LIMIT THEOREMS FOR THE MAXIMUM TERM IN STATIONARY SEQUENCES[1]

By Simeon M. Berman

Columbia University

1. Introduction and summary. Let $\{X_n, n = 0, \pm 1, \cdots\}$ be a real valued discrete parameter stationary stochastic process on a probability space $(\Omega, \mathfrak{F}, P)$; for each $n = 1, 2, \cdots$, let $Z_n = \max(X_1, \cdots, X_n)$. We shall find general conditions under which the random variable Z_n has a limiting distribution function (d.f.) as $n \to \infty$; that is, there exist sequences $\{a_n\}$ and $\{b_n\}$, $a_n > 0$, and a proper nondegenerate d.f. $\Phi(x)$ such that

$$\lim_{n\to\infty} P\{Z_n \leqq a_n x + b_n\} = \Phi(x) \tag{1.1}$$

for each x in the continuity set of $\Phi(x)$.

The simplest type of stationary sequence $\{X_n\}$ is one in which the random variables are mutually independent with some common d.f. $F(x)$. In this case, Z_n has the d.f. $F^n(x)$ and (1.1) becomes

$$\lim_{n\to\infty} F^n(a_n x + b_n) = \Phi(x). \tag{1.2}$$

It is well known that in (1.2) $\Phi(x)$ is of one of exactly three types; necessary and sufficient conditions on F for the validity of (1.2) are also known [9]. The three types are usually called extreme value d.f.'s [10].

Theorem 2.1 gives the limiting d.f. of Z_n in a stationary sequence satisfying a certain condition on the upper tail of the conditional d.f. of X_1, given the "past" of sequence: the limiting d.f. is a simple mixture of extreme value d.f.'s of a single type. These are the same kind of d.f.'s found by us [3] to be the limiting d.f.'s of maxima in sequences of exchangeable random variables. The conditions of Theorem 2.1 are specialized to exchangeable and Markov sequences, and Theorem 2.2 extends the methods of Theorem 2.1 to general (not necessarily stationary) Markov sequences. It is shown that stationary Gaussian sequences, except for the trivial case of independent, identically distributed Gaussian random variables, do not obey the requirements of the hypothesis of Theorem 2.1: hence, Sections 3, 4, and 5 are devoted to a detailed study of the maximum in a stationary Gaussian sequence.

Theorem 3.1 provides conditions on the rate of convergence of the covariance sequence to 0 which are sufficient for Z_n to have the same extreme value limiting d.f. as in the case of independence, namely, $\exp(-e^{-x})$. The relation of these conditions to the spectral d.f. of the process is also discussed. A weaker condition on the covariance sequence ensures the "relative stability in probability"

Received 8 August 1963; revised 10 November 1963.

[1] Research supported by National Science Foundation Grant NSF-GP-316 at Columbia University.

of Z_n (Theorem 4.1). Theorem 5.1 describes the behavior of Z_n when the spectrum has a discrete component with "not too many large jumps" and a "smooth" continuous component: when properly normalized, Z_n converges in probability to a random variable representing the maximum of the process corresponding to the discrete spectral component. A special case was given by us in [2].

We now summarize some known results used in the sequel. The extreme value d.f.'s are continuous, so that (1.2) holds for *all* x; furthermore, this holds if and only if it holds for all x satisfying $0 < \Phi(x) < 1$. (1.2) implies that for all such x

$$0 < F^n(a_n x + b_n) < 1, \qquad \text{for all large } n,$$

and

$$\lim_{n\to\infty} F(a_n x + b_n) = 1. \tag{1.3}$$

Let x_∞ be the supremum of all real numbers x' for which $F(x') < 1$; then, for all x satisfying $0 < \Phi(x) < 1$, we have

$$\lim_{n\to\infty} a_n x + b_n = x_\infty . \tag{1.4}$$

From (1.3), and the asymptotic relation $-\log F \sim (1 - F)$, $F \to 1$, we see that (1.2) holds if and only if

$$\lim_{n\to\infty} n[1 - F(a_n x + b_n)] = -\log \Phi(x). \tag{1.5}$$

2. Limiting d.f. **for the non-Gaussian case.** Let $\{X_n, n = 0, \pm 1, \cdots\}$ be a stationary sequence on $(\Omega, \mathfrak{F}, P)$. Let $\mathfrak{F}_k$ be the sub-σ-field of $\mathfrak{F}$ generated by the family of random variables

$$\{X_n, -\infty < n \leqq k\}, \quad k = 0, \pm 1, \cdots; \qquad \text{put} \quad \mathfrak{F}_{-\infty} = \bigcap_k \mathfrak{F}_k \quad \text{and}$$

$$F_k(x \mid \mathfrak{F}_{k-1}) = P\{X_k \leqq x \mid \mathfrak{F}_{k-1}\}, \quad k = 0, \pm 1, \cdots .$$

The following theorem was suggested by the "comparison" technique, first used by Lévy [11] and extended by Loève [12], for the calculation of the limiting d.f. of partial sums of dependent random variables.

THEOREM 2.1. *Let* $\{X_n, n = 0, \pm 1, \cdots\}$ *be a stationary sequence on* $(\Omega, \mathfrak{F}, P)$ *and* $F(x)$ *a* d.f. *for which* (1.2) *holds for properly selected sequences* $\{a_n\}$ *and* $\{b_n\}$. *If there exists a nonnegative random variable* Y *on* Ω *such that*

(i) Y *is an invariant random variable* (for the definition, see [7], p. 457);

(ii) Y *has a finite expectation, i.e.,* $EY < \infty$;

(iii)

$$\lim_{x\to x_\infty} E\,|[1 - F_1(x \mid \mathfrak{F}_0)]/[1 - F(x)] - Y| = 0, \tag{2.1}$$

where x_∞ *is defined in terms of* F; *then,*

$$\lim_{n\to\infty} P\{Z_n \leqq a_n x + b_n\} = E[\Phi(x)]^Y, \tag{2.2}$$

at all continuity points of the latter. The conditions under which the right hand side is a proper, non-degenerate d.f. *are given in* [3].

PROOF. We shall first prove the following inequality.

$$|P\{Z_n \leqq x\} - EF^{nY}(x)| \leqq nE\,|F_1(x \mid \mathfrak{F}_0) - F^Y(x)|, \tag{2.3}$$

for $-\infty < x < \infty$, $n = 1, 2, \cdots$, and every nonnegative, invariant random variable Y, satisfying (i), (ii), and (iii). Let $U_1, U_2, \cdots$ be a sequence of random variables on $(\Omega, \mathfrak{F}, P)$ which are conditionally, given the sequence $\{X_n\}$, mutually independent with the common conditional d.f. $F^Y(x)$, that is, for every finite set of x's, $x_1, \cdots, x_m$,

$$P\{U_j \leqq x_j, j = 1, \cdots, m \mid \{X_n\}\} = \prod_{j=1}^{m} F^Y(x_j). \tag{2.4}$$

This definition of the conditional d.f. is justified because an invariant random variable is $\mathfrak{F}_{-\infty}$-measurable ([7], p. 459); furthermore, by this very fact, we have

$$P\{U_j \leqq x_j, j = 1, \cdots, m \mid \{X_n\}\} = P\{U_j \leqq x_j, j = 1, \cdots, m \mid \mathfrak{F}_{-\infty}\}. \tag{2.5}$$

Fix n and x, and define the events $A_j, B_j, j = 1, \cdots, n$ as $A_j = \{X_j \leqq x\}$ and $B_j = \{U_j \leqq x\}$; then,

$$\{Z_n \leqq x\} = \bigcap_{j=1}^{n} A_j, \qquad \{\max(U_1, \cdots, U_n) \leqq x\} = \bigcap_{j=1}^{n} B_j;$$

hence, the left hand side of (2.3) is $|P(\cap A_j) - P(\cap B_j)|$. The difference $P(\cap A_j) - P(\cap B_j)$ may be expressed as the sum of n differences of probabilities as

$$\begin{aligned} &\left[P\left(\bigcap_{j=1}^{n} A_j\right) - P\left(\bigcap_{j-1}^{n-1} A_j \cap B_n\right)\right] \\ &\quad + \left[P\left(\bigcap_{j=1}^{n-1} A_j \cap B_n\right) - P\left(\bigcap_{j=1}^{n-2} A_j \cap \bigcap_{j=n-1}^{n} B_j\right)\right] + \cdots \\ &\quad + \left[P\left(\bigcap_{j=1}^{k} A_j \cap \bigcap_{j=k+1}^{n} B_j\right) - P\left(\bigcap_{j=1}^{k-1} A_j \cap \bigcap_{j=k}^{n} B_j\right)\right] + \cdots \\ &\quad + \left[P\left(A_1 \cap \bigcap_{j=2}^{n} B_j\right) - P\left(\bigcap_{j=1}^{n} B_j\right)\right]. \end{aligned} \tag{2.6}$$

We seek a common bound on the terms in this sum. $\bigcap_{j=1}^{k-1} A_j$ is in $\mathfrak{F}_{k-1}$; hence,

$$\begin{aligned} &P\left(\bigcap_{j=1}^{k} A_j \cap \bigcap_{j=k+1}^{n} B_j\right) - P\left(\bigcap_{j=1}^{k-1} A_j \cap \bigcap_{j=k}^{n} B_j\right) \\ &\quad = \int_{\cap_{j=1}^{k-1} A_j} \left[P\left(A_k \cap \bigcap_{j=k+1}^{n} B_j \mid \mathfrak{F}_{k-1}\right) - P\left(\bigcap_{j=k}^{n} B_j \mid \mathfrak{F}_{k-1}\right)\right] dP. \end{aligned} \tag{2.7}$$

By the rule of composition of conditional expectations [7] p. 37, and by (2.4) and (2.5), there follows

$$P\left(\bigcap_{j=k}^{n} B_j \mid \mathfrak{F}_{k-1}\right) = E\left[P\left(\bigcap_{j=k}^{n} B_j \mid \{X_n\}\right) \mid \mathfrak{F}_{k-1}\right]$$
(2.8)
$$= E\left[P\left(\bigcap_{j=k}^{n} B_j \mid \mathfrak{F}_{-\infty}\right) \mid \mathfrak{F}_{k-1}\right] = [F(x)]^{(n-k+1)Y}.$$

Let I_{A_k} be the indicator function of A_k; then, since $A_k \varepsilon\, \mathfrak{F}_k$,

$$P\left(A_k \cap \bigcap_{j=k+1}^{n} B_j \mid \mathfrak{F}_{k-1}\right) = E\left[P\left(A_k \cap \bigcap_{j=k+1}^{n} B_j \mid \mathfrak{F}_k\right) \mid \mathfrak{F}_{k-1}\right]$$
$$= E\left[I_{A_k} P\left(\bigcap_{j=k+1}^{n} B_j \mid \mathfrak{F}_k\right) \mid \mathfrak{F}_{k-1}\right];$$

the latter, by (2.8), is equal to $P(A_k \mid \mathfrak{F}_{k-1})\,[F(x)]^{(n-k)Y}$. From this expression and from (2.8) one deduces that the right hand side of (2.7) is bounded in absolute value by $E\,|\,P(A_k \mid \mathfrak{F}_{k-1}) - F^Y(x)\,|$. The stationarity of $\{x_n\}$ and the invariance of Y imply that the sequence of random variables $\{\,|\,P(A_k \mid \mathfrak{F}_{k-1}) - F^Y(x)\,|,\ k = 0, \pm 1, \cdots\}$ is stationary; hence,

$$E|P(A_k \mid \mathfrak{F}_{k-1}) - F^Y(x)| = E|P(A_1 \mid \mathfrak{F}_0) - F^Y(x)|, \qquad k = 1, \cdots, n.$$

Each term in (2.6) is bounded by this quantity, and so (2.3) is confirmed.

Let x satisfy $0 < \Phi(x) < 1$ and let $a_n x + b_n$ take the place of the variable x in (2.3). We shall show that the right hand side tends to 0 for $n \to \infty$; this will complete the proof. The right hand side is written as the product of two factors

$$n[1 - F(a_n x + b_n)] \cdot \frac{E|F_1(a_n x + b_n \mid \mathfrak{F}_0) - F^Y(a_n x + b_n)|}{1 - F(a_n x + b_n)}.$$

By (1.5), the first factor converges to a positive finite limit. The second factor may be written as

$$E\left|\frac{1 - F_1(a_n x + b_n \mid \mathfrak{F}_0)}{1 - F(a_n x + b_n)} - \frac{1 - F^Y(a_n x + b_n)}{1 - F(a_n x + b_n)}\right|,$$

and, by the triangle inequality, is dominated by

$$E\left|\frac{1 - F_1(a_n x + b_n \mid \mathfrak{F}_0)}{1 - F(a_n x + b_n)} - Y\right| + E\left|\frac{1 - F^Y(a_n x + b_n)}{1 - F(a_n x + b_n)} - Y\right|.$$

By (1.4) and (2.1), the first term above tends to 0. To prove that the second term also does, it suffices for us, by (1.3), to show that

$$\lim_{u \nearrow 1} E|[(1 - u^Y)/(1 - u)] - Y| = 0$$

for every nonnegative random variable Y with a finite expectation. By the law of the mean, there exists a (random) value $\bar{u}$, $u < \bar{u} < 1$, such that $1 - u^Y = Y\bar{u}^{Y-1}(1 - u)$; hence,

$$|[(1 - u^Y)/(1 - u)] - Y| = Y|\bar{u}^{Y-1} - 1| \leqq Y/u.$$

The assertion follows by taking expectations, letting $u \to 1$, and applying the dominated convergence theorem.

We shall now discuss the validity and implications of (2.1) in special kinds of stationary sequences.

Exchangeable sequences. It was noted in Section 1 that the class of limiting d.f.'s in (2.2) coincides with a subclass of the class of limiting d.f.'s for the maximum of exchangeable random variables [3]. There is a connection between (2.1) and the necessary and sufficient condition for convergence in the exchangeable case. The latter condition can be shown to be equivalent to the *convergence in distribution* of $[1 - F_1(x \mid \mathfrak{F}_{-\infty})]/[1 - F(x)]$ as $x \to x_\infty$; hence, for the special case of an exchangeable sequence, Theorem 2.1 implies the previously given necessary and sufficient condition.

Stationary Gaussian sequences. (2.1) does not hold in this case unless $Y \equiv 0$ or the covariance sequence is identically 0. Let $\{X_n\}$ be stationary and Gaussian with

$$EX_n = 0, \qquad EX_n^2 = 1, \qquad E(X_1^2 \mid \mathfrak{F}_0) - E^2(X_1 \mid \mathfrak{F}_0) = \sigma^2,$$

where σ^2 is the conditional variance given the "past", and $0 < \sigma < 1$. Let $\psi(x)$ be the standard Gaussian d.f. It is well known that

$$F_1(x \mid \mathfrak{F}_0) = \psi([x - E(X_1 \mid \mathfrak{F}_0)]/\sigma).$$

From the asymptotic formula for $1 - \psi(x)$, $x \to \infty$ ([8] p. 166), one obtains

$$1 - F_1(x \mid \mathfrak{F}_0) \sim (1 - \psi(x/\sigma)) \cdot \exp[-(1/2\sigma^2)E^2(X_1 \mid \mathfrak{F}_0) + (x/\sigma)E(X_1 \mid \mathfrak{F}_0)].$$

When divided by any decreasing function of x, this expression does not have a limiting d.f. which is neither improper nor degenerate at 0 unless $E(X_1 \mid \mathfrak{F}_0) = 0$ with probability 1; *a fortiori*, this expression cannot converge in the mean to a finite random variable which is not degenerate at 0, unless $E(X_1 \mid \mathfrak{F}_0) = 0$; hence, there exists no d.f. $F(x)$ and appropriate random variable Y (not identically 0) for which (2.1) holds, unless $E(X_1 \mid \mathfrak{F}_0) = 0$. In the latter case, the covariance sequence is identically 0.

Markov sequences. Let $\{X_n\}$ be a stationary Markov sequence: the finite dimensional d.f.'s of the sequence are determined by the (stationary) absolute d.f. and the transition probability function $F(x \mid X_k) = P\{X_{k+1} \leqq x \mid X_k\}$. By the Markov property, the expression $F_1(x \mid \mathfrak{F}_0)$ in (2.1) is equal to $F(x \mid X_0)$, and (2.1) becomes

$$\lim_{x \to x_\infty} E\,|[(1 - F(x \mid X_0))/(1 - F(x))] - Y| = 0. \tag{2.9}$$

If $\{X_n, n = 0, 1, \cdots\}$ is a Markov sequence and if the d.f. of X_0 is not a stationary d.f., then $\{X_n, n = 0, 1, \cdots\}$ is not a stationary sequence; nevertheless, a suitable version of (2.9) is still sufficient for the conclusion of Theorem 2.1 if the transition function is subject to certain conditions which guarantee the existence of a stationary initial d.f. We shall prove a simple form of such a limit theorem; more general forms can be established with the same techniques.

THEOREM 2.2. *Let* $\{X_n, n = 0, 1, \cdots\}$ *be a Markov sequence with the stationary transition function* $F(x \mid X)$. *We assume that there exist positive numbers* γ *and* ρ, $\rho < 1$, *such that for any initial* d.f. *and for any bounded random variable* f, $|f| \leqq M < \infty$, *which is measurable with respect to the* σ-*field generated by* $\{X_n, n \geqq k\}$, *we have*

$$|E'f - Ef| \leqq 2\gamma M\rho^k, \tag{2.10}$$

where E' *and* E *are the expectation operators with respect to the measures induced by the initial d.f. and the stationary d.f.* (*which is assumed to exist*), *respectively. If there exists* F *satisfying* (1.1) *such that for every* u

$$\lim_{x \to x_\infty} F(x \mid u) = 1, \tag{2.11}$$

and if (2.9) *holds with* $Y \equiv 1$, *then the conclusion* (2.2) *follows with* $Y \equiv 1$.

PROOF. We remark that (2.10) and the existence of a stationary d.f. are implied by condition (D_0), used to prove the central limit theorem in [7] p. 221.

The inequality

$$|P\{Z_n \leqq x\} - F^n(x)| \leqq \sum_{j=1}^{n} E' \,|F(x \mid X_{j-1}) - F(x)| \tag{2.12}$$

for any initial d.f. is constructed by using the method of proof of Theorem 2.1 with $Y \equiv 1$ and $F_k(x \mid \mathfrak{F}_{k-1}) = F(x \mid X_{k-1})$. If the initial d.f. of X_0 coincides with the stationary d.f., the assertion of the theorem follows from (2.9). Even if the initial d.f. is not the stationary d.f., then, as we shall show, the sum on the right hand side of (2.12) is asymptotically ($n \to \infty$, $x \to x_\infty$) independent of the initial d.f., so that the behavior of (2.12) is the same as in the case of a stationary initial d.f.

Let us estimate the sum of differences

$$\sum_{j=1}^{n} \{E' \,|F(x \mid X_{j-1}) - F(x)| - E\,|F(x \mid X_{j-1}) - F(x)|\}. \tag{2.13}$$

Let m be an arbitrary fixed position integer: consider the part of the sum (2.13) over indices from $m + 1$ to $n > m$. Put $f = |F(x \mid \cdot) - F(x)|$ and apply (2.10) with $M = 1$: the part of the indicated sum is bounded above by

$$2\gamma \sum_{k=m+1}^{n} \rho^k \sim 2\gamma\, \rho^{m+1}(1 - \rho)^{-1}, \qquad n \to \infty.$$

This can be made small by choosing m to be large. Now insert $a_n x + b_n$ in the place of x in the partial sum from 1 to m in (2.13): each term of the finite partial sum tends to 0 by (1.4) and (2.11).

A more general formulation of this theorem would let Y be a random variable which is measurable with respect to the σ-field generated by the "ergodic decomposition" of the state space [7, p. 232].

3. Limiting extreme value d.f. **for the maximum in a Gaussian sequence.** Let $\{X_n, n = 0, \pm 1, \cdots\}$ be a stationary Gaussian sequence (henceforth abbreviated as "S.G. sequence") with

$$EX_n = 0, \qquad EX_n^2 = 1, \qquad n = 0, \pm 1, \cdots \tag{3.1}$$

and with covariance sequence $\{r_n, n = 1, 2, \cdots\}$:

$$r_n = EX_0X_n, \qquad n = 1, 2, \cdots. \tag{3.2}$$

Let sequences $\{a_n\}$ and $\{b_n\}$ be defined as

$$\begin{aligned} a_n &= (2 \log n)^{-\frac{1}{2}} \\ b_n &= (2 \log n)^{\frac{1}{2}} - \tfrac{1}{2}(2 \log n)^{-\frac{1}{2}}(\log \log n + \log 4\pi). \end{aligned} \tag{3.3}$$

It is known that when $r_n \equiv 0$ (the X's are mutually independent)

$$\lim_{n\to\infty} P\{Z_n \leqq a_n x + b_n\} = \exp(-e^{-x}) \tag{3.4}$$

for all x, [5] p. 375. This result was generalized to the case $r_n = 0$ for all but finitely many n (case of m-dependence) by Watson [14]. In what follows we shall replace these conditions by *asymptotic* conditions on r_n for $n \to \infty$.

It is necessary to collect some results on probabilities associated with k-dimensional Gaussian d.f.'s. Let (r_{ij}) be a $k \times k$ symmetric positive definite matrix with 1's along the diagonal, and let $\varphi_k(x_1, \cdots, x_k; r_{ij}, 1 \leqq i < j \leqq k)$ be the k-dimensional Gaussian density function with mean vector $\mathbf{0}$ and covariance matrix (r_{ij}); φ_k is a function of the x's and the $k(k-1)/2$ parameters r_{ij}. Define:

$$Q_k(c; \{r_{ij}\}) = \int_{-\infty}^{c} \cdots \int_{-\infty}^{c} \varphi_k(x_1, \cdots, x_k; \{r_{ij}\}) \prod_{j=1}^{k} dx_j \tag{3.5}$$

The partial derivative with respect to r_{hl} is obtained by a direct adaptation of a method of Slepian ([13] Section 2.1):

$$\begin{aligned} &\frac{\partial Q_k}{\partial r_{hl}} \\ &= \int_{-\infty}^{c} \cdots \int_{-\infty}^{c} \varphi_k(x_1, \cdots, x_{h-1}, c, x_{h+1}, \cdots, x_{l-1}, c, x_{l+1}, \cdots, x_k; \{r_{ij}\}) \\ &\qquad \cdot \prod_{j\neq h, j\neq l} dx_j. \end{aligned} \tag{3.6}$$

This expression is positive; thus Q_k is an increasing function of the arguments $\{r_{ij}\}$; therefore, for δ, $0 < \delta < 1$, such that $r_{ij} \leqq \delta$, $1 \leqq i < j \leqq k$, we have

$$Q_k(c; \{r_{ij}\}) \leqq Q_k(c; \{\delta\}). \tag{3.7}$$

If the upper limit of integration $(c, \cdots, c)$ in the $(k-2)$-fold integral in (3.6) is replaced by $(\infty, \cdots, \infty)$, then the value of the integral is increased. Its augmented value is obtained by integration over the $(k-2)$ variables, and is equal to

$$\varphi_2(c, c; r_{hl}) = (2\pi)^{-1}(1 - r_{hl}^2)^{-\frac{1}{2}} \exp[-c^2/(1 + r_{hl})]; \tag{3.8}$$

hence,

$$(\partial/\partial r_{hl})\, Q_k(c; \{r_{ij}\}) \leqq \varphi_2(c, c; r_{hl}). \tag{3.9}$$

Suppose that the covariance r_{ij} is a function of the difference $j - i$, $i < j$; we write $r_{j-i} = r_{ij}$. The function P_k is defined as

$$P_k(c; r_1, \cdots, r_{k-1}) = Q_k(c; \{r_{ij}\}); \tag{3.10}$$

the partial derivatives are given by the chain rule as

$$\partial P_k/\partial r_j = \sum_{l-h=j} \partial Q_k/\partial r_{hl}. \tag{3.11}$$

In the following lemma, we compare the d.f. of the maximum in a S.G. sequence with a general covariance sequence to the maximum in the particular S.G. sequence with $r_n \equiv 0$.

LEMMA 3.1. *Let* $\{X_n\}$ *and* $\{Y_n\}$ *be* S.G. *sequences satisfying*

$$EX_n = EY_n = 0, \qquad EX_n^2 = EY_n^2 = 1, \; n = 0 \pm 1, \cdots$$
$$EX_0X_n = r_n, \; EY_0Y_n = 0, \qquad n = 1, 2, \cdots.$$

For every real number c, and every positive integer n,

$$|P\{\max(X_1, \cdots, X_n) \leqq c\} - P\{\max(Y_1, \cdots, Y_n) \leqq c\}| \leqq \sum_{j=1}^{n-1} |r_j|\,(n-j)\,\varphi_2(c, c; |r_j|). \tag{3.12}$$

PROOF. From the definition of P_k (3.10), there follows

$$P\{\max(X_1, \cdots, X_n) \leqq c\} = P_n(c; r_1, \cdots, r_{n-1})$$
$$P\{\max(Y_1, \cdots, Y_n) \leqq c\} = P_n(c; 0, \cdots, 0).$$

By the law of the mean, there exist numbers r_i' between 0 and r_i, $i = 1, \cdots, n-1$, such that

$$P_n(c; r_1, \cdots, r_{n-1}) - P_n(c; 0, \cdots, 0) = \sum_{j=1}^{n-1} r_j(\partial P_n/\partial r_j)(c; r_1', \cdots, r_{n-1}'). \tag{3.13}$$

(If $r_i = 0$ for some i, put $r_i' = 0$.) The sum in (3.11) has $k - j$ terms, and, by (3.9), each term is bounded by $\varphi_2(c, c; r_{hl})$; hence, (3.12) follows from the monotonicity of φ_2 as a function of r.

LEMMA 3.2. *Let* $\{X_n\}$ *be a* S.G. *sequence satisfying* (3.1) *and* (3.2). *If*

$$(3.14)\quad \lim_{n\to\infty} \sum_{j=1}^{n-1} |r_j|\ (n-j)(1-r_j^2)^{-\frac{1}{2}} n^{-2/(1+|r_j|)} (\log n)^{1/(1+|r_j|)} = 0,$$

then (3.4) *holds for* Z_n.

PROOF. In view of the result (3.4) for the special case $r_n \equiv 0$, we apply Lemma 3.1, put $c = a_n x + b_n$, and let $n \to \infty$. From (3.3) we obtain

$$(a_n x + b_n)^2 = 2 \log n - \log \log n + O(1).$$

If this is substituted for c in the right hand side of (3.12), the latter is dominated by a constant multiple of the left hand side of (3.14).

Sufficient conditions on $\{r_n\}$ for the validity of (3.14) are now given.

THEOREM 3.1. (3.14) *holds if either*

$$(3.15)\quad \lim_{n\to\infty} r_n \log n = 0,$$

or

$$(3.16)\quad \sum_{n=1}^{\infty} r_n^2 < \infty.$$

PROOF. (3.15) or (3.16) imply $r_n \to 0$; therefore, the stationarity of $\{X_n\}$ excludes the possibility that either $|r_n| = 1$ for any n or that $\limsup_{n\to\infty} |r_n| = 1$; hence there is a positive number δ such that

$$(3.17)\quad \sup_n |r_n| = \delta < 1.$$

Define: $\delta(n) = \sup_{k\geqq n} |r_k|$; then (3.15) implies

$$(3.18)\quad \lim_{n\to\infty} \delta(n) \log n = 0.$$

By (3.17), $(1 - r_j^2)^{-\frac{1}{2}}$ is bounded above by $(1 - \delta^2)^{-\frac{1}{2}}$, so that (3.14) is implied by

$$(3.19)\quad (\log n) \sum_{j=1}^{n-1} |r_j|\ (n-j) n^{-2/(1+|r_j|)} \to 0.$$

Let α be a real number satisfying $0 < \alpha < (1-\delta)/(1+\delta)$; let $[u]$ be the integral part of the real number u. Let n be large: split the sum in (3.19) into a first sum over indices up to $[n^\alpha]$ and a second sum over indices from $[n^\alpha] + 1$ to $n - 1$. The first sum is dominated by $\delta n^{1+\alpha-2/(1+\delta)}$, which when multiplied by $(\log n)$, tends to 0 because of the choice of α. The second sum, multiplied by $(\log n)$, may be written as

$$n^{-2}(\log n) \sum_{j=[n^\alpha]+1}^{n-1} |r_j|\ (n-j)\exp[(2 \log n)\ |r_j|/(1 + |r_j|\ )],$$

and is dominated by $(\log n)\ \delta(n^\alpha) \exp\ [2\delta(n^\alpha) \log n]$, which, by (3.18), tends to 0. The sufficiency of (3.15) is verified.

In order to demonstrate the sufficiency of (3.16), we have only to show that the *second* sum in (3.19), multiplied by (log n), tends to 0, as the proof for the *first* sum is the same as before. By the Cauchy-Schwarz inequality, the *square* of the second sum, multiplied by $(\log n)^2$, is dominated by

$$(\log n)^2 \cdot \sum_{j=[n^\alpha]+1}^{n-1} r_j^2 \cdot \sum_{j=[n^\alpha]+1}^{n-1} (n-j)^2 n^{-4/(1+|r_j|)};$$

the latter is dominated by

$$(\log n)^2 \cdot n^{3-4/(1+\delta(n^\alpha))} \sum_{j=[n^\alpha]+1}^{n-1} r_j^2,$$

which, by (3.16), tends to 0.

It is of interest to give conditions on the spectral d.f. of the process which are sufficient for (3.15) and (3.16). A simple sufficient condition for (3.15) is that the spectral d.f. be differentiable, and that its derivative (the spectral density function) satisfy a Lipschitz condition of order α, for some $\alpha > 0$. In fact, the stated condition implies $r_n = O(n^{-\alpha})$, which implies (3.15), [15] p. 46. A sufficient condition for (3.16) is, by the Bessel inequality, the square integrability of the spectral density function [15] p. 13.

We note that Theorem 3.1 contains, as a special case, our earlier result on the "law of large numbers" [1]. It would be interesting to make a corresponding generalization of Cramér's result [6] in the continuous parameter case.

It may be observed that the assumption of stationarity is not critical and that Theorem 3.1 may be generalized to nonstationary Gaussian sequences. Finally, there is no apparent overlap between our results and those of Chibisov [4], which were announced without proof.

4. Relative stability of the maximum in the Gaussian case. The maximum Z_n is said to be *relatively stable in probability* if there is a sequence of constants A_n such that $Z_n/A_n \to 1$ in probability [9]. In the S.G. sequence relative stability is implied by (3.4), with $A_n = (2 \log n)^{\frac{1}{2}}$. One may ask if it is possible to extend relative stability of Z_n to sequences satisfying conditions weaker than either (3.15) or (3.16). This is done in the following theorem.

THEOREM 4.1. *Let* $\{X_n, n = 0, \pm 1, \cdots\}$ *be a* S.G. *sequence satisfying* (3.1) *and* (3.2). *If*

$$\lim_{n\to\infty} r_n = 0, \tag{4.1}$$

then

$$Z_n/(2 \log n)^{\frac{1}{2}} \to 1 \text{ in probability.} \tag{4.2}$$

PROOF. (4.2) is equivalent to

$$\begin{aligned} &P\{Z_n > (2 \log n)^{\frac{1}{2}} (1 + \epsilon)\} \to 0 \\ &P\{Z_n > (2 \log n)^{\frac{1}{2}} (1 - \epsilon)\} \to 1, \qquad \epsilon > 0. \end{aligned} \tag{4.3}$$

The first part of (4.3) has a direct proof:

$$P\{Z_n > (2 \log n)^{\frac{1}{2}}(1 + \epsilon)\} = P\left(\bigcup_{k=1}^{n} \{X_k > (2 \log n)^{\frac{1}{2}}(1 + \epsilon)\}\right)$$

$$\leqq \sum_{k=1}^{n} P\{X_k > (2 \log n)^{\frac{1}{2}}(1 + \epsilon)\} = n[1 - \psi((2 \log n)^{\frac{1}{2}}(1 + \epsilon))] \to 0,$$

where the last assertion follows from the well known expansion for $1 - \psi(x)$, $x \to \infty$ ([8] p. 166).

The nontrivial part of the proof is the verification of the second part of (4.3). Let ρ be a real number satisfying $0 < \rho < 1$, and let $\{V_n, n = 0, \pm 1, \cdots\}$ be a S.G. sequence satisfying (3.1) and (3.2) with $r_n = \rho$, $n = 1, 2, \cdots$. ($\{V_n\}$ is a family of equally correlated Gaussian random variables.) We have shown in a previous note [2] that $\max(V_1, \cdots, V_n) - (2(1 - \rho) \log n)^{\frac{1}{2}}$ converges in probability, as $n \to \infty$, to a random variable which has a Gaussian d.f. with mean 0 and variance ρ; hence $\max(V_1, \cdots, V_n)/((2(1 - \rho) \log n)^{\frac{1}{2}} \to 1$ in probability, so that

$$P\{\max(V_1, \cdots, V_n) > (2(1 - \rho)\log n)^{\frac{1}{2}}(1 - \epsilon)\} \to 1, \qquad \epsilon > 0. \tag{4.4}$$

The main idea of the proof of the second part of (4.3) is showing that, for any $\rho > 0$, Z_n is "asymptotically stochastically larger" than $\max(V_1, \cdots, V_n)$, so that (4.3) follows from (4.4).

By (4.1), there exists for every ρ, $0 < \rho < 1$, an integer m such that

$$|r_n| \leqq \rho, \qquad n \geqq m. \tag{4.5}$$

The subsequence $\{X_{nm}, n = 0, \pm 1, \cdots\}$ is also a S.G. sequence, and its covariance sequence $\{r_{nm}, n = 1, 2, \cdots\}$ is, by (4.5), dominated by ρ; hence, by (3.7) and (4.4),

$$P\{\max_{1 \leqq j \leqq n} X_{jm} > (2(1 - \rho) \log n)^{\frac{1}{2}}(1 - \epsilon)\}$$

$$\geqq P\{\max(V_1, \cdots, V_n) > (2(1 - \rho) \log n)^{\frac{1}{2}}(1 - \epsilon)\} \to 1.$$

On the other hand, $Z_{nm} > \max_{1 \leqq j \leqq n} X_{jm}$, so that we also have

$$\lim_{n \to \infty} P\{Z_{nm} > (2(1 - \rho) \log n)^{\frac{1}{2}}(1 - \epsilon)\} = 1.$$

The convergence here is uniform in the factor $1 - \epsilon$, since the function on the left hand side is monotonic in $1 - \epsilon$ and the limit on the right hand side is continuous (constant) in $1 - \epsilon$. Using the uniform convergence, and the inequality $(x + y)^{\frac{1}{2}} - x^{\frac{1}{2}} \leqq \frac{1}{2} x^{-\frac{1}{2}} y$, $x > 0$, $y > 0$, (this is established by the law of the mean), we have

$$\liminf_{n \to \infty} P\{Z_{nm} > (2(1 - \rho) \log nm)^{\frac{1}{2}}(1 - \epsilon)\}$$

$$\geqq \lim_{n \to \infty} P\{Z_{nm} > (2(1 - \rho) \log n)^{\frac{1}{2}}(1 - \epsilon)\ [1 + \log m/2 \log n]\} = 1.$$

For any $n \geqq m$, $m\,[n/m] \leqq n \leqq m\,([n/m] + 1)$; hence, $Z_n \geqq Z_{m[n/m]}$, and so

$$\liminf_{n\to\infty} P\{Z_n > (2(1-\rho)\log n)^{\frac{1}{2}}(1-\epsilon)\}$$
$$\geqq \liminf_{n\to\infty} P\{Z_{m[n/m]} > (2(1-\rho)\log m\,([n/m]+1))^{\frac{1}{2}}(1-\epsilon)\}.$$

The latter, by the previous argument, is equal to 1; hence,

$$\lim_{n\to\infty} P\{Z_n > (2\log n)^{\frac{1}{2}}\,(1-\epsilon)(1-\rho)^{\frac{1}{2}}\} = 1.$$

Since ϵ and ρ are arbitrary, the second assertion in (4.3) holds.

The absolute continuity of the spectral d.f. is sufficient for (4.1); this is an immediate consequence of the Riemann-Lebesgue Theorem ([15] p. 45).

5. A convergence theorem for Z_n in the presence of a discrete spectral component. The conditions of Theorems 3.1 and 4.1 imply the continuity of the spectral d.f.; in fact, they require $r_n \to 0$, and, consequently, $n^{-1}(r_1^2 + \cdots + r_n^2) \to 0$, which is necessary and sufficient for the continuity of the spectrum ([7] p. 494). In this section, we study the asymptotic behavior of Z_n when the spectrum has discontinuities. First, we consider a process with a purely discrete spectrum.

LEMMA 5.1. *Let* $\{V_n, n = 0, \pm 1, \cdots\}$ *be a* S.G. *sequence whose spectral* d.f. *is a step function on* $[-\pi, \pi]$ *with a countable number of jumps of heights* c_j^2, $j = 1, 2, \cdots$, *at the respective points* λ_j, $j = 1, 2, \cdots$. *If*

$$\sum_{j=1}^{\infty} |c_j| < \infty, \tag{5.1}$$

then

$$\theta = \sup_{n\geqq 1} V_n \tag{5.2}$$

is nonnegative and finite with probability 1.

PROOF. It is known that V_n is representable with probability 1 in the form

$$V_n = \sum_{j=1}^{\infty} |c_j|\,(\xi_j \cos n\,\lambda_j + \eta_j \sin n\,\lambda_j) \tag{5.3}$$

where $\{\xi_n\}$ and $\{\eta_n\}$ are independent sequences of mutually independent standard Gaussian random variables ([7] p. 488); hence,

$$|V_n| \leqq \sum_{j=1}^{\infty} |c_j|\,(|\xi_j| + |\eta_j|), \qquad n = 1, 2, \cdots,$$

where the series converges with probability 1.

Now we consider a process composed of a process of the type in Lemma 5.1 and one of the type in Theorem 3.1.

THEOREM 5.1. *Let* $\{X_n\}$ *be a* S.G. *sequence satisfying* (3.1) *with a spectral* d.f. $G(x)$ *satisfying the following conditions*: $G(x) = G_1(x) + G_2(x)$, *where* $G_1(x)$ *is a spectral* d.f. *of the type in Lemma* 5.1, *with* $c^2 = \sum_{j=1}^{\infty} c_j^2$, $0 < c^2 < 1$; *and*

$G_2(x)$ which is a (continuous) spectral d.f. *of variation* $1 - c^2$, *has Fourier-Stieltjes coefficients*

$$r_n = 2\int_0^\pi \cos n\lambda \, dG_2(\lambda)$$

satisfying (3.15) *or* (3.16). *Then, for* $n \to \infty$,

$$Z_n - (2(1 - c^2)\log n)^{\frac{1}{2}} \to \theta \tag{5.4}$$

in probability, where θ *is defined by* (5.2).

PROOF. Let $\{U_n\}$ and $\{V_n\}$ be the (independent) S.G. sequences corresponding to the spectra G_2 and G_1, so that $X_n = U_n + V_n$, $n = 0, \pm 1, \cdots$. Let M be a fixed, arbitrary positive integer, and let the sum representing V_n in (5.3) be decomposed into a first sum $V_n^{(1)}$ over indices up to and including M, and a second sum $V_n^{(2)}$ over indices greater than M. As in the proof of Lemma 5.1, there exists a nonnegative random variable θ_M such that

$$E\theta_M = 2(2/\pi)^{\frac{1}{2}} \sum_{j=M+1}^{\infty} |c_j|, \qquad |V_n^{(2)}| \leqq \theta_M, \quad n = 0 \pm 1, \cdots; \tag{5.5}$$

hence, Z_n satisfies the double inequality

$$\max_{1 \leqq j \leqq n}(U_j + V_j^{(1)}) - \theta_M \leqq Z_n \leqq \max_{1 \leqq j \leqq n}(U_j + V_j^{(1)}) + \theta_M. \tag{5.6}$$

Suppose first that the λ's are rational multiples of 2π; this assumption will be dropped later. $\{V_n^{(1)}\}$ is a periodic sequence, so that it has a (random) period p which depends on M. Since $\{U_n\}$ and $\{V_n^{(1)}\}$ are mutually independent, $\{U_n\}$ is independent of the random variable p; thus, the sequence $\{U_{np}, n = 0, \pm 1, \cdots\}$ is, conditionally given p, a S.G. sequence with $EU_{np} = 0$, $EU_{np}^2 = 1 - c^2$, $EU_0U_{np} = r_{np}$. (3.4) implies $Z_n - (2\log n)^{\frac{1}{2}} \to 0$ in probability; if the variance of the X's is changed to $1 - c^2$, we get instead $Z_n - (2(1 - c^2)\log n)^{\frac{1}{2}} \to 0$ in probability. Now the covariance sequence $\{r_{np}\}$ satisfies (3.15) or (3.16) if $\{r_n\}$ does; hence, by our previous remark

$$\max(U_p, U_{2p}, \cdots, U_{np}) - (2(1 - c^2)\log n)^{\frac{1}{2}} \to 0$$

in probability.

The same argument shows that

$$\max(U_j, U_{p+j}, \cdots, U_{(n-1)p+j}) - (2(1 - c^2)\log n)^{\frac{1}{2}} \to 0, \tag{5.7}$$

in probability, $j = 1, \cdots, p$.

We now hold the family $\{(V_n^{(1)}, V_n^{(2)}), n = 0, \pm 1, \cdots\}$ fixed and consider the corresponding conditional probabilities. $\max_{1 \leqq j \leqq np}(U_j + V_j^{(1)})$ may be written as

$$\max_{1 \leqq k \leqq p} \max_{0 \leqq j \leqq n-1}(U_{jp+k} + V_{jp+k}^{(1)}) = \max_{1 \leqq k \leqq p}(\max_{0 \leqq j \leqq n-1} U_{jp+k} + V_k^{(1)}),$$

since $V_k^{(1)} = V_{jp+k}^{(1)}$; hence, by (5.7),

$$\max_{1 \leq k \leq p} (\max_{0 \leq j \leq n-1} U_{jp+k} + V_k^{(1)}) - (2(1 - c^2) \log np)^{\frac{1}{2}} \to \max_{1 \leq k \leq p} V_k^{(1)}.$$

in probability.

The analysis used in the proof of Theorem 4.1 shows also that

$$\max_{1 \leq j \leq n} (U_j + V_j^{(1)}) - (2(1 - c^2) \log n)^{\frac{1}{2}} \to \max_{1 \leq k \leq p} V_k^{(1)},$$

in probability; hence, by (5.6), the inequality

$$|\max_{1 \leq k \leq p} V_k^{(1)} - Z_n + (2(1 - c^2) \log n)^{\frac{1}{2}}| < \theta_M$$

will hold with probability close to 1 if n is sufficiently large. (5.1) and (5.5) show that θ_M is arbitrarily small with high probability if M is sufficiently large; furthermore, $\max_{1 \leq k \leq p} V_k^{(1)}$ is arbitrarily close to θ with high probability if M is sufficiently large. We may conclude that (5.4) holds *conditionally* in probability; hence, by bounded convergence, it also holds in probability.

The assumption that the $\lambda_j/2\pi$ are rational can be dropped. For each sequence $\{\xi_n\}$ and $\{\eta_n\}$ in (5.3), the corresponding sequence $\{V_n^{(1)}\}$ with arbitrary $\lambda_1, \cdots, \lambda_M$ is uniformly approximable by another such sequence having the same ξ's and η's but *rational* $\lambda_j/2\pi$. It now follows from Egorov's theorem that, with probability arbitrarily close to 1, the sequence $\{V_n^{(1)}\}$ with arbitrary λ_j is uniformly approximable by another such sequence with rational $\lambda_j/2\pi$; hence, the respective maxima are close with high probability. The λ_j have little effect on the sequence $\{V_n^{(2)}\}$ because θ_M in (5.5) is independent of them.

The writer is indebted to Professor Harald Cramér for several helpful conversations, to the late Professor Marek Fisz for the reference [4] and to the referee for spotting several errors in the first version of this paper.

REFERENCES

[1] Berman, Simeon M. (1962). A law of large numbers for the maximum in a stationary Gaussian sequence. *Ann. Math. Statist.* **33** 93–97.

[2] Berman, Simeon M. (1962). Equally correlated random variables. *Sankhyā* A **24** 155–156.

[3] Berman, Simeon M. (1962). Limiting distribution of the maximum term in sequences of dependent random variables. *Ann. Math. Statist.* **33** 894–908.

[4] Chibisov, D. M. (1962). Remarks on the work of N. V. Smirnov, "Limiting distribution laws for the terms of a variational series." *Theor. Probability Appl.* **7** 357–358 (Russian).

[5] Cramér, Harald (1946). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[6] Cramér, Harald (1962). On the maximum of a normal stationary stochastic process. *Bull. Amer. Math. Soc.* **68** 512–516.

[7] Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[8] Feller, William (1957). *An Introduction to Probability Theory and Its Applications*, **1** (2nd ed.). Wiley, New York.

[9] Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44** 423–453.

[10] GUMBEL, E. J. (1958). *Statistics of Extremes*. Columbia Univ. Press.
[11] LÉVY, P. (1937). *Théorie de l'Addition des Variables Aléatoires*. Gauthier-Villars, Paris.
[12] LOÈVE, MICHEL (1960). *Probability Theory*, (2nd ed.). Van Nostrand, Princeton.
[13] SLEPIAN, DAVID (1962). The one-sided barrier problem for Gaussian noise. *Bell System Tech. J.* **41** 463–501.
[14] WATSON, G. S. (1954). Extreme values in samples from m-dependent stationary stochastic processes. *Ann. Math. Statist.* **25** 798–800.
[15] ZYGMUND, A. (1959). *Trigonometric Series*. Cambridge Univ. Press.