The Annals of Probability
1978, Vol. 6, No. 3, 421-432

MAXIMA OF PARTIAL SAMPLES IN GAUSSIAN SEQUENCES[1]

BY YASHASWINI MITTAL

Stanford University

Let $\{X_n, n \geqq 0\}$ be a Gaussian sequence with $EX_i \equiv 0$; $V(X_i) \equiv 1$ and $EX_i X_j = r_{ij}$. Define $M_n = \max_{0 \leqq i \leqq n} X_i$ and $m_{n,k} = \max(X_i; i \in G_n)$ where $G_n = (t_1, \cdots, t_{n'})$ is a subset of $(0, 1, 2, \cdots, n)$ and $n' = [n/k]$ for some integer $k \geqq 1$, $[x]$ being the integral part of x. We show that $P\{c_n(M_n - m_{n,k}) \leqq x\} \to (1 + (k-1)e^{-x})^{-1}$ as $n \to \infty$ for all $x \geqq 0$ where $c_n = (2 \log n)^{\frac{1}{2}}$, if the sequence is "moderately dependent," namely if

(1) (i) $\sup_{ij} |r_{ij}| < \delta < 1$

(ii) $|r_{ij}| \leqq \rho(i-j)$ for $|i-j| > N_0$ such that $\rho_n \log n = o(1)$.

Somewhat surprisingly the same result holds even though the sequence is "strongly dependent," namely if

(2) (i) $r_{ij} = r(i-j)$; r_n convex for $n \geqq 0$ and $r_n = o(1)$

(ii) $(r(n) \log n)^{-1}$ is monotone and $o(1)$.

1. Introduction. Let $\{X_n, n \geqq 0\}$ be a Gaussian sequence with $EX_n \equiv 0$; $V(X_n) \equiv 1$ which will be called "standard sequence" from now on. Define $M_n = \max_{0 \leqq i \leqq n} X_i$ and $m_{n,k} = \max(X_i; i \in G_n)$ where $G_n = (t(1), \cdots, t(n'))$ is a subset of $(0, 1, \cdots, n)$ and $n' = [n/k]$ for some integer $k \geqq 1$, $[x]$ being the integral part of x. We assume that

(1.1) (i) $\sup_{i,j} |EX_i X_j| < \delta < 1$

(ii) $|EX_i X_j| \leqq \rho(i-j)$ for $|i-j| > N_0$ such that $\rho_n \log n = o(1)$.

By Berman (1964, Theorem 3.1), $c_n(M_n - b_n)$ converges in distribution to X where $c_n = (2 \log n)^{\frac{1}{2}}$; $b_n = c_n - \log(4\pi \log n)/2c_n$ and $P(X \leqq x) = e^{-e^{-x}}$ for $-\infty < x < \infty$. Throughout the paper $\log n$ denotes the natural logarithm of n. It is easy to see that for large n, M_n and $m_{n,k}$ will be approximately the same in size. In Section 2 we show that $c_n(M_n - m_{n,k})$ converges in distribution to a truncated logistic distribution under condition (1.1). (The word "truncated" is used rather freely here. In fact this distribution is the same as the logistic distribution on the positive axis while the mass of the negative axis is pooled into an atom at the origin. All references to the truncated logistic distributions should be interpreted in this manner.)

In Section 3 we deal with "strongly dependent" Gaussian sequences but require

Received August 17, 1976.

[1] Research supported in part by NSF Grant No. MPS74-21416 and a grant from SIMS.

AMS 1970 *subject classifications.* Primary 6030, 6050; Secondary 6285.

Key words and phrases. Maxima, Gaussian sequences, strong and weak correlations, partial samples, logistic distribution.

them to be stationary. Instead of (1.1) we assume there

(1.2) (i) $EX_iX_j = r(i-j)$; r_n convex for $n \geqq 0$ and $r_n = o(1)$
(ii) $(r(n) \log n)^{-1}$ is monotone and $o(1)$.

(These requirements easily imply (1.1)(i).) It is shown via a rather long proof that under (1.2), $c_n(M_n - m_{n,k})$ has the same truncated logistic limit distribution. This fact seems counterintuitive at first and is much harder to prove. The beginning of Section 3 has some comments in this direction.

The problem arose from some practical considerations of collecting air pollution data. Air pollution standards are generally in terms of the maximum concentrations. Various practical difficulties (cost of sampling or missing data) are at times responsible for having only a partial sample available for making inferences. The behavior of the difference of two maxima considered here can thus be used to retract information about the original maximum via the maximum of only a partial sample.

I want to thank my colleague, Paul Switzer, for suggesting this problem.

2. Asymptotically independent sequences. The basic result is given by

THEOREM 2.1. *Let $\{X_n, n \geqq 0\}$ be a standard sequence with $r_{ij} = EX_iX_j$ and M_n and $m_{n,k}$ as defined in the beginning of the last section. If* (1.1) *holds then*

$$(2.1) \qquad P\left\{M_n - m_{n,k} \leqq \frac{x}{c_n}\, ;\, m_{n,k} \leqq b_{n,k} + \frac{z}{c_n}\right\} \to \frac{\exp\{-e^{-z}(1+(k-1)e^{-x})\}}{1+(k-1)e^{-x}}$$

as $n \to \infty$ *where* $c_n = (2 \log n)^{\frac{1}{2}}$; $b_{n,k} = (2 \log (n/k))^{\frac{1}{2}} - \log (4\pi \log n)/2c_n$ *and* $0 \leqq x < \infty$; $-\infty < z < \infty$.

REMARK. As pointed out by the referee, this theorem can be proved as follows. Let $\xi_n = \max_{i \in G_n} X_i$ and $\xi_n' = \max_{i \in G_n^c} X_i$. We are looking for the limit of the joint distribution of the properly normalized Z_n and ξ_n where

$$\begin{aligned} Z_n &= 0 && \text{if } \xi_n' \leqq \xi_n \\ &= \xi_n' - \xi_n && \text{if } \xi_n' > \xi_n \end{aligned}$$

and this is equivalent to finding the limit of the joint distribution of properly normalized ξ_n and ξ_n'. If $r_{ij} \equiv 0$ for $0 \neq j$ then ξ_n and ξ_n' are independent and each converges to double exponential distribution after proper normalization. When r_{ij} satisfies above conditions we can compare the joint distribution of ξ_n and ξ_n' to the independent case by Berman (1971, (4–5) page 932) and show that the error term tends to zero. This approach is very direct and clear. Unfortunately, however, it cannot be extended to the strongly dependent case for reasons to be discussed in the beginning of Section 3. The following approach to the proof, even though long and cumbersome, is directly extended to the strongly deqendent case.

PROOF. Let us denote $\eta_{0,n} = b_{n,k} - g(n)/c_n$ and $\eta_{z,n} = b_{n,k} + z/c_n$. We choose $g(n) \to \infty$ such that $\{\max(\rho_n \log n); \log \log n/(\log n)^{\frac{1}{2}}\}e^{4g(n)} = o(1)$. We know by

Berman (1964) that $P\{m_{n,k} \leqq \eta_{0,n}\} = o(1)$, hence the probability in (2.1) can be replaced by

$$P\left\{M_n - m_{n,k} \leqq \frac{x}{c_n}\,;\, \eta_{0,n} \leqq m_{n,k} \leqq \eta_{z,n}\right\}. \tag{2.2}$$

Define $m^i_{n,k} = \max\{X_j; j \in G_n; j \neq t(i)\}$. Then

$$\begin{aligned}\{\eta_{0,n} \leqq m_{n,k} \leqq \eta_{z,n}\} &= \textstyle\bigcup_{i=1}^{n'} \{m^i_{n,k} \leqq X_{t(i)};\, \eta_{0,n} \leqq X_{t(i)} \leqq \eta_{z,n}\} \\ &= \textstyle\bigcup_{i=1}^{n'} A_i\,, \quad \text{say.}\end{aligned}$$

Now $P\{\bigcup_{i=1}^{n'} A_i\} = \sum_{n=1}^{n'} P(A_i)$ since the intersection of any two of the above sets, e.g., $A_i \cap A_j$, is a subset of the set $\{X_{t(i)} = X_{t(j)}\}$. But this set has probability zero unless X_{t_i} and X_{t_j} have correlation 1 and for $i \neq j$ this is impossible by assumption (1.1)(ii). Thus the probability in (2.2) is equal to

$$\textstyle\sum_{i=1}^{n'} P\left\{M_n{}^D \leqq X_{t_i} + \frac{x}{c_n}\,;\, m^i_{n,k} \leqq X_{t(i)};\, \eta_{0,n} \leqq X_{t(i)} \leqq \eta_{z,n}\right\} \tag{2.3}$$

where $M_n{}^D = \max\{X_u, u \in S_D\}$; $S_D = \{u, 0 \leqq u \leqq n \text{ but } u \notin G_n\}$ is the maximum of the sample having deleted about n/k variables. Conditioning on X_{t_i}, we get the quantity in (2.3) to be equal to

$$\textstyle\sum_{i=1}^{n'} \int_{\eta_{0,n}}^{\eta_{z,n}} P\left\{M_n{}^D \leqq y + \frac{x}{c_n}\,;\, m^i_{n,k} \leqq y \,|\, X_{t(i)} = y\right\} \varphi(y)\,dy\,. \tag{2.4}$$

Note that $\varphi(y) = (2\pi)^{-\frac{1}{2}} \exp\{-y^2/2\}$ is the standard normal density. Let us define $Y_u{}^i = ((X_u \,|\, X_{t(i)} = y) - r_{u,t(i)}y))/(1 - r^2_{u,t(i)})^{\frac{1}{2}}$ $u = 0, 1, \cdots, n$, $u \neq t(i)$. We know that $Y_u{}^i$ are standard Gaussian with

$$EY_u{}^iY_v{}^i = \frac{r_{uv} - r_{u,t(i)}r_{v,t(i)}}{(1 - r^2_{u,t(i)})^{\frac{1}{2}}(1 - r^2_{v,t(i)})^{\frac{1}{2}}} = \gamma^i_{uv}\,, \quad \text{say.}$$

The probability in (2.4) is equal to

$$P\left\{Y_u{}^i \leqq y\left(\frac{1 - r_{u,t(i)}}{1 + r_{v,t(i)}}\right)^{\frac{1}{2}} + \frac{x}{c_n}\,,\, u \in S_D;\, Y_v{}^i \leqq y\left(\frac{1 - r_{v,t(i)}}{1 + r_{u,t(i)}}\right)^{\frac{1}{2}},\, v \in S_i\right\} \tag{2.5}$$

where $S_i = G_n - \{t(i)\}$. The absolute difference of (2.5) and the probability in (2.5) when the variables $Y_u{}^i$ for all u such that $|u - t(i)| < n^\alpha$ $(1 > \alpha > 0)$ are excluded from consideration is at most

$$\begin{aligned} 1 - P&\left\{Y_u{}^i \leqq y\left(\frac{1 - r_{u,t(i)}}{1 + r_{u,t(i)}}\right)^{\frac{1}{2}};\, |u - t(i)| < n^\alpha\right\} \\ &\leqq 1 - P\left\{\max\nolimits_{|u-t(i)|<n^\alpha} Y_u{}^i \leqq y\left(\frac{1-\delta}{1+\delta}\right)^{\frac{1}{2}}\right\}. \end{aligned} \tag{2.6}$$

The last inequality follows because looking at (2.4) we see that only positive values of y need to be considered and by (1.1)(ii) $(1 - r_{u,t(i)})/(1 + r_{u,t(i)}) \geqq (1 - \delta)/(1 + \delta)$. The r.h.s. of (2.6) is at most $2n^\alpha(1 - \Phi(y((1 - \delta)/(1 + \delta))^{\frac{1}{2}}))$, where $\Phi(y) = \int_{-\infty}^{y} \varphi(u)\,du$. Thus if we delete X_u for $|u - t(i)| < n^\alpha$ from consideration in (2.4) the absolute difference will be bounded above by

$$n' \cdot 2n^\alpha \int_{\eta_{0,n}}^{\eta_{z,n}} \left(1 - \Phi\left(y\left(\frac{1-\delta}{1+\delta}\right)^{\frac{1}{2}}\right)\right)\varphi(y)\,dy$$

$$\leqq (\text{const.}) n^{1+\alpha}\left(1 - \Phi\left(\eta_{0,n}\left(\frac{1-\delta}{1+\delta}\right)^{\frac{1}{2}}\right)\right)(1 - \Phi(\eta_{0,n}))$$

$$\leqq (\text{const.}) \frac{n^{1+\alpha}}{\log n} \exp\left\{-\frac{\eta_{0,n}^2}{1+\delta}\right\}$$

$$\leqq (\text{const.}) \exp\left\{\left(1 + \alpha - \frac{2}{1+\delta}\right)\log n + 2g(n)\right\}.$$

By choosing $\alpha < (1-\delta)/(1+\delta)$ and noticing that $g(n)/\log n \to 0$, we see that the r.h.s. above is $o(n^{-\beta})$ for some $\beta > 0$. Finally, for $|u - t(i)| > n^\alpha$, $|r_{u,t(i)}| \leqq \rho(n^\alpha)$ and $(1 - r_{u,t(i)})/(1 + r_{u,t(i)}) \geqq (1 - \rho(n^\alpha))/(1 + \rho(n^\alpha))$. Thus the quantity in (2.4) is almost

$$(2.7)\qquad \sum_{i=1}^{n'} \int_{\eta_{0,n}}^{\eta_{z,n}} P\left\{\max_{u\in S_D^0} Y_u^i \leqq y\left(\frac{1-\rho(n^\alpha)}{1+\rho(n^\alpha)}\right)^{\frac{1}{2}} + \frac{x}{c_n};\ \max_{v\in S_i^0} Y_v^i \leqq y\left(\frac{1-\rho(n^\alpha)}{1+\rho(n^\alpha)}\right)^{\frac{1}{2}}\right\}\varphi(y)\,dy$$

apart from error terms of $o(n^{-\beta})$. We note that $S_D^0 = S_D \cap \{|u - t(i)| > n^\alpha\}$ and $S_i^0 = S_i \cap \{|u - t_i| > n^\alpha\}$. Define $CM_n^D = \max_{u\in S_D^0} Y^i$ and $cm_{n,k}^i = \max_{v\in S_i^0} Y_v^i$. Since y is at least $\eta_{0,n} = O((\log n)^{\frac{1}{2}})$, we can write

$$y\left(\frac{1-\rho(n^\alpha)}{1+\rho(n^\alpha)}\right)^{\frac{1}{2}} = y + \frac{o(1)}{c_n} = y_2, \quad \text{say,}$$

and let $y + (x + o(1))/c_n = y_1$. By Berman ((1971), 4.2, page 931) we can write the expression in (2.7) as

$$(2.8)\qquad \sum_{i=1}^{n'} \int_{\eta_{0,n}}^{\eta_{z,n}} P\{CM_n^D \leqq y_1\}P\{Cm_{n,k}^i \leqq y_2\}\varphi(y)\,dy + \sum_{i=1}^{n'} \int_{\eta_{0,n}}^{\eta_{z,n}} E_{i,y}\varphi(y)\,dy$$

where

$$(2.9)\qquad |E_{i,y}| \leqq \sum_{u=0}^{n}\sum_{v=0}^{n} \frac{|\gamma_{uv}^i|}{(1-\gamma_{uv}^{i2})^{\frac{1}{2}}} \exp\left\{-\frac{(y_1^2 - 2\gamma_{uv}^i y_1 y_2 + y_2^2)}{2(1-\gamma_{uv}^{i2})}\right\},$$

γ_{uv}^i is defined above (2.5). Also by Berman (1964, Lemma 3.1)

$$(2.10)\qquad |P\{CM_n^D \leqq y_1\} - \Phi^{n_1}(y_1)| \leqq \sum\sum_{u,v\in S_D^0} \frac{|\gamma_{uv}^i|}{(1-\gamma_{uv}^{i2})^{\frac{1}{2}}} \exp\left\{-\frac{y_1^2}{1+\gamma_{uv}^i}\right\}$$

and

$$(2.11)\qquad |P\{Cm_{n,k}^i \leqq y_2\} - \Phi^{n_2}(y_2)| \leqq \sum\sum_{u,v\in S_i^0} \frac{|\gamma_{uv}^i|}{(1-\gamma_{uv}^{i2})^{\frac{1}{2}}} \exp\left\{-\frac{y_2^2}{1+\gamma_{uv}^i}\right\}.$$

Here $n_1 = |S_D^0|$ is cardinality of S_D^0 and $n_2 = |S_i^0|$. Also $n_1/n \to (k-1)/k$ and $n_2/n \to 1/k$. We notice that in the r.h.s. of (2.9), (2.10), and (2.11), y_1 and y_2 could be replaced by y since $y = O((\log n)^{\frac{1}{2}})$. In the following we bound the sum in the r.h.s. of (2.9) by splitting it in two parts $|u - v| \leqq n^\alpha$ and $|u - v| > n^\alpha$. It's easy to see that the same bounds will work for the r.h.s. of (2.10) and (2.11). Since for each i, $|u - t(i)| > n^\alpha$, $\gamma_{uv}^i \leqq (r_{uv} + \rho(n^\alpha))/(1 - \rho(n^\alpha)) = \gamma_{uv}$, say, and $|\gamma_{uv}| \leqq \delta < 1$ for large n as well as $|\gamma_{uv}| \leqq (\rho(u - v) + \rho(n^\alpha))/(1 - \rho^2(n^\alpha))$. Thus we rewrite (2.8) as

$$\left[\frac{n}{k}\right] \int_{\eta_{0,n}}^{\eta_{z,n}} \Phi^{n_1}(y_1)\Phi^{n_2}(y_2)\varphi(y)\, dy + E \tag{2.12}$$

where

$$\begin{aligned} |E| &\leqq (\text{const.})n \cdot \int_{\eta_{0,n}}^{\eta_{z,n}} \sum_{u=0}^{n} \sum_{v=0}^{n} |\gamma_{uv}| \exp\{-y^2/(1 + \gamma_{uv})\}\varphi(y)\, dy \\ &\leqq (\text{const.})n \cdot \sum_{u=0}^{n} \sum_{v=0}^{n} |\gamma_{uv}| \exp\{-\eta_{0,n}^2/(1 + \gamma_{uv})\}(1 - \Phi(\eta_{0,n})) \\ &\leqq (\text{const.})e^{g(n)}\{\textstyle\sum\sum_{|u-v|\leqq n^\alpha} |\gamma_{uv}| \exp\{-\eta_{0,n}^2/(1 + \gamma_{uv})\} \\ &\quad + \textstyle\sum\sum_{|v-u|>n^\alpha} |\gamma_{uv}| \exp\{-\eta_{0,n}^2/(1 + \gamma_{uv})\}\}\,. \end{aligned} \tag{2.13}$$

The first term in the r.h.s. is at most $n^{1+\alpha}\delta \exp\{-\eta_{0,n}^2/(1 + \delta)\}$ which is $o(n^{-\beta})$ for some $\beta > 0$ by the same argument as before. For $|u - v| > n^\alpha$, $|\gamma_{uv}| < 3\rho(n^\alpha)$ and the second sum in the r.h.s. of (2.13) is at most

$$\begin{aligned} &(\text{const.})n^2\rho(n^\alpha) \exp\{-\eta_{0,n}^2/(1 + 3\rho(n^\alpha))\} \\ &\quad \leqq (\text{const.})\rho(n^\alpha) \log n \exp\left\{-2 \log n \left(\frac{1}{1 + 3\rho(n^\alpha)} - 1\right) + 2g(n)\right\} \\ &\quad = (\text{const.})\rho(n^\alpha) \log n \exp\{\rho(n^\alpha) \log n + 2g(n)\}\,. \end{aligned}$$

Substituting in (2.13), $|E| \to 0$ as $n \to \infty$ via choice of $g(n)$. It remains to show that the first term in (2.12) approaches the right limit as $n \to \infty$.

$$\begin{aligned} &\left[\frac{n}{k}\right] \int_{\eta_{0,n}}^{\eta_{z,n}} \Phi^{n_1}(y_1)\Phi^{n_2}(y_2)\varphi(y)\, dy \\ &\quad \sim \frac{n}{k} \int_{-g(n)}^{z} \Phi^{n_1}\left(b_{n,k} + \frac{x + y}{c_n}\right) \Phi^{n_2}\left(b_{n,k} + \frac{y}{c_n}\right) \frac{\varphi(b_{n,k} + y/c_n)}{c_n}\, dy\,. \end{aligned} \tag{2.14}$$

Using the definition of $b_{n,k}$, we see that

$$\begin{aligned} \Phi^{n_1}\left(b_{n,k} + \frac{x + y}{c_n}\right) &= \exp\left\{n_1 \log \Phi\left(b_{n,k} + \frac{x + y}{c_n}\right)\right\} \\ &\sim \exp\left\{-\frac{n_1\varphi\left(b_{n,k} + \frac{x + y}{c_n}\right)}{b_{n,k} + \frac{x + y}{c_n}}\right\}. \end{aligned}$$

Thus for $-g(n) \leqq y \leqq z$,

$$\begin{aligned} &\Phi^{n_1}\left(b_{n,k} + \frac{x + y}{c_n}\right) \exp\{(k - 1)e^{-(x+y)}\} \\ &\quad \leqq \exp\left\{(k - 1)e^{-(x+y)}\left[1 - \exp\left[-\frac{g(n) \log \log n}{c_n}\right]\right]\right\} + o((\log n)^{-\frac{1}{2}}) \\ &\quad \leqq \exp\left\{(\text{const.}) \cdot \frac{g(n) \log \log n}{c_n}\, e^{g(n)}\right\}. \end{aligned}$$

The r.h.s. above tends to 1 as $n \to \infty$ by choice of $g(n)$. Similar calculations show that

$$\left(\frac{n}{k}\Phi^{n_1}\left(b_{n,k}+\frac{x+y}{c_n}\right)\Phi^{n_2}\left(b_{n,k}+\frac{y}{c_n}\right)\frac{\varphi(b_{n,k}+y/c_n)}{c_n}\right)(e^{-(k-1)e^{-(x+y)}}e^{-e^{-y}}e^{-y})^{-1}$$

tends to 1 as $n \to \infty$ uniformly in y for $-g(n) \leqq y \leqq z$. Also the second term above is integrable in y. Applying the dominated convergence theorem we see that the l.h.s. of (2.14) tends to

$$\int_{-\infty}^{z} e^{-(1+(k-1)e^{-x})e^{-y}}e^{-y}\,dy = \frac{1}{1+(k-1)e^{-x}}\int_{(1+(k-1)e^{-x})e^{-z}}^{\infty} e^{-u}\,du$$
$$= \left\{\frac{\exp\{-e^{-z}(1+(k-1)e^{-x})\}}{1+(k-1)e^{-x}}\right\}.$$

Hence the result.

REMARK. The rate of convergence of (2.1) can be found in the above discussion. We note that via choice of the function $g(n)$ and in showing the second term in (2.13) is $o(1)$ the rate of convergence depends on the function $\rho(n)$ (and consequently on r_{ij}). If it can be assumed that $\rho(n)(\log n)^{\alpha} \to 0$ for some $0 < \alpha < 1$, then the rate of convergence could be shown to be $o((\log n)^{-\beta})$ for some $\beta > 0$. This is quite slow for practical calculation.

3. Strongly dependent sequences. In this section we assume throughout that $\{X_n, n \geqq 0\}$ is a standard stationary Gaussian sequence with $r_n \equiv EX_iX_{i+n}$. We assume that (1.2) holds. The aim (realized in Theorem 3.1) will be to show that the same limiting distribution as in (2.1) holds for $M_n - m_{n,k}$ under these conditions where M_n and $m_{n,k}$ are the same as before. (However, definitions of various other quantities are different in this section from the last and should be carefully noted.)

As was seen in Mittal and Ylvisaker (1975), the limit distribution of the normalized maxima will now change due to the strong interdependence of the variables. The fact that such dependence plays no part in the distribution considered in Theorem 2.1 seems odd at first but could be explained if we look at the representation of the sequence considered in [3]. We can write for $0 \leqq i \leqq n$

$$X_i = (1-r_n)^{\frac{1}{2}}\xi_i + (r_n)^{\frac{1}{2}}U$$

where $\{\xi_i, 0 \leqq i \leqq n\}$ are standard normal with covariance sequence $r_i' = (r_i - r_n)/(1-r_n)$ and U is again standard normal independent of $\{\xi_i, 0 \leqq i \leqq n\}$. (The dependence of the variables ξ_i and r_i' on n is suppressed for the sake of simpler notation.) Thus

$$M_n = (1-r_n)^{\frac{1}{2}}\max\nolimits_{0\leqq i\leqq n}\xi_i + (r_n)^{\frac{1}{2}}U = (1-r_n)^{\frac{1}{2}}M_n' + (r_n)^{\frac{1}{2}}U\,, \quad \text{say,}$$
$$m_{n,k} = (1-r_n)^{\frac{1}{2}}\max\nolimits_{i\in G_n}\xi_i + (r_n)^{\frac{1}{2}}U = (1-r_n)^{\frac{1}{2}}m_{n,k}' + (r_n)^{\frac{1}{2}}U\,, \quad \text{say,}$$

and

$$C_n(M_n - m_{n,k}) = C_n(1-r_n)^{\frac{1}{2}}(M_n' - m_{n,k}')\,, \tag{3.1}$$

the part $(r_n)^{\frac{1}{2}}U$ which makes the process strongly dependent cancels out and allows us to see the same behavior as in the last section for the difference of the two maxima. Even though the ideas are the same the proofs are longer and more involved due to the bulky nature of the covariance function.

The following theorem is stated somewhat differently than Theorem 2.1 because the second order terms of $m_{n,k}$ are not comparable with the difference $M_n - m_{n,k}$.

THEOREM 3.1. *If* (1.2) *holds for* $\{X_n, n \geqq 0\}$ *as described in the beginning of this section then*

$$P\left\{M_n - m_{n,k} \leqq \frac{x}{c_n}\right\} \to \frac{1}{1 + (k-1)e^{-x}}, \tag{3.2}$$

as $n \to \infty$ *for all* $0 \leqq x < \infty$.

PROOF. Looking at (3.1), it is sufficient to show that $P(M_n' - m'_{n,k} \leqq x/c_n(1-r_n)^{\frac{1}{2}})$ converges to the r.h.s. of (3.2). Let us define $\eta_{0,n} = b_{n,k} - g(n)/c_n$ where $g(n) \to \infty$ is chosen such that $(r_n)^{\frac{1}{2}}e^{2g(n)} = o(1)$. (Notice the different definitions of $\eta_{0,n}$ and $g(n)$.) In order to follow the proof of Section 2, we first have to show that $P\{m'_{n,k} < \eta_{0,n}\} = o(1)$. The proof of this follows that of (2.17) in [3]. Define

$$\begin{aligned} \tau(-1) &\equiv 0 \\ \tau(0) &= [n^\alpha] \qquad \text{for} \quad 0 < \alpha < \frac{1-\delta}{(1+\delta)^2} \\ \tau(i) &= n \exp\{-r_n^{i/2} \log n\} \qquad i = 1, 2, \cdots, q(n) \\ \tau(q(n)+1) &= n \end{aligned} \tag{3.3}$$

where $q(n)$ is chosen so that

$$r_n^{(q(n)+1)/2} > \frac{(r_n)^{\frac{1}{4}}}{\log n} \geqq r_n^{(q(n)+2)/2}. \tag{3.4}$$

As illustrated in [3], (1.2)(ii) implies that

$$r_{\tau(j)} - r_n \leqq r_n^{(j+2)/2}.$$

By Berman's lemma [1]

$$\begin{aligned} &P\{m'_{n,k} < \eta_{0,n}\} \\ &\leqq \Phi^{n'}(\eta_{0,n}) + (\text{const.}) \textstyle\sum_i \sum_j r'(i-j) \exp\left\{-\frac{\eta_{0,n}^2}{1+r'(i-j)}\right\} \\ &\leqq \Phi^{n'}(\eta_{0,n}) + (\text{const.}) n \textstyle\sum_{i=0}^{q(n)} \tau(i+1) r'_{\tau(i)} \exp\left\{-\frac{\eta_{0,n}^2}{1+r'_{\tau(i)}}\right\}. \end{aligned} \tag{3.5}$$

Direct computations will show that the first term of the r.h.s. of (3.5) is $o(1)$ for any $g(n) \to \infty$, and that the second term is $o(1)$ can be seen by some modification of the procedure in [3]. The same method is also used in (3.14). Hence we exclude the details at this stage. Thus it is sufficient to show that

$$(3.6)\qquad P\left\{M_n' - m'_{n,k} \leqq \frac{x}{c_n(1-r_n)^{\frac{1}{2}}}\,;\ m'_{n,k} > \eta_{0,n}\right\}$$

tends to the desired limit in (3.2) as $n \to \infty$. Conditioning on $\xi_{t_i} \in G_n$ and proceeding as in Section 2, the quantity in (3.6) is equal to

$$(3.7)\qquad \sum_{i=1}^{n'} \int_{\eta_{0,n}}^{\infty} P\left\{Y_u^{\,i} \leqq y\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}} + \frac{x}{c_n(1-r_n)^{\frac{1}{2}}}\quad u \in S_D,\right.$$
$$\left. Y_v^{\,i} \leqq y\left(\frac{1-r'_{v,t(i)}}{1+r'_{v,t(i)}}\right)^{\frac{1}{2}}\quad v \in S_i\right\}\varphi(y)\,dy$$

where $\{Y_u^{\,i},\ 0 \leqq u \leqq n\}$ are standard Gaussian with

$$EY_u^{\,i}Y_v^{\,i} = \frac{r'_{u,v} - r'_{u,t(i)}r'_{v,t(i)}}{(1-r'^2_{u,t(i)})^{\frac{1}{2}}(1-r'^2_{v,t(i)})^{\frac{1}{2}}} = \gamma^i_{uv}\,,\quad \text{say.}$$

By Berman (1971, (4.5), page 932) we can write (3.7) as

$$(3.8)\qquad \sum_{i=1}^{n'} \int_{\eta_{0,n}}^{\infty} \prod_{u\in S_D} \Phi\left(y\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}} + \frac{x}{c_n(1-r_n)^{\frac{1}{2}}}\right)$$
$$\times \prod_{v\in S_i} \Phi\left(y\left(\frac{1-r'_{v,t(i)}}{1+r'_{v,t(i)}}\right)^{\frac{1}{2}}\right)\varphi(y)\,dy + \sum_{i=1}^{n'}\int_{\eta_{0,n}}^{\infty} E_{i,y}\varphi(y)\,dy$$

where the error $E_{i,y}$ is such that

$$(3.9)\qquad |E_{i,y}| \leqq \sum_u \sum_v \left|\int_0^{\gamma^i_{uv}} \varphi(y_1, y_2, \lambda)\,d\lambda\right|$$

and

$$\varphi(y_1, y_2, \lambda) = (2\pi)^{-1}(1-\lambda^2)^{-\frac{1}{2}}\exp\left\{-\frac{y_1^{\,2} - 2\lambda y_1 y_2 + y_2^{\,2}}{2(1-\lambda^2)}\right\}.$$

We know that y_1 is either $y_u = y((1-r'_{u,t(i)})/(1+r'_{u,t(i)}))^{\frac{1}{2}}$ or $y_u + (x/c_n(1-r_n)^{\frac{1}{2}})$ and y_2 is either $y_v = y((1-r'_{v,t(i)})/(1+r'_{v,t(i)}))^{\frac{1}{2}}$ or $y_v + (x/c_n(1-r_n)^{\frac{1}{2}})$. Noticing that $|\gamma^i_{uv}| \leqq \delta < 1$ for large n for each i, we can see that

$$(3.10)\qquad |E_{i,y}| \leqq (\text{const.}) \sum_u \sum_v |\gamma^i_{uv}| \exp\left\{-\frac{y_u^{\,2} - 2|\gamma^i_{uv}|y_u y_v + y_v^{\,2}}{2(1-\gamma^{i2}_{uv})}\right\}$$
$$= (\text{const.}) \sum_u \sum_v h(u, v)\,,\quad \text{say.}$$

Define

$$\Delta_0 = \sum_{0<|u-t(i)|\leqq\tau(0)} \sum_v h(u, v) = \sum_u \sum_{0<|v-t(i)|\leqq\tau(0)} h(u, v)\,.$$

The last equality follows by symmetry. We notice that the sum of $h(u, v)$ over the set $\{|u - t(i)| \leqq \tau(0)$ or $|v - t(i)| \leqq \tau(0)\}$ is at most $2\Delta_0$. Also define

$$\Delta_j = \text{sum of}\quad h(u, v)\quad \text{on set}$$
$$\{\tau(j-1) < |u - t(i)| \leqq \tau(j)\} \cap \{\tau(j-1) < |v - t(i)| \leqq \tau(0)\}$$

$j = 1, 2, \cdots, (q(n) + 1)$. Then by the same argument as above the sum in the r.h.s of (3.10) is at most $2\sum_{j=0}^{q(n)+1}\Delta_j$. Next we will estimate Δ_j, $j = 0, 1, \cdots, (q(n)+1)$ for $y > \eta_{0,n}$. First

$$\Delta_0 = \sum_{|u-t(i)|\leqq\tau(0)}\sum_{|v-t(i)|\leqq 2\tau(0)} h(u, v) + \sum_{|u-t(i)|\leqq\tau(0)}\sum_{|v-t(i)|>2\tau_0} h(u, v)\,. \tag{3.11}$$

To bound the exponent in $h(u, v)$,

$$\begin{aligned}\frac{y_u^2 - 2|\gamma_{uv}^i|y_u y_v + y_v^2}{2(1-\gamma_{uv}^{i2})} &= y^2\left\{\tfrac{1}{2}(1-\gamma_{uv}^{i2})^{-1}\left[\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}} - \left(\frac{1-r'_{v,t(i)}}{1+r'_{v,t(i)}}\right)^{\frac{1}{2}}\right]^2\right. \\ &\quad \left. + (1+|\gamma_{uv}^i|)^{-1}\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}}\left(\frac{1-r'_{v,t(i)}}{1+r'_{v,t(i)}}\right)^{\frac{1}{2}}\right\} \\ &\geqq \frac{y^2}{1+|\gamma_{uv}^i|}\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}}\left(\frac{1-r'_{v,t(i)}}{1+r'_{v,t(i)}}\right)^{\frac{1}{2}}.\end{aligned} \tag{3.12}$$

We will use this bound for all sums Δ_j except the second term in the r.h.s. of (3.11). For this term notice that when $|u - t(i)| \leqq \tau(0)$; $|v - t(i)| > 2\tau(0)$ we have $|u - v| > \tau(0)$ and $|\gamma_{uv}^i| \leqq (\text{const.})r'(n^\alpha)$. Thus the l.h.s. of (3.12) is at least

$$\frac{y^2}{2}\left(\frac{1-\delta}{1+\delta} + \frac{1-r'(n^\alpha)}{1+r'(n^\alpha)} - (\text{const.})r'(n^\alpha)\right).$$

Substituting, we get

$$\begin{aligned}\Delta_0 &\leqq (\text{const.})\left\{\tau^2(0)\exp\left(-y^2\frac{(1-\delta)}{(1+\delta)^2}\right)\right. \\ &\quad \left. + n\tau(0)\exp\left(-y^2\left(\frac{1}{1+\delta} - (\text{const.})r'(n^\alpha)\right)\right)\right\} \\ &\leqq (\text{const.})e^{2g(n)}\left\{\exp\left(2\log n\left(\alpha - \frac{1-\delta}{(1+\delta)^2}\right)\right)\right. \\ &\quad \left. + \exp\left(\log n\left(\alpha + (\text{const.})r'(n^\alpha) - \frac{1-\delta}{1+\delta}\right)\right)\right\}.\end{aligned}$$

Since $r'(n^\alpha) = o(1)$, by choosing $\alpha < (1-\delta)/(1+\delta)^2 - \beta$ for some $\beta > 0$ we see that the r.h.s. above is $o(n^{-\beta})$. Now

$$\begin{aligned}\Delta_j &\leqq (\text{const.})\sum\sum_{\tau(j-1)<|u-t(i)|;|v-t(i)|\leqq\tau(j)}|\gamma_{uv}^i| \\ &\quad \times \exp\left\{-\frac{y^2}{1+\gamma_{uv}^i}\left(\frac{1-r'_{\tau(j-1)}}{1+r'_{\tau(j-1)}}\right)\right\}.\end{aligned} \tag{3.13}$$

Before proceeding to find upper bounds for Δ_j let us notice that

$$\gamma_{uv}^i \leqq r'_{uv}(1-r'^2_{u,t(i)})^{-\frac{1}{2}}(1-r'^2_{v,t(i)})^{-\frac{1}{2}} \leqq 2(r_{u,v} - r_n)$$

for large n if $|u - t(i)| > \tau(0)$ and $|v - t(i)| > \tau(0)$. In the following discussion we will bound $2(r_{u,v} - r_n)$ above and use the same bound for $|\gamma_{uv}^i|$. This is justified because of the following. We will bound γ_{uv}^i below by, say, d and then show that $|d|$ is smaller than the smallest upper bound we use for $(r_{u,v} - r_n)$.

Now $\gamma^i_{uv} \geqq r'_{u,v} - r'_{u,t(i)} r'_{v,t(i)}$. Suppose $|u - t(i)| = a$ and $|v - t(i)| = b$; then the lowest value of $r'_{u,v}$ will be when $|u - v| = a + b$. Suppose $a \leqq b$; then $b \geqq (a + b)/2$. Thus

$$r'_{u,v} - r'_{u,t(i)} r'_{v,t(i)} \geqq r'_{a+b} - r_{n^\alpha} r'_{(a+b)/2}$$
$$= (r_{a+b} - r_n - r_{n^\alpha}(r_{(a+b)/2} - r_n))(1 - r_n)^{-1} .$$

But $r_{(a+b)/2} \leqq r_{a+b}(1 + \log 2/\log (a + b)/2)$ because of condition (1.2)(ii). Thus the r.h.s. above is at least

$$(1 - r_n)^{-n}\left[(r_{a+b} - r_n)(1 - r_{n^\alpha}) - \frac{r_{n^\alpha} r_{a+b} \log 2}{\log (a + b)/2}\right] \geqq -\frac{r_{n^\alpha} r_{a+b} \log 2}{(\log (a + b)/2)(1 - r_n)} .$$

We will see later that the lowest upper bound used for $r'_{uv} - r_n$ is $r_n{}^{(q(n)+2)/2}$. But

$$r_n{}^{(q(n)+2)/2} - \frac{r_{n^\alpha} r_{a+b} \log 2}{(\log (a + b)/2)(1 - r_n)} = r_n{}^{(q(n)+2)/2}(1 - (\text{const.})(r_n{}^{(q(n)-2)/2} \log n)^{-1}) .$$

By choice of $q(n)$ as in (3.4) we see that $r_n{}^{(q(n)-2)/2} \log n \to \infty$ as $n \to \infty$. Returning to (3.13), for $y \geqq \eta_{0,n}$

$$(\text{const.})\Delta_j \leqq \tau(j) \textstyle\sum_{|v-t(i)|>\tau(j-1)} |\gamma^i_{uv}| \exp\left\{-\frac{\eta^2_{0,n}}{1 + |\gamma^i_{uv}|}(1 - 2r'_{\tau(j-1)})\right\}$$

$$\text{(3.14)} \qquad \leqq \tau(j) \left\{\delta \exp\left(-\frac{\eta^2_{0,n}}{1 + \delta}(1 - 2r'_{\tau(j-1)})\right)\right.$$
$$+ \textstyle\sum_{l=1}^{j-1} \tau(l) r'_{\tau(l-1)} \exp\left(-\frac{\eta^2_{0,n}}{1 + 2r'_{\tau(l-1)}}(1 - 2r'_{\tau(l-1)})\right)$$
$$\left. + n r'_{\tau(j-1)} \exp(-\eta^2_{0,n}(1 - 4r'_{\tau(j-1)}))\right\} .$$

The first term above is at most

$$(\text{const.}) \log n \exp\left\{2g(n) - \log n\left(\frac{1 - \delta}{1 + \delta} - 4r'_{\tau(j-1)} + r_n{}^{j/2}\right)\right\} = O(n^{-\beta})$$

for some $\beta > 0$. The last term is at most

$$(\text{const.}) r'_{\tau(j-1)} \log n \exp\left\{2g(n) - r_n{}^{j/2} \log n\left(1 - \frac{8r'_{\tau(j-1)}}{r_n{}^{j/2}}\right)\right\}$$
$$\leqq (\text{const.}) r_n{}^{(j+1)/2} \log n \exp\left\{-\frac{r_n{}^{j/2} \log n}{2}\right\}$$

because of the bound shown right after (3.4). Lastly, the middle term in the r.h.s. of (3.14) is at most

$$(\text{const.}) e^{2g(n)} \textstyle\sum_{l=1}^{j-1} r_n{}^{(l+1)/2} \log n \exp\{-\log n(r_n{}^{l/2} + r_n{}^{j/2} - 4r'_{(l-1)} - 8r'_{(j-1)})\}$$
$$\leqq (\text{const.}) e^{2g(n)} \textstyle\sum_{l=1}^{j-1} r_n{}^{(l+1)/2} \log n \exp\left\{-\frac{r_n{}^{l/2} \log n}{2}\right\}$$

since $\tau'_{\tau(l-1)} \leqq 2r_n{}^{(l+1)/2}$ and $l \leqq j - 1$. Substituting in (3.14) we see that by choice of $g(n)$

$$\Delta_j \leqq (\text{const.})\left\{o(n^{-\beta}) + \sum_{l=1}^{j} r_n^{(l+1)/2} \log n \exp\left(-\frac{r_n^{l/2}}{2}\log n\right)\right\}$$
$$\leqq (\text{const.})\left\{o(n^{-\beta}) + \sum_{l=1}^{j} \exp\left(-\frac{r_n^{l/2}\log n}{4}\right)\right\}$$

since

$$-\frac{r_n^{l/2}\log n}{2} + \log(r_n^{(l+1)/2}\log n) \leqq -\frac{r_n^{l/2}\log n}{2} + 2\log(r_n^{l/2}\log n)$$

and $l \leqq j \leqq q(n)$ implies $r_n^{l/2}\log n \to \infty$ as $n \to \infty$. Thus the sum in the r.h.s. above is at most

$$\sum_{l=1}^{j} \frac{4}{r_n^{l/2}\log n}\,.$$

Choice of $q(n)$ implies $(r_n)^{\frac{1}{2}l}\log n > (r_n)^{-\frac{1}{2}(q(n)-l+\frac{1}{2})}$ and the upper bound for the above will be

$$4\sum_{l=1}^{j} (r_n)^{\frac{1}{2}(q(n)-l+\frac{1}{2})} = 4(r_n)^{\frac{1}{2}(q(n)-j+\frac{1}{2})} \sum_{l=0}^{j-1} (r_n)^{\frac{1}{2}l}$$
$$\leqq 8(r_n)^{\frac{1}{2}(q(n)-j+\frac{1}{2})}\,.$$

Thus

$$\sum_{j=1}^{q(n)} \Delta_j \leqq (\text{const.}) \sum_{j=1}^{q(n)} (r_n)^{\frac{1}{2}(q(n)-j+\frac{1}{2})} \leqq (\text{const.})(r_n)^{\frac{1}{4}}\,.$$

For the last sum $\Delta_{q(n)+1}$ we proceed as in (3.14). The only term that needs attention is the last in the r.h.s. of (3.14). Similar procedures will show that it is at most

$$(\text{const.})r_n^{(q(n)+2)/2}\log n \exp\{-r_n^{(q(n)+1)/2}\log n\} \leqq (\text{const.})(r_n)^{\frac{1}{4}}$$

by choice of $q(n)$. Thus

$$\sum_{i=1}^{n'} \int_{\eta_{0,n}}^{\infty} E_{i,y}\varphi(y)\,dy \leqq (\text{const.})(r_n)^{\frac{1}{4}}\exp(2g(n))$$

which by choice of $g(n)$ is $o(1)$. It remains to be seen that the first term in (3.8) approaches the right limit as $n \to \infty$.

Let us write $y = b_{n,k} + w/c_n$ and look at

$$\prod_{u\in S_D} \Phi\left(b_{n,k}\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}} + \frac{x+w+o(1)}{c_n}\right)$$
$$\simeq \exp\left\{-\sum_{u\in S_D} \frac{\varphi\left(b_{n,k}\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}} + \frac{x+w+o(1)}{c_n}\right)}{b_{n,k}\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}} + \frac{x+w+o(1)}{c_n}}\right\}.$$

We will split the sum in the exponent above as usual and see that it is approximately equal to

$$\sum_{j=0}^{q(n)+1} \sum_{\tau(j-1)\leqq|u-t(i)|<\tau(j);\,u\in S_D} (\log n)^{-\frac{1}{2}}\varphi\left(b_{n,k}\left(\frac{1-r'_{u,t(i)}}{1+r'_{u,t(i)}}\right)^{\frac{1}{2}} + \frac{x+w}{c_n}\right) \tag{3.15}$$

for $-g(n) \leqq w < \infty$. The above sums for $0 \leqq j \leqq q(n)$ tend to zero and for $j = q(n) + 1$ it tends to $(k-1)e^{-(x+w)}$. The proof of this is very much similar to that of (3.14). In fact we can see that the first two terms of the r.h.s. of (3.14) could be taken as the upper bound for $\sum_{j=0}^{q(n)}$ in (3.15). The sum for $j = q(n) + 1$ is taken over all $u \in S_D$ such that $\tau(q(n)) \leqq |u - t(i)| < n$. It is easy to see that the number of terms in the sum are asymptotically equal to $(k-1)n/k$. Thus the sum for $j = q(n) + 1$ is about

$$\frac{(k-1)n}{k(\log n)^{\frac{1}{2}}} \exp\left\{-\frac{b_{n,k}^2}{2} - (x+w) + o(1)\right\}$$

which tends to $(k-1)\exp(-(x+w))$. Similarly we can show that

$$\prod_{v \in S_i} \Phi\left(b_{n,k}\left(\frac{1 - r_{v,t(i)}}{1 + r_{v,t(i)}}\right)^{\frac{1}{2}} + \frac{w}{c_n}\right)$$

tends to $\exp(-e^{-w})$. We follow the same steps as in the end of Section 2 to complete the proof of Theorem 3.1. Notice that the additional condition required there, namely $\log\log n/(\log n)^{\frac{1}{2}} \exp(2g(n)) = o(1)$, is automatically satisfied here since $\log\log n/(\log n)^{\frac{1}{2}}(r_n)^{\frac{1}{2}} = o(1)$ in view of (1.2)(ii) and the choice of $g(n)$.

REFERENCES

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

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

[3] Mittal, Yash and Ylvisaker, Donald (1975). Limit distributions for the maxima of stationary Gaussian processes. *Stochastic Processes Appl.* **3** 1–18.

Department of Statistics
Stanford University
Stanford, California 94305