I}_t^{(i)}\}$ on $(\Omega, \mathcal{F}, P)$. A version of the latter process is obtained by setting

$$\hat{I}_t^{(i)} = I(\eta_t > p^{(i)}), \quad i = 1, \ldots, r,\ t = 1, \ldots, n ,$$

and further it is readily seen that by defining

$$I_1^{(i)} = I(\eta_t > P_1^{(i)}) = I(\eta_t > p^{(i)}),\ i = 1, \ldots, r,$$

and then recursively, for $t = 2, \ldots, n$, $i = 1, \ldots, r$,

$$I_t^{(i)} = I(\eta_t > P_t^{(i)}(I_1^{(1)}, \ldots, I_{t-1}^{(r)})) ,$$

one obtains a version of the former process.

Hence, defining $N_n^{(i)}$, $\hat{N}_n^{(i)}$ by $N_n^{(i)}(\{t/n\}) = I_t^{(i)}$, $\hat{N}_n(\{t/n\}) = \hat{I}_t^{(i)}$, $i = 1, \ldots, r$, $t = 1, \ldots, n$, and $N_n^{(i)}(\{1/n, \ldots, n/n\}^c) = \hat{N}_n^{(i)}(\{1/n, \ldots, n/n\}^c) = 0$ it follows that $N_n = (N_n^{(1)}, \ldots, N_n^{(r)})$ and $\hat{N}_n = (\hat{N}_n^{(1)}, \ldots, \hat{N}_n^{(r)})$ have the same distributions as the point processes of exceedances of the levels $u^{(1)}, \ldots, u^{(r)}$ by $\{\xi_t\}$ and $\{\hat{\xi}_t\}$ respectively. Furthermore,

$$P(I_t^{(i)} \neq \hat{I}_t^{(i)}) \leqq E\{E\{|I_t^{(i)} - \hat{I}_t^{(i)}| \,\|\, \eta_1,\ldots,\eta_{t-1}\}\}$$
$$= E|P_t^{(i)} - p^{(i)}|$$
$$= E|P(I_t^{(i)} = 0 \,\|\, \mathcal{B}_{t-1}) - P(I_t^{(i)} = 0)|$$

and since

$$P(N_n \neq \hat{N}_n) \leqq \sum_{t=1}^{n} \sum_{i=1}^{r} P(I_t^{(i)} \neq \hat{I}_t^{(i)})$$

it follows by Lemma 2.3 (ii) that (4.2) holds for $u^{(r)} \leqq 2(1+\rho)^{-1/2}(\log n/K)^{1/2}$. Further, for $u^{(r)} > 2(1+\rho)^{-1/2}(\log n/K)^{1/2}$

$$P(N_n \neq \hat{N}_n) \leqq P(M_n > u^{(r)}) + P(\hat{M}_n > u^{(r)})$$
$$\leqq 4R_n \quad ,$$

as was shown in the proof of Theorem 3.1, and hence (4.2) holds also in this case.

The proof of part (ii) is quite similar. For $i = 1,\ldots,r$ let $\pi_m^{(i)} = \Sigma_{k-m}^{\infty} \exp\{-\tau^{(i)}/n\} \cdot (\tau^{(i)}/n)^k/k!$ be the probability that a Poisson variable with mean $\tau^{(i)}/n$ is larger than m, and define $\tilde{I}_1^{(1)},\ldots,\tilde{I}_n^{(r)}$ by requiring that $\tilde{I}_t^{(i)} = m$ on the set $\pi_{m+1}^{(i)} < \eta_t \leqq \pi_m^{(i)}$. It is then immediate that $\tilde{N}_n = (\tilde{N}_n^{(1)},\ldots,\tilde{N}_n^{(r)})$ has the required distribution if $\tilde{N}_n^{(i)}$ is defined by $\tilde{N}_n^{(i)}(\{t/n\}) = \tilde{I}_t^{(i)}$, $t = 1,\ldots,n$, and $\tilde{N}_n^{(i)}(\{1/n,\ldots,n/n\}^c) = 0$. Furthermore,

$$P(\hat{N}_n^{(i)} \neq \tilde{N}_n^{(i)}) \leqq \sum_{t=1}^{n} P(\hat{I}_t^{(i)} \neq \tilde{I}_t^{(i)})$$
$$\leqq n(|\pi_1^{(i)} - p^{(i)}| + \pi_2^{(i)})$$
$$\leqq n(|\tau^{(i)}/n - \tau_n^{(i)}/n| + (\tau^{(i)}/n)^2)$$
$$= |\tau^{(i)} - \tau_n^{(i)}| + \frac{1}{n}(\tau^{(i)})^2 \quad ,$$

where we have used that $|\pi_1^{(i)} - p^{(i)}| = |1 - e^{-\tau(i)/n} - \tau_n^{(i)}/n| \leqq |\tau^{(i)}/n - \tau_n^{(i)}/n| + (\tau^{(i)}/n)^2/2$ and that $\pi_2^{(i)} \leqq (\tau^{(i)}/n)^2/2$ in the third step. The inequality (4.3) now follows at once from part (i), since $P(N_n \neq \tilde{N}_n) \leqq P(N_n \neq \hat{N}_n) + \Sigma_{i=1}^r P(\hat{N}_n^{(i)} \neq \tilde{N}_n^{(i)})$. □

The variation distance, d, between the distributions of two r-dimensional integervalued random variables $X = \{X^{(k)};\ 1 \leqq k \leqq r\}$ and $Y = \{Y^{(k)};\ 1 \leqq k \leqq r\}$ is defined as

$$d(X,Y) = \frac{1}{2} \sum_{z \in Z^r} |P(X=z) - P(Y=z)| ,$$

where Z^r is the r-dimensional integer lattice. It is immediate that d is a metric on the set of distributions on Z^r, and that it is a metric for convergence in distribution. The interest and usefulness of the metric can be seen from the easily obtained relations

$$\text{(4.4)} \qquad d(X,Y) = \frac{1}{2} \sup_{|h| \leqq 1} |Eh(X) - Eh(Y)|$$

$$= \sup_A |P(X \in A) - P(Y \in A)|$$

$$\geqq \sup_{z \in Z^r} |P(X \leqq z) - P(Y \leqq z)|$$

The distance d only depends on the marginal distributions of X and Y, but if X and Y have a joint distribution it follows from the second inequality in (4.4) that

$$\text{(4.5)} \qquad d(X,Y) \leqq P(X \neq Y) .$$

This at once gives the first part of the following corollary to theorem 4.1.

Corollary 4.2 Suppose that the hypothesis of Theorem 4.1 is satisfied, let A_{ij}, $j = 1,\dots,k_i$, $i = 1,\dots,r$ be Borel subsets of

$[0,1]$, and write $\nu_n = \{N_n^{(i)}(A_{ij});\ j=1,\ldots,k_i,\ i=1,\ldots,r\}$ and let $\hat{\nu}_n, \tilde{\nu}_n, \nu$ be defined similarly, but with N replaced by $\hat{N}_n$, $\tilde{N}_n$, and N, respectively.

(i) Then

$$d(\nu_n,\hat{\nu}_n) \leqq 16\, r^2 R_n,$$

$$d(\nu_n,\tilde{\nu}_n) \leqq 16\, r^2 R_n + \sum_{i=1}^{r} |\tau_n^{(i)} - \tau^{(i)}| + \frac{1}{n}\sum_{i=1}^{r} (\tau^{(i)})^2 \quad ,$$

and if furthermore the A_{ij}'s are intervals, then

$$d(\nu_n,\nu) \leqq 16\, r^2 R_n + \sum_{i=1}^{n} |\tau_n^{(i)} - \tau^{(i)}| + \frac{1}{n}\sum_{i=1}^{r} \{2k_i\, \tau^{(i)} + (\tau^{(i)})^2\}.$$

(ii) Suppose $r=1$, and write $N_n = N_n^{(1)}$ for the time-normalized point process of exceedances of the level $u = u^{(1)}$, let N be a Poisson process with parameter τ, and suppose that $\tau_n = n(1-\Phi(u)) \leqq K$, with $n/K \geqq e$. Further let $h : [0,1] \to \mathbb{R}$ be bounded by one, $|h| \leqq 1$, and have modulus of continuity $\delta(\varepsilon) = \sup\{|h(t)-h(t')|;\ t,t' \in [0,1], |t-t'| \leqq \varepsilon\}$. Then

$$P(|\textstyle\int h dN_n - \int h dN| > \varepsilon) \leqq 16 r^2 R_n + |\tau_n - \tau| + \tau^2/n + \displaystyle\sum_{k>\varepsilon/\delta(1/n)} e^{-\tau}\tau^k/k!\ .$$

<u>Proof</u> (i) The first two bounds follow at once from Theorem 4.1 and (4.5). Furthermore,

$$(4.6) \qquad P(\tilde{\nu}_n \neq \nu) \leqq \sum_{i=1}^{r}\sum_{j=1}^{k_i} P(\tilde{N}_n(A_{ij}) \neq N(A_{ij}))$$
$$\leqq \sum_{i=1}^{r}\sum_{j=1}^{k_i} 2(1-e^{-\tau^{(i)}/n})$$
$$\leqq \sum_{i=1}^{r} 2k_i\, \tau^{(i)} \ ,$$

since clearly $\tilde{N}_n^{(i)}(A_{ij}) = N(A_{ij})$ if $N^{(i)}$ does not have any points in the two intervals $((t-1)/n, t/n]$ which contain the endpoints of the interval A_{ij}. The last inequality in part (i) is an immediate consequence of (4.6), the second inequality in part (i), and the triangle inequality for the metric d.

(ii) By (4.3) there exist versions of N_n, $\tilde{N}_n$ such that

$$P(N_n \neq \tilde{N}_n) \leqq 16R_n + |\tau_n - \tau| + \tau^2/n \quad . \tag{4.7}$$

Furthermore, recalling the definition of $\tilde{N}_n$ from the Poisson process N with intensity τ, we have

$$\begin{aligned} \int h d\tilde{N}_n - \int h dN &\leqq \sum_{t=1}^{n} \int_{s=(t-1)/n}^{t/n} |(h(t/n) - h(s))| dN(s) \\ &\leqq \sum_{t=1}^{n} \delta(1/n) \int_{s=(t-1)/n}^{t/n} dN(s) \\ &= \delta(1/n) N([0,1]) \quad , \end{aligned}$$

and hence

$$P(|\int h d\tilde{N}_n - \int h dN| > \varepsilon) \leqq \sum_{k>\varepsilon/\delta(1/n)} e^{-\tau}\tau^k/k! \quad . \tag{4.8}$$

Part (ii) now follows at once from (4.7) and (4.8). □

As a last corollary we will give an approximation of the joint distribution of the $M_n^{(k)}$'s, the k-th largest of $\xi_1,\dots,\xi_n$. The proof is immediate from Corollary 4.2 (i) and the obvious relation

$$\{M_n^{(1)} \leqq u^{(1)}, \dots, M_n^{(r)} \leqq u^{(r)}\} = \{N_n^{(1)}([0,1]) \leqq 0, \dots, N_n^{(r)}([0,1]) \leqq r-1\}$$

and its counterpart for $\hat{M}_n^{(k)}$, the k-th largest amongst $\hat{\xi}_1,\dots,\hat{\xi}_n$.

<u>Corollary 4.3</u> Suppose that the hypothesis of Theorem 4.1 is satisfied. Then

$$|P(M_n^{(1)} \leqq u^{(1)}, \ldots, M_n^{(r)} \leqq u_r) - P(\hat{M}_n^{(1)} \leqq u^{(1)}, \ldots, \hat{M}_n^{(r)} \leqq u_r)| \leqq 16r^2 R_n,$$

and

$$|P(M_n^{(1)} \leqq u^{(1)}, \ldots, M_n^{(r)} \leqq u_r) - P(N^{(1)}([0,1]) \leqq 0, \ldots, N^{(r)}([0,1]) \leqq r-1)|$$

$$\leqq 16r^2 R_n + \sum_{i=1}^{r} |\tau_n^{(i)} - \tau^{(i)}| + \frac{1}{n} \sum_{i=1}^{r} (\tau^{(i)})^2 .$$

In particular, for $\tau_n = n(1-\Phi(u)) \leqq K$,

$$|P(M_n^{(k)} \leqq u) - \sum_{i=0}^{k-1} e^{-\tau_n} \tau_n^i / i!| \leqq 16R_n + \frac{1}{n} \tau_n^2 .$$

5. A comment on improved approximations

As noted above, the approximation errors in Theorems 3.1 and 4.1 are roughly of the order $(1/n)^{(1-\rho)/(1+\rho)}$, with $\rho=\max\{0,r_1,r_2,\ldots\}$, provided r_t decreases sufficiently fast as $t\to\infty$. For ρ small, this seems quite satisfactory, but for ρ close to one it would be useful to have more accurate approximations. Considering e.g. $P(M_n \leqq u)$, and $P(\hat{M}_n \leqq u)$ as a first order term in the approximation of it, one possibility would be to find a second order term, similar to the right hand side of (3.8), and then to find a bound for the error in the second order approximation of $P(M_n \leqq u)$. This, although perhaps feasible, seems likely to incur considerable extra complications to the already somewhat involved calculations of this paper.

Another possibility would be to approximate M_n by the maximum, say $\overline{M}_n$, of some other dependent stationary sequence $\{\overline{\xi}_t\}$ which in some way is easier to handle than the original sequence $\{\xi_t\}$. If $\{\overline{\xi}_t\}$ has zero means, unit variances, and covariances $\overline{r}_t = E\,\overline{\xi}_s\overline{\xi}_{s+t}$, an analogue of (2.4) is valid, namely

$$P(M_n \leqq u) - P(\overline{M}_n \leqq u) = \sum_{1\leqq s<t\leqq n} \{r_{s-t} - \overline{r}_{s-t}\} \int_0^1 \int_{-\infty}^{u} \cdots \int f_h(x_s = x_t = u)\, d\underline{x}'dh \;,$$

where $f_h(x_s = x_t = u)$ now is obtained by putting $x_s = x_t = u$ in the density function of n standard normal variables with covariances $hr_t + (1-h)\overline{r}_t$, c.f. e.g. Leadbetter, Lindgren, & Rootzén (1979), p. 47. Let ρ,ρ' be as on p. 6 and define $\overline{\rho},\overline{\rho}'$ similarly from $\{\overline{r}_t\}$. Calculations parallel to those in Sections 2 and 3 then show that, under suitable conditions, the rate of convergence to zero of $P(M_n \leqq u) - P(\overline{M}_n \leqq u)$ is similar to that of $P(M_n \leqq u) - P(\hat{M}_n \leqq u)$, but with ρ replaced by $\max(\rho,\overline{\rho})$ if $\rho \neq \overline{\rho}$, and by max

$(\rho', \bar{\rho}')$ if $\rho = \bar{\rho}$, $\rho' \neq \bar{\rho}'$. Thus, the order of the approximation is improved only if the maximal covariance $\bar{\rho}$ in the approximating process is *exactly* equal to ρ.

One consequence of the convergence of N_n for any r, is that asymptotically the locations of the k-th largest values are uniformly distributed. An interesting question is to find a bound for the rate of this convergence, which is conjectured to be of the same order as the convergences treated in this paper.

Acknowledgement I want to thank Jan Lanke for many useful discussions on the subject of this paper, and Philip Hougaard for reading the manuscript.

References

[1] Berman, S.M.: Limit theorems for the maximum term in stationary normal sequences. Ann. Math. Statist. 35, 502-516 (1964).

[2] Fisher, R.A. & Tippet, L.H.C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Phil. Soc. 24, 180-190 (1928).

[3] Gumbel, E.J.: Statistics of extremes. Columbia University press, New York, 1958.

[4] Hall, P.: The rate of convergence of normal extremes. J. Appl. Prob. 16, 433-439 (1979).

[5] Kallenberg, O.: Random measures. Academic Press, New York, 1976.

[6] Leadbetter, M.R., Lindgren, G. & Rootzén, H.: Conditions for convergence in distribution of maxima of stationary normal processes. Stochastic Processes Appl. 8, 131-139 (1978).

[7] Leadbetter, M.R., Lindgren, G. & Rootzén, H.: Extremal and related properties of stationary processes, Part. I. Report 1979-2, Dept. of Math. Statistics, Umeå, Sweden.

[8] Leadbetter, M.R., Lindgren, G. & Rootzén, H.: Extremal and related properties of stationary processes. Springer, Berlin, 1982.

[9] Nair, A.K.: Asymptotic distribution and moments of sample extremes. Ann. Probability 9, No.1 (1981).

[10] Serfling, R.J.: A general Poisson approximation theorem. Ann. Probability 3, 726-731 (1976).

[11] Slepian, D.: The one-sided barrier problem for Gaussian noise. Bell Syst. Techn. J. 41, 468-501 (1962).

[12] Watts, V., Rootzén, H. & Leadbetter, M.R.: On limiting distributions of intermediate order statistics from stationary sequences. To appear in Ann. Probability.

PREPRINTS 1980

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Olkin, Ingram and Væth, Michael: Maximum Likelihood Estimation in a Two-way Analysis of Variance with Correlated Errors in One Classification.

No. 2 Skovgaard, Ib M.: Transformation of an Edgeworth Expansion by a Sequence of Smooth Functions.

No. 3 Asmussen, Søren: Equilibrium Properties of the M/G/1 Queue.

No. 4 Johansen, Søren and Keiding, Susanne: A Family of Models for the Elimination of Substrate in the Liver.

No. 5 Skovgaard, Ib M.: Edgeworth Expansions of the Distributions of Maximum Likelihood Estimators.

No. 6 Tjur, Tue: A Connection Between Rasch's Item Analysis Model and a Multiplicative Poisson Model.

No. 7 Lauritzen, Steffen L. : Examples of Extreme Point Models in Statistics.

PREPRINTS 1981

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Johansen, Søren: Asymptotic Inference in Random Coefficient Regression Models.

No. 2 Asmussen, Søren: On the Role of a Certain Eigenvalue in Estimating the Growth Rate of a Branching Process.

No. 3 Lauritzen, Steffen L.: Time Series Analysis in 1880. A Discussion of Contributions made by T.N. Thiele.

No. 4 Asmussen, Søren: Conditioned Limit Theorems Relating a Random Walk to its Associate with Applications to Risk Reserve Processes and the GI/G/1 Queue.

No. 5 Johansen, Søren: The Statistical Analysis of a Markov Branching Process.

No. 6 Asmussen, Søren: Time - Dependent Approximations in some Queueing Sysetems with Imbedded Markov Chains Related to Random Walks.

No. 7 Skovgaard, Ib M.: A Second-Order Investigation of Asymptotic Ancillarity.

No. 8 The Rate of Convergence of Extremes of Stationary Normal Sequences.