The Annals of Statistics
1981, Vol. 9, No. 1, 66–77

SMOOTHING OF SAMPLES FOR MAXIMA

BY YASHASWINI MITTAL

Stanford University

Smoothing of the data by averaging is suggested in order to study the maximum. The maximum of the smoothed data is approximated by that of a Gaussian sample and thus is more robust against outliers.

1. Introduction. Let $X_1, \cdots, X_n$ be a sequence of random variables with mean zero and

$$Y_i = \sum_{j=(i-1)L(n)+1}^{iL(n)} X_j / \mathrm{Var}^{1/2} \left(\sum_{j=(i-1)L(n)+1}^{iL(n)} X_j\right), \quad i = 1, 2, \cdots, [n/L(n)]$$

where $L(n) \uparrow \infty$ with n and $[\cdot]$ denotes the integral value. If the central limit theorem is assumed to hold for $\{X_i\}$, then each of Y_i is asymptotically Gaussian with mean zero and variance one. With suitable assumptions on the dependence of $\{X_i\}$ and careful estimates of the error terms, we show here that the asymptotic distribution of $\Pi_n = \max_{1\le i\le[n/L(n)]} Y_i$ is the same as $\max_{1\le i\le[n/L(n)]} Z_i$ where Z_i are independent standard normal variables.

Besides being of theoretical interest, the maximum Π_n of the smoothed sample $Y_1, \cdots, Y_{[n/L(n)]}$ has practical interest. Smoothing of samples is done in practice as a routine procedure in a number of cases before looking at the maximum. Examples of this are abundant in air pollution data. As pollutant concentrations are observed with a high frequency and are known to exhibit a distinctly nonstationary behavior due to trends and weather variables, it is unrealistic to assume that they form an i.i.d. sample. Averaging is used in the hope of making the data more "Gaussian" and "independent."

From a practical point of view, it is best to work under verifiable assumptions on the data. Strictly speaking, however, there are very few "verifiable" assumptions. Suitable practical assumptions are those that ask for reasonable faith and reasonable verification from the statistician (e.g., lack of correlations is taken to mean independence or "quickly vanishing"; covariance function is taken to indicate stationarity). It is hoped that the assumptions made on the original sample $X_1, \cdots, X_n$ in the following sections would be more reasonable than the traditional assumptions in asking for the statistician's faith, even though they are no more "verifiable" than the traditional ones. We have aimed here at conditions which are reasonable, mathematically tractable and lend themselves to the investigation of robustness in the behavior of Π_n.

The method of smoothing suggested in the first paragraph above may not be practical in certain situations since it requires that the length of each block size increase with n, thus making the recomputation of the averages Y_i necessary with each addition of new data. To get around this difficulty, we have suggested in Section 3 a scheme for selection of block sizes L_i. This scheme allows a large range for selection of L_i. The block sizes L_i are required to increase to infinity with n only if i increases to infinity with n. For fixed values of i, one has a choice of selecting L_i to be fixed or a function of n, depending on what seems prudent in the particular situation.

The proofs of the theorems in Sections 2 and 3 are based on the large deviations type results for the central limit theorem. In Section 2, we do the case of the i.i.d. sample to illustrate the simple main idea of the proof. To reduce the length of the exposition, we have assumed there that all block lengths are equal. In Theorem 2.1, we give the "smallest value" of the block length $L(n)$ $(= (\ln n)^{3+\delta}$ for some $\delta > 0)$ that would achieve the intended convergence. We also find the rate of convergence there. There are a few large deviations results available

Received August 1978; revised July 1979.
AMS 1970 *subject classification.* Primary 60F05, 60F99; secondary 60G15.
Key words and phrases. Maxima, smoothed samples, central limit theorem, large deviations.

for dependent sequences. Any of these can be used to get a corresponding result of the type of Theorem 3.1. For other results on large deviations and an extensive bibliography, see Statulevičius ([8], [9]).

The averages computed in Theorem 3.1 involve the knowledge of the variances of the block sums. Since it is unreasonable to assume that these would be known in practice, we substitute the sample estimates for them in Theorem 3.2, viz., sum of squares of the block observations. Since the averages will be computed by taking the ratios (the block sum)/(sum of squares of the block observations)$^{1/2}$, care needs to be taken in dealing with very small values of the denominator. A method suggested to achieve this is that of truncating the denominator away from zero. See Section 3 for more details.

In the last section, we find a joint distribution of M_n (maximum of the original sample) and Π_n (maximum of the smoothed sample) in case $\{X_i\}$ is assumed to be a stationary Gaussian sequence with a smooth covariance function.

2. The case of i.i.d. variables. Let $\{X_i, i \geq 1\}$ be a sequence of i.i.d. random variables with $EX_i = 0$; $EX_i^2 = \sigma^2$ and $Ee^{tX_i} < \infty$ for all i and $|t| < t_0$ making all the moments of $\{X_i\}$ finite for all i. Since this section only illustrates the idea of the proof, we will average X_i over blocks of equal length $L(n)$ to produce an i.i.d. smoothed sample $Y_1, \cdots, Y_{[n/L(n)]}$. This will reduce the length of the proof considerably. We choose $L(n)$ to be the smallest value for which

$$(2\ln[n/L(n)] - \ln(4\Pi\ln[n/L(n)]))^{3+\delta} \leq L(n) \tag{2.1}$$

for some $\delta > 0$.

Remark 1. The choice of $L(n)$ in (2.1) needs to be made by trial and error. The "best" value of $L(n)$ is the one that barely satisfies the inequality. Asymptotically this choice of $L(n)$ becomes equal to about $(2\ln n)^{3+\delta}$. But for values of n even as large as 10,000, the choice in (2.1) leads to a considerably smaller value of $L(n)$ than $(2\ln n)^{3+\delta}$.

Remark 2. Note that throughout, when no mention of an argument is made in a limiting statement, as in "$L(n) \uparrow \infty$," it is to be taken to read "as $n \to \infty$." Also, the symbols "O" and "o" will be used only when $n \to \infty$.

Remark 3. Hereafter the constants c_k and b_k are defined by

$$c_k = (2\ln k)^{1/2};\ b_k = c_k - \frac{\ln(4\Pi\ln k)}{2c_k} \qquad k = 1, 2, \cdots. \tag{2.2}$$

Let us define $Y_i = \sum_{j=(i-1)L(n)+1}^{iL(n)} X_j/(\sigma(L(n))^{1/2})$; $\Pi_n = \max_{1\leq i\leq[n/L(n)]} Y_i$ and $u_n = b_{[n/L(n)]} + x/c_{[n/L(n)]}$.

Theorem 2.1. *For the quantities defined above*

$$lt_{n\to\infty} P(\Pi_n \leq u_n) = \exp(-e^{-x}) \tag{2.3}$$

for $-\infty < x < \infty$.

Proof. In the following, we keep track of both the second and the third order terms. The third order terms are of some interest.

Write $K_n = [n/L(n)]$. Since Y_i are i.i.d. $P(\Pi_n \leq u_n) = \{1 - P(Y_i > u_n)\}^{K_n}$. Using Theorem 2, page 520, of Feller [5], we have $P(Y_i > u_n) = (1 - \Phi(u_n))E_n$ where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} \exp(-u^2/2)\,du$ and $E_n = \exp\{u_n^2\lambda(u_n/L^{1/2}(n))\}[1 + o(u_n/L^{1/2}(n)]$. The function $\lambda(x)$ is defined by the equation

$$x^2\lambda(x) = \lambda_1 x^3 + \lambda_2 x^4 + \cdots,$$

λ_i being constants and $\lambda_1 = EX_i^3/6\sigma^3$. For the choice of $L(n)$ which gives asymptotic equality in (2.1), we have $u_n^3/L^{1/2}(n) \simeq u_n^{-\delta}$, and $u_n/L^{1/2}(n) \approx u_n^{-\delta-2}$. Thus

$$\begin{aligned} u_n^2\lambda(u_n/L^{1/2}(n)) &= L(n)\left\{\left(\frac{u_n}{L^{1/2}(n)}\right)^2 \lambda\left(\frac{u_n}{L^{1/2}(n)}\right)\right\} \\ &= L(n)\left\{\lambda_1\left(\frac{u_n}{L^{1/2}(n)}\right)^3 + O\left(\frac{u_n}{L^{1/2}(n)}\right)^4\right\} \\ &= \lambda_1(u_n^3/L^{1/2}(n)) + O\left(\frac{u_n^4}{L(n)}\right) \\ &= \lambda_1 u_n^{-\delta} + O(u_n^{-2(1+\delta)}). \end{aligned}$$

And

$$E_n = 1 + \lambda_1 u_n^{-\delta} + O(u_n^{-\delta_0}),$$

where $\delta_0 = \min(2\delta, \delta + 2)$. Substituting, we get

$$P(\Pi_n \le u_n) = \{1 - (1 - \Phi(u_n))(1 + \lambda_1 u_n^{-\delta} + O(u_n^{-\delta_0}))\}^{K_n}.$$

Now $K_n(1 - \Phi(u_n)) = \exp(-x)(1 + o(u_n^{-2+\theta}))$ for any $\theta > 0$ small. Thus

$$\begin{aligned} P(\Pi_n \le u_n) &= \{\Phi(u_n) - \lambda_1 u_n^{-\delta} K_n^{-1}\exp(-x)(1 + o(u_n^{-2+\theta})) + O(u_n^{-\delta_0})K_n^{-1}\exp(-x)\}^{K_n} \\ &= \Phi^{K_n}(u_n)\{1 - \lambda_1 u_n^{-\delta}\exp(-x)(K_n\Phi(u_n))^{-1} + O(u_n^{-\delta_1})K_n^{-1}\exp(-x)\}^{K_n} \end{aligned}$$

where $\delta_1 = \min(\delta_0, 2 + \delta - \theta)$ for some $\theta > 0$ small. Substituting, we get

$$P(\Pi_n \le u_n) = \Phi^{K_n}(u_n) - \lambda_1 u_n^{-\delta}\exp(-x)\Phi^{K_n-1}(u_n) + O(u_n^{-\delta_1})\exp(-x). \tag{2.4}$$

REMARK. Simple arithmetic calculations show that $\Phi^{K_n}(u_n) = \exp(-e^{-x}) + O(\exp(-x)(\ln\ln K_n)^2 u_n^{-2})$. Thus theoretically the error rate for the convergence in (2.3) is $\max\{u_n^{-\delta}; (\ln\ln K_n)^2 u_n^{-2}\}$. In practice, the second term of the right-hand side of (2.4) can be considered as the error term since $\Phi^{K_n}(u_n)$ is computable for a given sample.

This concludes the proof of Theorem 2.1 and illustrates the relationship between the choice of $L(n)$ and the error rate.

3. The main theorems. In this section we assume that the sequence of variables $\{X_i, i \ge 1\}$ is ϕ-mixing, i.e., there exists a nonincreasing sequence ϕ_n of numbers such that $\phi_n \to 0$ and

$$|P(A \cap B) - P(A)P(B)| \le \phi_n P(A) \tag{E}$$

for all $A \in \mathscr{F}_1^k$ and $B \in \mathscr{F}_{k+n}^\infty$, $k \ge 1$. By $\mathscr{F}_1^k$ and $\mathscr{F}_{k+n}^\infty$ we denote the σ-fields generated by $(X_1, \cdots, X_k)$ and $(X_{k+n}, X_{k+n+1}, \cdots)$ respectively. Let $\{L_i\}$ be a nondecreasing sequence of integers. Define $\Delta_i = \sum_{k=1}^i L_k$ and K_n is such that $\Delta_{K_n-1} < n \le \Delta_{K_n}$. We choose $\{L_i\}$ such that $\Delta_{i-1}^{\gamma_1} \le L_i < n^{\gamma_2}$ for some $0 < \gamma_1 < \gamma_2 < 1$. Notice that since $\Delta_i \ge i$, $L_i \ge i^{\gamma_1}$ and L_i is required to increase to infinity (with n) only if $i \uparrow \infty$ (with n), thus making possible the choice of fixed block sizes for fixed values of i. On the other hand, one can choose equal block sizes (i.e., $L_1 = L_2 = \cdots = L_{K_n}$), but in this case L_i will be required to increase to infinity fast enough. E.g., it is possible that $L_i = n^\gamma$ for some $0 < \gamma < 1$ and all $i = 1, 2, \cdots, K_n$, but it is not possible that $L_i = \ln n$ for all $i = 1, 2, \cdots, K_n$, since for, say, $i = n^{1/2}$, we will violate the requirement $L_i \ge \Delta_{i-1}^{\gamma_1}$.

In the first theorem we prove the sums of $\{X_i\}$ over various blocks will be normalized using its variances which are unknown in practice. In Theorem 3.2 we will replace these by sample estimates.

Let us define $Y_i = \sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j / V_i^{1/2}$ where $V_i = \mathrm{Var}(\sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j)$, $i = 1, 2, \cdots, K_n$ and $\Pi_n = \max_{1\le i\le K_n} Y_i$. We recall the definitions of c_k and b_k given by (2.2) and let $u_n = b_{K_n} + x/c_{K_n}$.

THEOREM 3.1. *Let*

$$EX_j = 0 \text{ for all } j \ge 1. \tag{3.1}$$

$$\sum_{n=1}^\infty \phi_n^{1/2} < \infty. \tag{3.2}$$

(3.3a) $$A_1 \le \sigma_n^2 \le A_2 \text{ for large } n \text{ and some constants } A_2 > A_1 > 0 \text{ where } \sigma_n^2 = \operatorname{Var}(X_n).$$

(3.3b) $$\liminf_i V_i/L_i > 0.$$

(3.4) $$E|X_j|^{2+c} \le M \text{ for all } j \ge 1 \text{ and some constants } M > 1 \text{ and } c > 2. \text{ (We will take } \gamma_1 \text{ in the choice of } L_i \text{ such that } \gamma_1 c/2 > 2.)$$

Then

(3.5) $$lt_{n\to\infty} P(\Pi_n \le u_n) = \exp(-e^{-x}) \qquad \text{for } -\infty < x < \infty.$$

REMARK. Notice that (3.1) can be replaced by the requirement that EX_j be "close enough" to some fixed known constant μ. The expected values will be "close enough" if $\sum_{j\in L_i} (EX_j - \mu)/V_i^{1/2} = o((\ln n)^{-1/2})$ for large i. This would make the exposition considerably more cumbersome and not much more practical. Thus we choose to work with the condition (3.1).

PROOF OF THEOREM 3.1. As mentioned in the introduction, we make repeated use of the results of Babu, Ghosh and Singh [1]. For convenience we reproduce their Theorems 2 and 5 below. Conditions (3.1)–(3.4) above are very similar to their conditions except somewhat stronger. The first theorem below is the large deviations result and Theorem B is stated in [1] as a moderate deviations result. We need Theorem B in case L_i is quite large.

THEOREM A. *Let* $\{X_i\}$ *be a* ϕ*-mixing sequence for which* (3.1) *and* (3.2) *hold. Also let*

(3.3)′ $$\inf_{n\ge 1} n \operatorname{Var}\left(\textstyle\sum_{j=1}^n X_j\right) > 0$$

(3.4)′ $$E|X_j|^{2+c} \le M \qquad \text{for some} \quad c > 0 \qquad \text{and} \quad M > 1.$$

Then for all $t^2 > (c+1)\ln n$ *and some* $K > 0$,

(3.6) $$|P(\eta_n \le t) - \Phi(t)| \le Kn^{-c/2} t^{-2-c} (\ln t)^{2+2c}$$

where

$$\eta_n = \textstyle\sum_{j=1}^n X_j / \operatorname{Var}^{1/2}\left(\sum_{j=1}^n X_j\right).$$

THEOREM B. *For* $\{X_j\}$ *as in Theorem* A *and* $t^2 \le (c+1)\ln n$,

(3.7) $$|P(\eta_n \le t) - \Phi(t)| \le Kn^{-\lambda} \exp(-t^2/2) + O(n^{-c/2}(1 + |t|)^{-2-c})$$

for some $\lambda > 0$.

We turn now to the proof of Theorem 3.1. We use standard arguments that have been used in the literature before. Divide $Y_1, \cdots, Y_{mK_n}$ into m blocks of length K_n each. Clip a small portion from the right-hand end of each to "separate" these intervals. On each block, find upper and lower bounds for the probability that the maximum is bounded by u_{nm} where $u_{nm} = b_{mK_n} + x/c_{mK_n}$. Finally, let K_n tend to infinity first and then take a limit as $m \to \infty$.

For any integer $m \ge 1$, let $I_j = \{(j-1)K_n + 1, (j-1)K_n + 2, \cdots, jK_n - [K_n^{1/2}]\}$, $I_j^* = \{jK_n - [K_n^{1/2}] + 1, \cdots, jK_n\}$ for $j = 1, 2, \cdots, m$ and $I = \cup_{j=1}^m I_j$; $I^* = \cup_{j=1}^m I_j^*$. Define $\Pi_n(m) = \max_{1\le i\le mK_n} Y_i$, $\Pi_{I_j} = \max_{i\in I_j} Y_i$ and $\Pi_I = \max_{i\in I} Y_i$. Now

(3.8) $$P\{\Pi_n(m) \le u_{nm}\} = \textstyle\prod_{j=1}^m P\{\Pi_{I_j} \le u_{nm}\} + P_1 + P_2$$

where

$$P_1 = P\{\Pi_n(m) \le u_{nm}\} - P\{\Pi_I \le u_{nm}\} \text{ and } P_2 = P\{\Pi_I \le u_{nm}\} - \textstyle\prod_{j=1}^m P(\Pi_{I_j} \le u_{nm}).$$

For the first step of the proof, we will show that upper bounds on both $|P_1|$ and $|P_2|$ are $o(1)$.

(3.9) $$|P_1| \le P(Y_i > u_{nm} \quad \text{for at least one} \quad i \in I^*) \le \textstyle\sum_{i\in I^*} P(Y_i > u_{nm}).$$

We can use Theorems A and B to estimate the probability in the right-hand side of (3.9) if $L_k \uparrow \infty$. The smallest value of i for $i \in I^*$ is $K_n - [K_n^{1/2}]$. The choice of L_i gives $L_i \geq \Delta_{i-1}^{\gamma_1} > (i-1)^{\gamma_1}$ for some $\gamma_1 > 0$. Thus Theorems A and B together give

$$(3.10) \qquad P(Y_i > u_{nm}) \leq 1 - \Phi(u_{nm}) + K \max\{K_n^{-(1+\delta)};\ L_i^{-c/2} u_{nm}^{-2-c} (\ln u_{nm})^{2+2c}\}$$

for some constants $\delta > 0$ and $K > 0$. The second term in the right-hand side of (3.10) is maximum of the error terms in Theorems A and B. We need Theorem B in case L_i is so large that $u_{nm}^2 \leq (c+1)\ln L_i$. We notice that if t is too small, Theorem B can give significantly bigger error term than that of Theorem A. However, in our case u_{nm} is large enough that it does not make much difference.

We choose γ_1 such that $\gamma_1 c/2 > 1$ and replace L_i by $|i-1|^{\gamma_1}$ in the right-hand side of (3.10). The total number of i in I^* is $m[K_n^{1/2}]$ and $i \geq K_n - [K_n^{1/2}]$. Substituting the value of u_{nm} and taking the sum over i in (3.10) we get that

$$(3.11) \qquad |P_1| = o(K_n^{-\delta})$$

for some $0 < \delta < \frac{1}{2}$. To find an upper bound for $|P_2|$, we use the ϕ-mixing property of $\{X_j\}$. Each I_j is separated by a distance of at least $[K_n^{1/2}]$ from I_l for $l \neq j$. Applying (E) m times we get that

$$(3.12) \qquad |P_2| \leq m\phi_{[K_n^{1/2}]} = o(1).$$

Thus for any integer $m \geq 1$,

$$(3.13) \qquad P(\Pi_n(m) \leq u_{nm}) = \prod_{j=1}^{m} P(\Pi_{I_j} \leq u_{nm}) + o(1).$$

Now, for any $z > 0$,

$$o(1) + \sum_{j=1}^{[z]+1} P(\Pi_{I_j} \leq u_{nz}) = P(\Pi_n([z]+1) \leq u_{nz})$$

$$(3.14) \qquad \leq P(\Pi_n(z) \leq u_{nz})$$

$$\leq P(\Pi_n([z]) \leq u_{nz}) = \prod_{j=1}^{[z]} P(\Pi_{I_j} \leq u_{nz}) + o(1).$$

We proceed now to find upper and lower bounds for $P(\Pi_{I_j} \leq u_{nz})$. For a lower bound, Boole's inequality gives $P(\Pi_{I_j} \leq u_{nz}) \geq 1 - \sum_{i \in I_j} P(Y_i > u_{nz})$ for all $j = 1, 2, \cdots, [z]+1$. Let us look at the case when $j = 1$. Let n_1 be some integer such that $n_1 = o(\ln K_n)^{1/2}$. For $1 \leq i \leq n_1$, we use Chebychev's inequality to get

$$P(Y_i > u_{nz}) \leq EY_i^2/u_{nz}^2 = u_{nz}^{-2} = O((\ln zK_n)^{-1}).$$

Thus $\sum_{i=1}^{n_1} P(Y_i > u_{nz}) = o(1)$. For $i > n_1$, we use Theorems A and B as in (3.10). Using the same kinds of arguments as after (3.10), we get

$$\sum_{i=n_1}^{K_n - [K_n^{1/2}]} P(Y_i > u_{nz}) \leq K_n \frac{\phi(u_{nz})}{u_{nz}} + K\{u_{nz}^{-2-c}(\ln u_{nz})^{2+2c} \sum_{i=n_1}^{K_n} |i-1|^{-\gamma_1 c/2} + K_n^{-\delta}\}$$

for some $\delta > 0$. We use the traditional notation $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$. (There does not seem to be any danger of confusion between the function ϕ_n in (E) and the standard normal density $\phi(x)$.) Substituting value for u_{nz} and taking limits we observe that

$$(3.15) \qquad lt_{n\to\infty} \sum_{i \in I_j} P(Y_i > u_{nz}) \leq \tau/z$$

where $\tau = \exp(-x)$ and $j = 1$. For $j = 2, 3, \cdots, [z]+1$, we can use Theorems A and B directly since $i \geq K_n$. Thus (3.15) holds for all $1 \leq j \leq [z]+1$. Substituting in (3.14) we get

$$(3.16) \qquad \liminf_{n\to\infty} P(\Pi_n(z) \leq u_{nz}) \geq \liminf \prod_{j=1}^{[z]+1} P(\Pi_{I_j} \leq u_{nz}) \geq (1 - \tau/z)^{[z]+1}.$$

Next, we will find an upper bound for $P(\Pi_{I_j} \le u_{nz})$. To do this we first need to "separate" Y_ι, $i \in I_j$. If $i \in \cup_{j=2}^m I_j$ then $i > K_n$, hence $i > [K_n^\delta]$, $0 < \delta < 1$. From the first block I_1, we exclude $Y_1, Y_2, \cdots, Y_{n_0}$; $n_0 = [K_n^\delta]$ from consideration, since doing this will only increase the probability $P(\Pi_{I_1} \le u_{nz})$. Thus we assume without loss of generality that $i \ge [K_n^\delta]$ for some $\delta > 0$. This ensures that any block lengths considered after this exclusion will be sufficiently large, viz., $L_\iota \ge \Delta_{\iota-1}^{\gamma_1} > K_n^{\gamma_1\delta}$. Let us fix $j = 1$ and notice that here $I_1 = \{n_0, n_0 + 1, \cdots, K_n - [K_n^{1/2}]\}$. The computations for $j = 2, 3, \cdots, [z] + 1$ are very similar. For z fixed and n large enough, define

$$V_\iota' = \mathrm{Var}(\textstyle\sum_{l=\Delta_{\iota-1}+1}^{\Delta_\iota-[z]} X_l) \qquad \text{and} \qquad Y_\iota' = \textstyle\sum_{l=\Delta_{\iota-1}+1}^{\Delta_\iota-[z]} X_l/(V_\iota')^{1/2}.$$

Also let

$$Y_\iota'' = \textstyle\sum_{l=\Delta_\iota-[z]+1}^{\Delta_\iota} X_l.$$

Then we can write

$$\Pi_{I_1} = \max_{\iota\in I_1} \{(V_\iota'/V_\iota)^{1/2} Y_\iota' + V_\iota^{-1/2} Y_\iota''\}. \tag{3.17}$$

Thus

$$P(\Pi_{I_1} \le u_{nz}) \le P\{\max_{\iota\in I_1} (V_\iota'/V_\iota)^{1/2} Y_\iota' \le u_{nz} + O((\ln K_n)^{-1+\theta})\}$$
$$+ P\{\max_{\iota\in I_1} (Y_\iota''/V_\iota^{1/2}) > (\ln K_n)^{-1}\}$$

for some $0 < \theta < ½$. Now, because $\sigma_j^2 \le A_2$ for all j, $\mathrm{Var}(Y_j) \le [z]^2 A_2$ and

$$|V_\iota - V_\iota'| \le [z] A_2 \{1 + \textstyle\sum_{l=1}^{L_\iota} \phi_l\}$$

using (E) and Lemma 1, page 170 of Billingsley [4]. Because of (3.2), $\sum_{l=1}^{L_\iota} \phi_l \le$ (constant) and because of (3.3b) $V_\iota \ge$ (constant) L_ι where both the constants are independent of i. Writing $(V_\iota'/V_\iota)^{1/2} = 1 + O(K_n^{-\delta\gamma_1})$ we get

$$P\{\max_{i\in I_1} (V_\iota'/V_\iota)^{1/2} Y_i' \le u_{nz} + O((\ln K_n)^{-1+\theta})\} = P\{\max_{\iota\in I_1} Y_\iota' \le u_{nz}'\} \tag{3.18}$$

where $u_{nz}' = u_{nz} + o((\ln K_n)^{-1+\theta})$. Since Y_ι'' is a sum of only $[z]$ number of X_l, each one of which has finite absolute moments up to order $2 + c$, we must have $E|Y_i''|^{2+c} < \infty$ for all i. Using Chebychev's inequality,

$$P\{\max_{\iota\in I_1} (Y_\iota''/V_\iota^{1/2}) > (\ln K_n)^{-1}\} \le \textstyle\sum_{\iota\in I_1} P\left(Y_\iota'' > \dfrac{V_\iota^{1/2}}{\ln K_n}\right) \le (\ln K_n)^{2+c} \textstyle\sum_{\iota\in I_1} V_\iota^{-(2+c)/2}.$$

Substituting i^{γ_1} for L_ι and taking the sum we see that the right-hand side above is $o(K_n^{-\delta})$ for some $\delta > 0$ and

$$P(\Pi_{I_1} \le u_{nz}) \le P(\max_{\iota\in I_1} Y_\iota' \le u_{nz}') + o(K_n^{-\delta}). \tag{3.19}$$

To finish the proof of (3.16), we need to show that $\limsup_{n\to\infty} P(\max_{i\in I_\iota} Y_\iota' \le u_{nz}') \le 1 - \tau/z +$ terms of smaller order than $1/z$). By Boole's inequality,

$$P(\max_{\iota\in I_1} Y_i' \le u_{nz}') \le 1 - \textstyle\sum_{\iota\in I_1} P(Y_\iota' > u_{nz}') + 2 \sum \sum_{i,l\in I_1;\iota\ne l} P(Y_i' > u_{nz}';\ Y_l' \ge u_{nz}'). \tag{3.20}$$

The second term in the right-hand side of (3.20) is computed in very similar manner to that of (3.15) which gives $\sum_{\iota\in I_1} P(Y_\iota' > u_{nz}') = (\tau/z)(1 + o(1))$. Events in the third term of the right-hand side of (3.20) are separated by a distance at least equal to $(l - i)[z]$. Using (E) we get an upper bound for this last sum in (3.20) to be

$$2 \textstyle\sum \sum_{i,l\in I_1;\iota\ne l} P(Y_\iota' > u_{nz}') P(Y_l' > u_{nz}')$$
$$+ 2 \textstyle\sum_{i\in I_1} P(Y_\iota' > u_{nz}') \sum_{l>\iota;l\in I_1} \phi_{(l-\iota)[z]}$$
$$\le 2(\textstyle\sum_{\iota\in I_1} P(Y_\iota' > u_{nz}'))^2 + 2g(z)(\tau/z)(1 + o(1))$$

where $g(z) = \sum_{l=1}^{\infty} \phi_{l[z]}$ is a function of z. Because of (E), it is easy to see that $lt_{z\to\infty} g(z) = 0$. Substituting in (3.14) we get

$$\limsup_{n\to\infty} P(\textstyle\prod_n (z) \le u_{nz}) \le (1 - \tau/z + \tau^2/z^2 + 2g(z)\tau/z)^{[z]}. \tag{3.21}$$

Combining (3.16), (3.21) and taking limit as $z \to \infty$ we get

$$lt_{z\to\infty}\, lt_{n\to\infty}\, P(\textstyle\prod_n (z) \le u_{nz}) = \exp(-e^{-x}) \tag{3.22}$$

for $-\infty < x < \infty$. Simple arithmetic then leads to (3.5). This concludes the proof of Theorem 3.1.

We now turn our attention to estimating the variances V_i which are unknown in practice. The averages Y_i in Theorem 3.1 were based on these variances V_i. In Theorem 3.2 below we compute new averages that are based entirely on the sample.

A reasonable estimate for V_i seems to be

$$Q_i = \textstyle\sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j^2,$$

especially if the dependence in $\{X_i\}$ is weak enough. In Theorem 3.2 we slightly strengthen the ϕ-mixing condition of Theorem 3.1 (thus making the dependence of $\{X_i\}$ even weaker).

As explained in the introduction, the new averages can be highly sensitive to small values of Q_i. In any case, while looking at the maximum of these averages we want to make sure that we don't get stuck with large values of the maximum that could be generated by small values of Q_i. Conceivably this could be done in many different ways. Below, we do this by truncating Q_i away from zero. Let $Q_i^\theta = \max\{\theta, Q_i\}$ for some $\theta > 0$ and

$$Y_i^S = \textstyle\sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j/(Q_i^\theta)^{1/2} = Y_i(V_i/Q_i^\theta)^{1/2} \qquad \text{for} \quad i = 1, 2, \cdots, K_n.$$

We define $\Pi_n^S = \max_{1\le i\le K_n} Y_i^S$.

THEOREM 3.2. *We assume that* (3.1)–(3.4) *hold. In addition, if*

$$V_i - EQ_i = o(L_i/\ln L_i) \tag{3.23}$$

and

$$\inf_i(\operatorname{Var} Q_i/L_i) \ge c_1 \tag{3.24}$$

for some constant $c_1 > 0$. *Then*

$$lt_{n\to\infty} P(\Pi_n^S \le u_n) = \exp(-e^{-x}) \tag{3.25}$$

for $-\infty < x < \infty$.

REMARK. Conditions (3.23) and (3.24) are on the correlations of (X_j, X_k) and (X_j^2, X_k^2). Using (E) and Lemma 1, page 170, of [4], it can easily be shown that the left-hand side of (3.23) is $O(L_i)$. Property (E) is related to the rate of decay of dependence between X_j and X_k. In (3.23) we assume that there is enough cancellation due to positive and negative correlations so that the left-hand side of (3.23) grows slower than $L_i/\ln L_i$. For (3.24) it is sufficient to assume that $\liminf_j \operatorname{Var}(X_j^2) > 0$ and

$$\textstyle\sum\sum_{j,k\in[\Delta_{i-1}+1,\Delta_i], j\ne k} E(X_j^2 - \sigma_j^2)(X_k^2 - \sigma_k^2) = o(L_i).$$

PROOF OF THEOREM 3.2. The use of Q_i^θ instead of Q_i is simply to avoid embarrassing situations when Q_i may be too small. We first show in Lemma 3.1 below that this does not occur if L_i is assumed to increase with n.

LEMMA 3.1. *If there exists a function* $f(n) \uparrow \infty$ *such that* $L_i/f(n) \uparrow \infty$ *for all* $1 \le i \le K_n$, *then*

$$\textstyle\sum_{i=1}^{K_n} P(Q_i \ne Q_i^\theta) = o(1). \tag{3.26}$$

PROOF OF LEMMA 3.1. We know that

$$P(Q_i \neq Q_i^\theta) = P(Q_i < \theta)$$

$$= P\left(\frac{Q_i - EQ_i}{(\mathrm{Var}\, Q_i)^{1/2}} < \frac{\theta - EQ_i}{(\mathrm{Var}\, Q_i)^{1/2}}\right).$$

We see that $(\mathrm{Var}\, Q_i)^{-1/2}(\theta - EQ_i) \leq -d_1 L_i^{1/2}$ for some constant $d_1 > 0$ because of (3.23) and (3.24). Applying Theorems A and B we have

$$(3.27) \quad \sum_{i=1}^{K_n} P(Q_i \neq Q_i^\theta) \leq \sum_{i=1}^{K_n} \{1 - \Phi(d_1 L_i^{1/2}) + K \max(L_i^{-(1+c)}(\ln L_i)^{2+2c};\, L_i^{-\lambda}\phi(d_1 L_i^{1/2}))\}$$

for some constants $K > 0$ and $\lambda > 0$. We will split the sum in the right-hand side of (3.27) into two parts, $1 \leq i \leq f(n)$ and $f(n) < i \leq K_n$. For the first part we write

$$L_i^{-(1+c)}(\ln L_i)^{2+2c} \leq (\max_{1\leq x<\infty} (x^{1/4}\ln x))^{2+2c} L_i^{-(1+c)/2} < d_2 L_i^{-(1+c)/2}$$

for some constant $d_2 > 0$. Thus the first part of the sum in the right-hand side of (3.27) is at most

$$\sum_{i=1}^{f(n)} \left\{\frac{\phi(d_1 f(n))}{d_1 f(n)} + K \max(d_2(f(n))^{-(1+c)/2};\, (f(n))^{-\lambda}\phi(d_1 f(n)^{1/2}))\right\} = o(1).$$

For $i > f(n)$, we will substitute i^{γ_1} for L_i. The second part of the sum under consideration is at most

$$(\mathrm{const}) \sum_{i=f(n)}^{K_n} \{i^{-\gamma_1/2}\exp(-d_1^2 i^{\gamma_1}/2) + (\ln i)^{2+2c} i^{-\gamma_1(1+c)}\}.$$

Because $\gamma_1 c/2 > 1$, the above is $o(1)$. This proves Lemma 3.1.

The proof of Theorem 3.2 is slightly tricky. We write

$$Y_i^S = Y_i + Y_i\left\{\frac{V_i}{Q_i} \cdot \frac{Q_i}{Q_i^\theta} - 1\right\}.$$

In view of Theorem 3.1 it is sufficient to show that the maximum of the second term above will converge to zero in probability. First let us assume that there exists $f(n)$ such that $L_i/f(n) \uparrow \infty$. Then Lemma 3.1 above would let us substitute 1 for Q_i/Q_i^θ. Remaining proof would follow from the basic fact that $(Q_i/V_i - 1)$ goes to zero sufficiently fast (shown at 3.32). But for this we need to assume that $i > n_0 = [K_n^\delta]$ for some $\delta > 0$. Thus, as a first step of the proof of Theorem 3.2, we exclude $Y_1^S, \cdots, Y_{n_0}^S$ from consideration in (3.25). That is, we want to show that the right-hand side in the following inequality is $o(1)$.

$$(3.28) \quad P(\max_{n_0<i\leq K_n} Y_i^S \leq u_n) - P(\Pi_n^S \leq u_n) \leq \sum_{i=1}^{n_0} P(Y_i^S > u_n).$$

In order to find the upper bound on $P(Y_i^S > u_n)$, we write

$$Y_i^S = Y_i\left(\frac{V_i}{Q_i} \cdot \frac{Q_i}{Q_v^\theta}\right)^{1/2} \simeq Y_i(V_i/Q_i)^{1/2}$$

in view of Lemma 3.1. At (3.31) below we prove a weaker version of (3.32) and this would allow us to substitute Y_i for Y_i^S, Y_i being asymptotically normal. Use of Theorems A and B would then allow exclusion of $Y_1^S, \cdots, Y_{n_0}^S$. (Notice that all through this we have assumed that $L_i/f(n) \uparrow \infty$.)

There may exist some $f(n)$ as above simply by choice of L_i. But because of the wide choice made available in the beginning of the section, L_i may be fixed for fixed values of i. The only way left to make sure that $L_i \uparrow \infty$ (with n) is to exclude some Y_i^S for small values of i from consideration. This argument may seem circular, since we noticed in the last paragraph that L_i was assumed to increase with n for exclusion of $Y_1^S, \cdots, Y_{n_0}^S$. But note that there we needed to exclude a specified number n_0 of Y_i^S. This is possible only by using more accurate information about the probabilistic structure of Y_i^S as we do in arguments leading to (3.32). A smaller number $n_1 = o((\ln K_n)^{1/2})$ of Y_i^S can be excluded by using cruder bounds given by, say, Chebychev's inequality and substituting the smallest possible value of Q_i^θ in Y_i^S viz., θ, which is what we do below. For the final twist of the argument, notice that Chebychev's inequality would give reasonable bounds on the new average $\sum_{j\in L_i} X_j/\theta$ if L_i is not too large. Thus we define $D = (1 \le i \le n_1 \mid L_i \le u_n)$ where n_1 is some integer such that $n_1 = o((\ln K_n)^{1/2})$. In case D is not empty, we exclude Y_i^S for $i \in D$ by using Chebychev's inequality, and for remaining i, (i.e., $i > n_1$) we have $L_i/n_1^{\gamma_1} \uparrow \infty$ (with n). If D is empty, then $L_i > u_n$; $1 \le i \le n_1$, and hence for all i, $1 \le i \le K_n$, since L_i is nondecreasing in i. Thus for $i \notin D$, there exists $f(n) = \min(n_1^{\gamma_1}, u_n)$ for which $L_i/f(n) \uparrow \infty$ (with n).

For $i \in D$, let us write $Y_i(V_i/\theta)^{1/2}$ instead of Y_i^S, since this would only increase the probability $P(Y_i > u_n)$. Also $V_i \sim (\text{const})L_i \lesssim (\text{const})u_n$ in view of (3.3a), (3.3b), (3.23) and the requirements of set D. Now

$$\sum_{i\in D} P(Y_i^S > u_n) \le \sum_{i\in D} P(Y_i > d_3 u_n^{1/2}) \le \sum_{i\in D} d_3^{-2} u_n^{-1} \tag{3.29}$$

for some constant $d_3 > 0$. The total number of i in D is at most $o((\ln K_n)^{1/2})$ and the right-hand side of (3.29) is $o(1)$. Thus in the remaining part of this section we can assume without loss of generality that there exists some $f(n) \uparrow \infty$ such that $L_i/f(n) \uparrow \infty$ (with n) for all i.

We now turn our attention to showing that the right-hand side of (3.28) is $o(1)$. In view of (3.26) we write $Y_i^S \simeq Y_i(V_i/Q_i)^{1/2}$ for $i \in [1, 2, \cdots, n_0]$. For any $\epsilon > 0$,

$$P(Y_i^S > u_n) \le P(Y_i > (1-\epsilon)u_n) + P(|(V_i/Q_i)^{1/2} - 1| > \epsilon). \tag{3.30}$$

Using Theorems A and B as in (3.10), we see that the sum for $i \in [1, \cdots, n_0]$ of the first term in the right-hand side of (3.30) is $o(1)$ for sufficiently small values of $\delta > 0$ (i.e., of $n_0 = [K_n^\delta]$). We will now show that

$$\sum_{i=1}^{n_0} P(Q_i > (1+2\epsilon)V_i) = o(1). \tag{3.31}$$

Note that $(\text{Var } Q_i)^{-1/2}(V_i - EQ_i) = O(L_i^{1/2}/\ln L_i)$ due to (3.23) and (3.24). Thus $P(Q_i > (1 + 2\epsilon)V_i) \simeq P((\text{Var } Q_i)^{-1/2}(Q_i - EQ_i) > (\text{const})\epsilon L_i^{1/2})$. Using the same estimates as in (3.27), (3.31) follows easily. The argument for $\sum_{i=1}^{n_0} P(Q_i < (1-2\epsilon)V_i) = o(1)$ are very similar. A few arithmetic calculations show that this is sufficient to conclude that the right-hand side of (3.28) is $o(1)$, i.e., we assume $Y_1^S, \cdots, Y_{n_0}^S$ as excluded from consideration.

To complete the proof of Theorem 3.2 we are now in a position to prove a stronger version of (3.31), since $i > [K_n^\delta]$. We write $Y_i^S \simeq Y_i + Y_i\{(V_i/Q_i)^{1/2} - 1\}$ (having substituted Q_i for Q_i^θ in view of 3.26). Theorem 3.1 implies that it is sufficient to show that $(\ln K_n) \max_{n_0 \le i \le K_n} \{(V_i/Q_i)^{1/2} - 1\} \to_p 0$ as $n \to \infty$. For this we will show that

$$lt_{n\to\infty} \sum_{i=n_0}^{K_n} P(Q_i/V_i > 1 + \epsilon/\ln K_n) = 0. \tag{3.32}$$

The corresponding result for $P(Q_i/V_i < 1 - \epsilon/\ln K_n)$ will then follow easily. Now

$$P(Q_i/V_i > 1 + \epsilon/\ln K_n) = P\left(\frac{Q_i - EQ_i}{(\text{Var } Q_i)^{1/2}} > \frac{V_i - EQ_i}{(\text{Var } Q_i)^{1/2}} + \frac{\epsilon V_i}{(\ln K_n)(\text{Var } Q_i)^{1/2}}\right). \tag{3.33}$$

Because $i > [K_n^\delta]$, $\ln L_i > \gamma_1 \delta \ln K_n$ and (3.23) implies that $(V_i - EQ_i)(\text{Var } Q_i)^{-1/2} = o(L_i^{1/2}(\ln K_n)^{-1})$. Thus the right-hand side of (3.33) is at most $P\{(\text{Var } Q_i)^{-1/2}(Q_i - EQ_i) > d_4 L_i^{1/2}(\ln K_n)^{-1}\}$ for some constant $d_4 > 0$. The remaining steps are by now routine applications of Theorems A and B. The sum in (3.33) is $o(1)$ and this completes the proof of Theorem 3.2.

4. Joint distribution of $\mathbf{M}_n$ and $\mathbf{\Pi}_n$. In this section we will assume that $\{X_j; j \ge 0\}$ is a stationary Gaussian sequence with mean zero, variance one and $r_j = EX_0X_j$. Let $M_n = \max_{0\le j\le n} X_j$. The limiting distribution of M_n is extensively studied under various conditions on $\{r_j\}$ (see, e.g., [2] and [7]). The smoothed sample $\{Y_i; i \ge 0\}$ in this case will again be a Gaussian process (nonstationary) whose convariance function can be computed from $\{r_j\}$.

Throughout this section, we assume that $r_k \ln k \to 0$ as $k \to \infty$. This condition is sufficient and practically necessary for the convergence of M_n as is apparent from looking at results of

[2] and [7]. Conditions for convergence of Π_n however, do not depend directly on $\{r_k\}$, but on the size and the smoothness of the variances V_i of the blocks. In Mittal ([6], Theorem 2.1), asymptotic behavior of the maximum of normalized sums is considered. Conditions (4.1)–(4.3) below are those of the "moderately dependent" case there. See [6] for some examples of covariance functions satisfying (4.1)–(4.3). The joint distribution of M_n and Π_n in the "strongly dependent" case (see Theorem 3.1 of [6]) is conjectured to be degenerate on the diagonal. The normalizing constants in this case will involve the covariance function. We choose to omit the calculations for the "strongly dependent" case.

The sequence $\{L_i\}$ described in the beginning of Section 3 will have now even a greater freedom of choice. To keep the choice sensible, we will assume that $L_i < n^{\gamma_2}$ for some $0 < \gamma_2 < 1$ and $L_i/L_{i+1}^{1-\gamma} \to \infty$ as $i \to \infty$ for all $0 < \gamma < 1$. In addition, assume that $\ln n/\ln K_n$ is bounded. (This allows the choice $L_i = (\ln n)^k$, $k \ge 1$ for all $1 \le i \le K_n$ which was not possible in Section 3. As before, K_n is such that $\sum_{i=1}^{K_n-1} L_i < n \le \sum_{i=1}^{K_n} L_i$.)

THEOREM 4.1. *Let $r_k \ln k \to 0$ as $k \to \infty$ and for sufficiently large k,*

$$\sum_{i=-k}^{k} r_i = k^{\delta}G(k) > 0, \tag{4.1}$$

$-1 < \delta < 1$ and $G(k)$ is a slowly varying function satisfying the following conditions.

(4.2) *For all sufficiently large l, $G(k)/G(l)$ is bounded away from 0 and ∞ if $l \ge k \ge l^{\theta}$, $1 \ge \theta > 0$ and*

(4.3) *If $l \ge k$, $k \to \infty$ such that $(l-k)/k^{1-\gamma} \to \infty$ for all $\gamma > 0$ and $\dfrac{l-k}{k} \to c$ for $0 \le c \le \infty$, then $lt_{k\to\infty} \dfrac{l}{k}(G(l)/G(l-k) - 1) = 0$.*

Then

$$lt_{n\to\infty} P(M_n \le u_n(x); \Pi_n \le u_{K_n}(y)) = \exp\{-e^{-x} - e^{-y}\} \tag{4.4}$$

for all $-\infty < x, y < \infty$ where $u_m(x) = b_m + x/c_m$ and b_m and c_m are defined in (2.2).

PROOF. The proof consists of showing that M_n and Π_n are asymptotically independent and that Π_n properly normalized converges to the double exponential distribution ($\exp(-e^{-y})$). Notice that conditions (4.1)–(4.3) are intuitively much weaker than the ϕ-mixing assumed in Section 3. This is to be expected since here we are dealing with Gaussian rather than general sequences.

We will first find the limiting behavior of Π_n. Without loss of generality, we can replace Π_n in (4.4) by $\max_{K_n^{\gamma} \le i \le K_n} Y_i$ for any $0 < \gamma < 1$ since $\sum_{i=1}^{K_n} P(Y_i > u_{K_n}(y)) \le K^{\gamma_n} \phi(u_{K_n}(y)/u_{K_n}(y)) = o(1)$. Thus from now on let us assume $\Pi_n = \max_{K_n^{\gamma} \le i \le K_n} Y_i$. The result follows from Theorem 3.1 of Berman [2] if we show that $|EY_iY_j| < 1$ (see 4.9 below) and that $|EY_iY_j| \ln |j - i| \to 0$ as $j - i \to \infty$ (shown at (4.11) and (4.12)).

Define $D_{ij} = \sum_{k=i}^{j-1} L_k$ and $B_i = \{\Delta_{i-1} + 1, \Delta_{i-1} + 2, \cdots, \Delta_i\}$ where Δ_i is the same as in Section 3. Then

$$\begin{aligned} EY_iY_j &= \sum_{l\in B_i} \sum_{k\in B_j} E(X_lX_k)/(V_iV_j)^{1/2} \\ &= \sum_{l=0}^{L_i-1} \sum_{k=D_{ij}}^{D_{ij}+L_j-1} r_{k-l}/(V_iV_j)^{1/2}. \end{aligned} \tag{4.5}$$

Using condition (4.1), we get the numerator on the right-hand side of (4.5) to be equal to

$$\sum_{l=0}^{L_i-1} \{(D_{ij} + L_j - l)^{\delta}G(D_{ij} + L_j - l) - (D_{ij} - l)^{\delta}G(D_{ij} - l)\}. \tag{4.6}$$

First, let us assume $j = i + 1$. Then we can rewrite (4.6) as

$$\sum_{l=0}^{L_i-1} (L_i + L_j - l)^{\delta}G(L_i + L_j - l) - V_i. \tag{4.7}$$

Here we have $L_j < L_i + L_j - l \le L_i + L_j$. Now $(L_i + L_j - l)/(L_j - l)^{1-\theta} \ge L_i^{\theta} \to \infty$ for all $\theta > 0$. Thus condition (4.3) implies that $G(L_i + L_j - l) \le (l + \epsilon)G(L_i)$ for large n and $0 \le l \le$

$L_i - l$. The variance $V_i \sim (\delta + 1)^{-1} L_i^{\delta+1} G(L_i)$. (See (2.5) in [6].) The following upper bound for (4.7) ignores the second order terms. For every $\epsilon > 0$, we can find n large enough that (4.7) is at most

$$G(L_i)(1 + \delta)^{-1}\{(L_i + L_j)^{\delta+1} - L_i^{\delta+1} - L_j^{\delta+1} + \epsilon((L_i + L_j)^{\delta+1} - L_j^{\delta+1})\}. \tag{4.8}$$

We recall the algebraic inequality $(a + b)^\theta - a^\theta - b^\theta < a^{\theta/2} b^{\theta/2}$ for all $a, b > 0$ and $0 < \theta < 2$. Also notice that $G(L_i)/G(L_j) \to 1$ as $n \to \infty$ because of (4.3) and the choice of $L_{i+1} = L_j$. Substituting in (4.5), we get the following strict upper bound for it.

$$(1 + \epsilon')(\rho + 2\epsilon) \tag{4.9}$$

where $\rho < 1$. We can choose ϵ and ϵ' small enough for n large so that (4.9) is strictly less than one.

Next let us assume $j \geq i + 2$. The (4.6) is equal to

$$\begin{aligned} &\textstyle\sum_{l=0}^{L_i-1} \{(D_{ij} + L_j - l)^\delta (G(D_{ij} + L_j - l) \\ &\qquad - G(D_{ij} - l)) + G(D_{ij} - l)((D_{ij} + L_j - l)^\delta (D_{ij} - l)^\delta)\}. \end{aligned} \tag{4.10}$$

Now, consider the first term in the sum of (4.10). We know that $(D_{ij} - l)/L_j^{1-\theta} \geq L_{j-1}/L_j^{1-\theta} \to \infty$ for all $\theta > 0$. Hence applying (4.3) we get

$$(D_{ij} + L_j - l)^\delta (G(D_{ij} + L_j - l) - G(D_{ij} - l) \leq \epsilon(D_{ij} - l)L_j(D_{ij} + L_j - l)^{-1+\delta}$$

for $\epsilon > 0$ and n large. Also

$$\begin{aligned} &(D_{ij} + L_j - l)^\delta - (D_{ij} - l)^\delta \\ &\qquad = (D_{ij} + L_j - l)/(D_{ij} + L_j - l)^{1-\delta} - (D_{ij} - l)^\delta < L_j(D_{ij} + L_j - l)^{-1+\delta}. \end{aligned}$$

Thus (4.10) is less than

$$(1 + \epsilon)L_j \textstyle\sum_{l=0}^{L_i-1} (D_{ij} + L_j - l)^{-1+\delta} G(D_{ij} - l).$$

Suppose that for all $0 < \theta < 1$, $L_i/(D_{ij} - 2L_i + 1)^{1-\theta}$ and $L_j/(D_{ij} - L_i - L_j + 1)^{1-\theta}$ both tend to infinity with n, then using (4.3) we know that $G(D_{ij} - l)/G(L_i)$ and $G(D_{ij} - l)/G(L_j)$ both tend to one with $n \to \infty$. Substituting in (4.5) we get

$$|EY_iY_j| \leq (1 + \epsilon')(L_iL_j)^{(1-\delta)/2}/((j - i - 1)L_i + L_j)^{1-\delta} < (1 + \epsilon')(j - i)^{-(1-\delta)/2}. \tag{4.11}$$

Now, if there exists $0 < \theta < 1$ such that $L_i/(D_{ij} - 2L_i + 1)^{1-\theta}$ is bounded as $n \to \infty$, then

$$\begin{aligned} |EY_iY_j| &\leq (\text{const}) \frac{\max_{0 \leq l \leq L_i - 1} G(D_{ij} - l)}{G^{1/2}(L_i)G^{1/2}(L_j)} \left(\frac{L_i}{(D_{ij} + L_j - L_i)^{(1-\theta)}}\right)^{(1-\delta)/2} \\ &\quad \cdot \left(\frac{L_j}{D_{ij} + L_j - L_i}\right)^{(1-\delta)/2} (D_{ij} + L_j - L_i)^{-\theta(1-\delta)/2} \\ &< \epsilon(j - i)^{\theta(1-\delta)/4} \end{aligned} \tag{4.12}$$

for large n. Similarly for the case when $L_j/(D_{ij} - L_i - L_j + 1)^{1-\theta}$ is bounded. Thus (4.11) and (4.12) imply that $|EY_iY_j| \ln|j - i| \to 0$ as $n \to \infty$ for all $i, j > K_n^\gamma$. This together with (4.9) and Theorem 3.1 of [2] implies that $lt_{n\to\infty} P(\Pi_n \leq u_{K_n}(y)) = \exp(-e^{-y})$ for $-\infty < y < \infty$.

To show the asymptotic independence of M_n and Π_n, notice that $Z^n = (Z_1^n, \cdots, Z_{n+K_n}^n)^T = (X_1, \cdots, X_n, Y_1, \cdots, Y_{K_n})^T$ has a joint normal distribution for each n. Let $\sum(n)$ denote the covariance matrix of Z^n. That is, the (i, j)th element of $\sum(n)$ is $\sigma_{ij}(n) = EZ_i^nZ_j^n$. Let $\sigma_{ij}^0(n) = EZ_i^nZ_j^n$, if both i and $j \leq n$ or both i and j are $> n$. Take $\sigma_{ij}^0(n) = 0$ otherwise. That is, $\sum^0(n) = ((\sigma_{ij}^0(n)))$ is the covariance matrix of Z^n where X_l and Y_m are assumed independent for each l and m. Slight modification of (4.12) in Berman [3] (as used in [6]) gives that

$$|P(M_n \leq u_n(x); \Pi_n \leq u_{K_n}(y)) - P(M_n \leq u_n(x))P(\Pi_n \leq u_{K_n}(y))|$$

$$(4.13)\qquad \leq (\text{const}) \textstyle\sum_{i=1}^{n+K_n} \sum_{j=1}^{n+K_n} | \sigma_{ij}(n) - \sigma_{ij}^0(n) | \cdot \exp\{-(u_n^2(x) - 2\sigma_{ij}(n)u_n(x)u_{K_n}(y) + u_{K_n}^2(y))/2(1 - \sigma_{ij}^2(n))\}.$$

We only need to compute $\sigma_{ij}(n) = \sigma_{ji}(n)$ for $i \leq n; j > n$, the sum in (4.13) being zero over the remaining terms. Now

$$(4.14)\qquad EX_iY_j = \textstyle\sum_{l \in B_{j-n}} EX_iX_l / V_{j-n}^{1/2}.$$

If $X_i \in B_{jn}$ then

$$(4.15)\qquad EX_iY_j \leq \frac{lnL_{j-n} + L_{j-n}^{\delta} G(L_{j-n})}{L_{j-n}^{(1+\delta)/2} G^{1/2}(L_{j-n})} = o(L_{j-n}^{-\epsilon})$$

for some $\epsilon > 0$ of $X_i \notin B_{j-n}$, then let A_{ij} denote the "distance" between X_j and B_{j-n}, i.e., for all l with $X_l \in B_{j-n}$, $l - i \geq A_{ij}$. Hence for $X_i \notin B_{j-n}$

$$(4.16)\qquad \begin{aligned} EX_iY_j &\leq \frac{(A_{ij} + L_{j-n})^{\delta} G(A_{ij} + L_{j-n}) - A_{ij}^{\delta} G(A_{ij})}{L_{j-n}^{(1+\delta)/2} G^{1/2}(L_{j-n})} \\ &\leq L_{j-n}^{(1-\delta)/2} (A_{ij} + L_{j-n})^{-1+\delta+\epsilon}. \end{aligned}$$

The above follows by similar computations as before. The right-hand side of (4.16) is $o(L_{j-n}^{-\epsilon})$ for some $\epsilon > 0$. However, if $A_{ij} \geq K_n^{\gamma}$ for some $\gamma > 0$, then we can write the right-hand side of (4.16) as $o(K_n^{-\gamma'})$ for some $\gamma' > 0$. Substituting in (4.13) we get an upper bound for the right-hand side there to be

$$(4.17)\qquad \begin{aligned} &n \textstyle\sum_{j=n+1}^{n+K_n^{\gamma}} L_{j-n}^{-\epsilon} \exp\left\{\frac{-u_n^2(x)}{2} - (1 - 2L_{j-n}^{-\epsilon} u_n(x)/u_n(y)) \frac{u_n^2(y)}{2}\right\} \\ &\quad + nK_n K_n^{-\gamma'} \exp\left\{\frac{-u_n^2(x)}{2} - (1 - 2K_n^{-\gamma'} u_n(x)/u_n(y)) \frac{u_n^2(y)}{2}\right\}. \end{aligned}$$

We have chosen K_n such that $u_n(x)/u_n(y)$ is bounded and $L_{j-n}^{-\epsilon} \to 0$ as $n \to \infty$. Thus the terms in (4.17) are $o(1)$ and the result follows.

Acknowledgment. I would like to thank Professor Wassily Hoeffding for many references and helpful discussions.

REFERENCES

[1] BABU, G., GHOSH, M. and SINGH, K. (1978). On rates of convergence to normality for ϕ-mixing processes. Indian Statistical Institute, Calcutta. Technical Report No. 3.

[2] BERMAN, S. M. (1964). Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35** 502–516.

[3] BERMAN, S. M. (1971). Asymptotic independence of the numbers of high and low level crossing of stationary Gaussian processes. *Ann. Math. Statist.* **42** 927–945.

[4] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[5] FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications,* Vol. II. Wiley, New York.

[6] MITTAL, Y. (1979). Extreme value distribution for normalized sums from stationary Gaussian sequences. *Stochastic Processes Appl.* **9** 67–84.

[7] MITTAL, Y. and YLVISAKER, D. (1975). Limit distributions for the maxima of stationary Gaussian processes. *Stochastic Processes Appl.* **3** 1–18.

[8] STATULEVIČIUS, V. A. (1966). On large deviations. *Z. Wahrsheinlichkeitstheorie und Verw. Gebiete* **6** 133–144.

[9] STATULEVIČIUS, V. A. (1974). Limit theorems for dependent random variables under various regularity conditions. *Proc. Int. Cong. Math.* Vancouver.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305