ON THE MAXIMUM DEVIATION OF THE SAMPLE DENSITY

By Michael Woodroofe[1]

Carnegie Institute of Technology

1. Introduction and summary. Let $X_1, X_2, \cdots$ denote independent random variables having a common density f. The present paper considers estimates of f of the following form ([8] and [10]):

$$(1.1) \qquad f_n(x) = \tau \int G(x, \tau(x - y))\, dF_n(y) = (\tau/n) \sum_{i=1}^n G(x, \tau(x - X_i))$$

where F_n denotes the sample distribution function of $X_1, \cdots, X_n$, $\tau = \tau(n) \to \infty$ and $\tau = o(n)$ as $n \to \infty$, and G is a non-negative function defined on R^2 satisfying regularity conditions to be listed in Section 2. More precisely, it considers the asymptotic behavior as $n \to \infty$ of the maximum deviation of $f_n(x)$ from $f(x)$ where x varies in a compact interval which without loss of generality we take to be $[-1, 1]$. The main result, Theorem 3.1, states that under regularity conditions

$$(1.2) \qquad p \lim_{n\to\infty} \max_{|x| \leqq 1} (n/2\tau \log \tau)^{\frac{1}{2}} |(f_n(x) - f(x))/\|G_x\|_2 (f(x))^{\frac{1}{2}}| = 1$$

where G_x denotes an x-section of G—i.e. $G_x(y) = G(x, y)$, $y \varepsilon R^1$—and $\|\cdot\|_p$ denotes the norm in $L_p = L_p(R^1$, Lebesgue measure). Also, a limiting distribution related to (1.2) is computed, and some sufficient conditions are given for the almost sure convergence of $\max_{|x| \leqq 1} |f_n(x) - f(x)|$ to zero as $n \to \infty$.

Since under mild regularity conditions $f_n(x)$ and $f_n(y)$, $x \neq y$, are asymptotically, independently, normally distributed when suitably normalized ([5]), one might expect (1.2) to follow from its analogue for normal random variables ([2]) by an argument involving the weak convergence of stochastic processes. However, having been unable to verify the necessary compactness conditions ([9]) with respect to any topology which makes the maximum functional in (1.2) continuous, we have adopted a more elementary approach. This approach uses a theorem on the large deviations of sums of independent, identically distributed random vectors to estimate the relevant probabilities directly. In order to develop the primary topic of the paper as quickly as possible, we have postponed the proof of the theorem on large deviations until Section 5, while using it in Sections 3 and 4 to prove our main theorems. Section 2 presents some preliminary material.

We are aware of only two other papers which have considered the uniform convergence of f_n to f, namely [8] and [7].

Received 17 May 1966; revised 7 November 1966.

[1] The research for this paper was supported by the Office of Naval Research (Nonr 225(52)) while the author was at Stanford University. This paper in whole or part may be reproduced for any purpose of the United States government.

2. Preliminaries. Throughout Sections 2, 3, and 4, f and G will be assumed to satisfy the conditions listed below. f is bounded on R^1 and is positive and continuous on some neighborhood of $[-1, 1]$. G is a non-negative, bounded, measurable function on R^2 for which

$$\int G(x, y)\, dy = 1, \qquad x \varepsilon R^1, \tag{2.1a}$$

$$\sup_{x \varepsilon R^1} \int_{|y| \geqq t} |y|\, G(x, y)\, dy \to 0 \quad \text{as} \quad t \to \infty. \tag{2.1b}$$

These assumptions obviously imply that the family of functions $\{G_x : x \varepsilon R^1\}$ form a norm-bounded subset of L_p for every $p \varepsilon [1, \infty]$ and that

$$\sup_{x \varepsilon R^1} \int |y|\, G(x, y)\, dy < \infty.$$

For $0 < \alpha, \beta \leqq 1$ we will write $f \varepsilon$ Lip (α) and $G \varepsilon$ Lip (β) respectively when we wish to assume that f satisfies a uniform Lipschitz condition of order α on some neighborhood of $[-1, 1]$ and that G satisfies a uniform Lipschitz condition of order β on R^2.

The mean and variance of $f_n(x)$ are

$$\mu_n(x) = \tau \int G(x, \tau(x - y)) f(y)\, dy, \tag{2.2a}$$

$$(\tau/n)\sigma_n^2(x) = (\tau/n)(\tau \int G^2(x, \tau(x - y)) f(y)\, dy - \tau^{-1}\mu_n(x)^2), \tag{2.2b}$$

respectively, as is obvious from (1.1). Their asymptotic behavior as $n \to \infty$ is described by the following, easily verified lemma:

LEMMA 2.1. *There is an* $\eta' > 1$ *such that for all* $p \varepsilon [1, \infty)$

$$t \int G^p(x, t(x - y)) f(y)\, dy = f(x) \|G_x\|_p^p + o(1)$$

uniformly (*in* x) *on* $[-1, \eta']$ *as* $t \to \infty$. *If* $f \varepsilon$ Lip (α), *then* $o(1)$ *may be replaced by* $O(t^{-\alpha})$.

Thus, there exist $\eta > 1$ and n_0 for which $\sigma_n(x)$ is bounded away from zero on $[-1, \eta]$ for $n \geqq n_0$. Let

$$X_n(x) = (n/\tau)^{\frac{1}{2}}((f_n(x) - \mu_n(x))/\sigma_n(x)) = n^{-\frac{1}{2}} \sum_{i=1}^{n} Z_{n,i}(x), \tag{2.3a}$$

$$Z_{n,i}(x) = \tau^{\frac{1}{2}}((G(x, \tau(x - X_i)) - \tau^{-1}\mu_n(x))/\sigma_n(x)), \qquad i = 1, \cdots, n, \tag{2.3b}$$

for $-1 \leqq x \leqq \eta$ and $n \geqq n_0$. The covariance $r_n(x, y) = E(X_n(x)X_n(y))$ is

$$[\tau \int G(x, \tau(x - u)) G(y, \tau(y - u)) f(u)\, du - \tau^{-1}\mu_n(x)\mu_n(y)]/\sigma_n(x)\sigma_n(y), \tag{2.4}$$

and it is an easy consequence of (2.1) and Lemma 2.1 that

$$\sup |r_n(x, y)| = o((\log \tau)^{-1}) \quad \text{as} \quad n \to \infty \tag{2.5}$$

where the supremum is taken over the set T_n consisting of all (x, y) for which $-1 \leqq x, y \leqq \eta$ and $|x - y| \geqq 2\tau^{-1} \log \tau$. Moreover, if $G(x, y) = 0$ whenever $|y| \geqq \frac{1}{2}$, then (2.5) holds with $O(\tau^{-1})$ replacing $o((\log \tau)^{-1})$ when the supremum is taken over all (x, y) for which $-1 \leqq x, y \leqq \eta$ and $|x - y| \geqq \tau^{-1}$.

The following lemma is an easy consequence of Theorem 5.1 and (2.5); in it we have let $T_{n,p}$ be the set of $\mathbf{x} = (x_1, \cdots, x_p)' \varepsilon R^p$ for which $|x_i| \leqq 1$, $i = 1, \cdots, p$, and $\min_{i \neq j} |x_i - x_j| \geqq \tau^{-1}$, and we have denoted by Φ and Φ_r the standardized, univariate normal distribution function and the standardized, bivariate normal distribution function with parameter r respectively.

LEMMA 2.2. *Let* $w = w_n \to \infty$ *with* $w = o(\log \tau)$ *and* $\tau^\gamma = o(n)$, $\gamma > 1$, *as* $n \to \infty$ *and let* $\epsilon = +1$ *or* -1; *then as* $n \to \infty$

(i) $P(\epsilon X_n(x) \geqq w) \sim (1 - \Phi(w))$,

(ii) $P(X_n(x) \geqq w, X_n(y) \geqq w) \sim \int_w^\infty \int_w^\infty d\Phi_r(u, v)$ *where* $r = r_n(x, y)$ *uniformly on* $[-1, \eta]$ *and* T_n *respectively. If, in addition,* $G(x, y) = 0$ *whenever* $|y| \geqq \frac{1}{2}$, *then*

(iii) $P(\epsilon_i X_n(x_i) \geqq w, i = 1, \cdots, p) \sim (1 - \Phi(w))^p$ *as* $n \to \infty$ *uniformly on* $T_{n,p}$.

3. Convergence in probability.

THEOREM 3.1. *Let* $f \varepsilon$ Lip (α) *and* $G \varepsilon$ Lip (β), $0 < \alpha, \beta \leqq 1$, *and let* $\tau^\gamma = o(n)$ *and* $n = o(\tau^\delta)$ *as* $n \to \infty$ *with* $1 < \gamma < \delta \leqq 1 + 2\alpha$; *then* (1.2) *holds.*

PROOF. Since $(n/\tau)^{\frac{1}{2}} |\mu_n(x) - f(x)| \to 0$ and $\sigma_n(x) \to \|G_x\|_2(f(x))^{\frac{1}{2}}$ uniformly on $[-1, 1]$ as $n \to \infty$ by Lemma 2.1, it will suffice to show that for every positive ϵ

$$P(\max_{|x| \leqq 1} X_n(x) \leqq (1 - \epsilon)(2 \log \tau)^{\frac{1}{2}}) \to 0, \tag{3.1a}$$

$$P(\max_p \max_{0 \leqq x \leqq 1} |X_{n,p}(x)| \geqq (1 + \epsilon)(2 \log \tau)^{\frac{1}{2}}) \to 0 \tag{3.1b}$$

as $n \to \infty$ where the $X_{n,p}(x)$ processes are defined as follows: for n so large that $\tau^{-1} < \eta - 1$, say $n \geqq n_1 \geqq n_0$, and for $p = 0, \cdots, [2\tau]$ where $[\cdot]$ denotes the greatest integer function, let $x_{n,p} = -1 + p\tau^{-1}$, and let

$$\sigma_{n,p}(x) = \sigma_n(x_{n,p} + x\tau^{-1})/\sigma_n(x_{n,p}), \qquad 0 \leqq x \leqq 1, \tag{3.2a}$$

$$X_{n,p}(x) = \sigma_{n,p}(x) X_n(x_{n,p} + x\tau^{-1}), \qquad 0 \leqq x \leqq 1. \tag{3.2b}$$

Using Lemma 2.2 and (2.5), one may establish (3.1a) by essentially the same argument that is used to establish the corresponding assertion in [4]. To establish (3.1b), we have first to find a sequence of finite subsets of [0, 1] on which the maxima will essentially be attained with probability approaching one.

LEMMA 3.1. *If* $G \varepsilon$ Lip (β), $0 < \beta \leqq 1$, *then there is a constant* C *for which*

$$E(\exp(|(X_{n,p}(x) - X_{n,p}(y))/|x - y|^{\beta/2}|)) \leqq C$$

for all $0 \leqq x, y \leqq 1$, $p = 0, \cdots, [2\tau]$, *and* $n \geqq n_1$.

PROOF. For fixed x, y, n, and p, let M denote the moment generating function of $U = (\sigma_{n,p}(x) Z_{n,1}(x_{n,p} + x\tau^{-1}) - \sigma_{n,p}(y) Z_{n,1}(x_{n,p} + y\tau^{-1}))$; then by (2.3) and (3.2)

$$E(\exp((X_{n,p}(x) - X_{n,p}(y))/|x - y|^{\beta/2})) = (1 + M''(t)/2n\,|x - y|^\beta)^n \tag{3.3}$$

where M'' denotes the second derivative of M and $0 < t < (1/n\,|x - y|^\beta)^{\frac{1}{2}}$. Now $G \varepsilon$ Lip (β) implies $|U| \leqq C_1 \tau^{\frac{1}{2}} |x - y|^\beta$ wp one and Var $(U) \leqq C_2 |x - y|^\beta$ with

C_1 and C_2 independent of x, y, n, and p. Therefore, $M''(t) \leqq C_3 |x - y|^\beta$, so that the right side of (3.3) does not exceed $\exp(C_3/2)$. Since the roles of x and y may be exchanged, the lemma follows.

If g is a continuous function on [0, 1], let $g^{(k)}$ denote the continuous broken line with vertices at $(i2^{-k}, g(i2^{-k}))$, $i = 0, \cdots, 2^k$; i.e. $g^{(k)}$ assumes the values $g(i2^{-k})$ at the points $i2^{-k}$, $i = 0, \cdots, 2^k$, and varies linearly between them.

Lemma 3.2. *Let $G \varepsilon \operatorname{Lip}(\beta)$, $0 < \beta \leqq 1$; then there exist constants d and D for which*

$$P(\max_{0 \leqq x \leqq 1} |X_{n,p}(x) - X_{n,p}^{(k)}(x)| \geqq u2^{-k\beta/4}) \leqq D \exp(-du2^{k\beta/4})$$

for all $p = 0, \cdots, [2\tau]$, $n \geqq n_1$, $u > 0$, and $k \geqq 1$.

Proof. The proof of Lemma 3.2 is nearly a verbatim repetition of the discussion in [9] on pp. 178–179. Simply use Lemma 3.1 in place of equation (2.8) of [9].

Now if $\epsilon = 3\epsilon_1$ is given and we let $2^k \sim (\log \tau)^{4/\beta}$ as $n \to \infty$ and $w_n = (1 + 2\epsilon_1)(2 \log \tau)^{\frac{1}{2}}$, $n \geqq 1$, then for n sufficiently large the left side of (3.1b) will not exceed

$$(3.4) \quad \sum_{p=1}^{[2\tau]} \sum_{i=0}^{2^k} P(|X_{n,p}(i2^{-k})| \geqq w_n) + 2D\tau \exp(-2\delta 2^{k\beta/4}) = o(\tau^{-2\epsilon_1}) + O(\tau^{-2\delta+1}) = o(1) \quad \text{as} \quad n \to \infty. \quad \text{QED}$$

If $\epsilon_1 \geqq \delta$, then (3.4) is the nth term of a summable series. Since we did not use the hypothesis that $f \varepsilon \operatorname{Lip}(\alpha)$ to establish (3.4), the following theorem is an easy consequence of Lemma 2.1 and the Borel-Cantelli lemmas.

Theorem 3.2. *Let $G \varepsilon \operatorname{Lip}(\beta)$, $0 < \beta \leqq 1$, and let $\tau^\gamma = o(n)$ and $n = o(\tau^\delta)$ as $n \to \infty$ with $1 < \gamma < \delta$. Then $\max_{|x| \leqq 1} |f_n(x) - f(x)| \to 0$ as $n \to \infty$ wp one. If $f \varepsilon \operatorname{Lip}(\alpha)$, $\frac{1}{2} < \alpha \leqq 1$, and if $\gamma > 2$, then $\max_{|x| \leqq 1} |f_n(x) - f(x)| = o(\tau^{-\frac{1}{2}})$ as $n \to \infty$ wp one.*

4. A limiting distribution. In this section we will use the technique of Watson ([11]) to compute the asymptotic distribution as $n \to \infty$ of

$$(4.1) \quad M_n = a_n^{-1}(\max_{|i| \leqq \tau} (n/\tau)^{\frac{1}{2}} |(f_n(x_i) - f(x_i))/\|G_{x_i}\|_2 (f(x_i))^{\frac{1}{2}}| - b_n)$$

where for $|i| \leqq \tau$ and $n \geqq 1$, $x_i = i\tau^{-1}$ and

$$(4.2a) \quad a_n = (2 \log 4\tau)^{-\frac{1}{2}},$$

$$(4.2b) \quad b_n = (2 \log 4\tau)^{\frac{1}{2}} - (\tfrac{1}{2})(2 \log 4\tau)^{-\frac{1}{2}}(\log \log 4\tau + \log 2\pi).$$

Theorem 4.1. *Let $f \varepsilon \operatorname{Lip}(\alpha)$, $0 < \alpha \leqq 1$; let $G(x, y) = 0$ whenever $|y| \geqq \frac{1}{2}$; and let $\tau^\gamma = o(n)$ and $n = o(\tau^\delta)$ as $n \to \infty$ with $1 < \gamma < \delta < 1 + 2\alpha$. Then $\lim_{n\to\infty} P(M_n < x) = \exp(-\exp(-x))$ for $x \varepsilon R^1$.*

Proof. By Lemma 2.1, it will suffice to prove the theorem with $M_n^* = a_n^{-1}(\max_{|i| \leqq \tau} |X_n(x_i)| - b_n)$ replacing M_n. (See Theorem 3.1, first line of the proof.) For $n \geqq n_0$, $p \geqq 1$, and $x \varepsilon R^1$, let $w_n = a_n x + b_n$ and

$$(4.3) \quad S_{n,p} = \sum_{(p)} P(|X_n(i_j\tau^{-1})| \geqq w_n, j = 1, \cdots, p)$$

where $\sum_{(p)}$ denotes summation over all subsets of size p drawn from $\{-[\tau], \cdots, [\tau]\}$. Then it is an easy consequence of Lemma 2.2 that $S_{n,p} \to (1/p!) \exp(-x)$ as $n \to \infty$ for each fixed p, so that the theorem follows, as in [11], from the inequalities on pp. 99–100 of [6]. QED

We remark that no generality would be gained by requiring G to vanish only for $|y| \geqq A$, for a factor of $2A$ may be absorbed in the sequence τ.

5. Large deviations. We will now present the theorem on large deviations from which Lemma 2.2 follows. Since the proof is a variation on that given in [3], it will only be sketched.

We will regard points in R^p as column vectors $\mathbf{x} = (x_1, \cdots, x_p)'$ where $'$ denotes transpose and will write $\mathbf{x}'\mathbf{y} = \sum_{i=1}^{p} x_i y_i$ for $\mathbf{x}, \mathbf{y} \in R^p$. $\Phi(\cdot:\mathbf{A})$ will denote the p-dimensional, normal distribution function with mean $\mathbf{0}$ and covariance matrix $\mathbf{A}$, $|\mathbf{A}|$ will denote the determinant of the matrix $\mathbf{A}$, and $\int \cdots \int_{\mathbf{x}}^{\infty}$ will abbreviate $\int_{x_1}^{\infty} \cdots \int_{x_p}^{\infty}$. Finally, we will call a vector $\mathbf{v}$ *admissible* with respect to a positive definite matrix $\mathbf{A}$ iff $\mathbf{u} = \mathbf{A}^{-1}\mathbf{v} > 0$ in the sense that $u_i > 0$, $i = 1, \cdots, p$.

THEOREM 5.1. *Let* $\{\mathbf{X}^{(n,k)}: k = 1, \cdots, k_n, n = 1, 2, \cdots\}$ *be a triangular array of p-dimensional random vectors which are independent and identically distributed in each row. Suppose that* (1) $E(\mathbf{X}^{(n,1)}) = \mathbf{0}, n \geqq 1,$ *and* $\mathbf{A}_n = E(\mathbf{X}^{(n,1)}\mathbf{X}^{(n,1)\prime}) \to \mathbf{A}_0$ *as* $n \to \infty$ *where* $\mathbf{A}_n$ *is positive definite for* $n \geqq 0$; (2) $\max_{i=1,\cdots,p} |X_i^{(n,1)}| \leqq c_n$ *wp one*, $n \geqq 1$; (3) $0 < w_n \to \infty$ *and* $c_n^2 w_n^{2q} = o(k_n)$ *as* $n \to \infty$ *where* $q = \max\{p, 3\}$; *and* (4) $\mathbf{v}$ *is admissible with respect to* $\mathbf{A}_0$. *Then as* $n \to \infty$

$$P(k_n^{-\frac{1}{2}} \textstyle\sum_{k=1}^{k_n} X_i^{(n,k)} \geqq v_i w_n,\ i = 1, \cdots, p) \sim \int \cdots \int_{\mathbf{v}w_n}^{\infty} d\Phi(\mathbf{x}:\mathbf{A}_n).$$

PROOF. For $n \geqq 1$ let $M^{(n)}$ denote the moment generating function of $\mathbf{X}^{(n,1)}$— i.e. $M^{(n)}(\mathbf{t}) = E(\exp(\mathbf{t}'\mathbf{X}^{(n,1)}))$, $\mathbf{t} \in R^p$—and let $\varphi^{(n)}$ and $\mathbf{B}^{(n)}$ denote respectively the vector and matrix functions of first and second order partial derivatives of $L^{(n)} = \log M^{(n)}$. Next $\mathbf{u}_n = \mathbf{A}_n^{-1}\mathbf{v}$, $\mathbf{t}_n = (w_n k_n^{-\frac{1}{2}})\mathbf{u}_n$, $\varphi_n = \varphi^{(n)}(\mathbf{t}_n)$, $\mathbf{B}_n = \mathbf{B}^{(n)}(\mathbf{t}_n)$, and $L_n = L^{(n)}(\mathbf{t}_n)$. Finally, define H_n and G_n by

$$G_n(\mathbf{x}) = \exp(-L_n) \int \cdots \int_{-\infty}^{\mathbf{x}} \exp(\mathbf{t}_n'\mathbf{y})\, dF_n(\mathbf{y}), \tag{5.1a}$$

$$H_n(\mathbf{x}) = G_n^{*k_n}(\mathbf{x}k_n^{-\frac{1}{2}} + k_n\varphi_n) \tag{5.1b}$$

where F_n denotes the distribution function of $\mathbf{X}^{(n,1)}$ and $*$ denotes convolution. We observe that by hypothesis (2)

$$\max_{t't \leqq c_n^{-2}} \max_{i,j,k} |(\partial^3/\partial s_i \partial s_j \partial s_k) L^{(n)}(s)|_{s=t}| \leqq Dc_n \tag{5.2}$$

where D is independent of n. We also observe that if $\mathbf{Q}_n \to \mathbf{A}_0$ as $n \to \infty$, then by the admissibility of $\mathbf{v}$ and the dominated convergence theorem

$$\begin{aligned} &w_n^p \int \cdots \int_{\mathbf{0}}^{\infty} \exp(-w_n\mathbf{x}'\mathbf{u}_n)\, d\Phi(\mathbf{x}:\mathbf{Q}_n) \\ &\qquad \to (|\mathbf{A}_0|^{-\frac{1}{2}}/(2\pi)^{p/2})(\textstyle\prod_{i=1}^{p} u_{0,i})^{-1} \quad \text{as} \quad n \to \infty \end{aligned} \tag{5.3}$$

where $\mathbf{u}_0 = \mathbf{A}_0^{-1}\mathbf{v} = (u_{0,1}, \cdots, u_{0,p})'$.

If one now introduces the Esscher transformation ([3]) and applies the multivariate normal approximation theorem ([1]) to H_n, one finds, after some routine analysis using (5.2), (5.3), hypothesis (3), and Taylor's theorem, that

$$\begin{aligned} &P(k_n^{-\frac{1}{2}} \textstyle\sum_{k=1}^{k_n} X_i^{(n,k)} \geqq v_i w_n,\ i = 1, \cdots, p) \\ (5.4)\qquad &= \exp(k_n(L_n - \mathbf{t}_n'\varphi_n)) \int \cdots \int_{\mathbf{z}_n}^{\infty} \exp(-w_n \mathbf{x}'\mathbf{u}_n)\, dH_n(\mathbf{x}) \\ &\sim \exp(k_n(L_n - \mathbf{t}_n'\varphi_n)) \int \cdots \int_{\mathbf{0}}^{\infty} \exp(-w_n \mathbf{x}'\mathbf{u}_n)\, d\Phi(\mathbf{x}:\mathbf{B}_n) \\ &= \exp(k_n(L_n - \mathbf{t}_n'\varphi_n + \tfrac{1}{2}\mathbf{t}_n'\mathbf{B}_n t_n)) \int \cdots \int_{\mathbf{v}_n w_n}^{\infty} d\Phi(\mathbf{x}:\mathbf{B}_n) \end{aligned}$$

as $n \to \infty$ where $\mathbf{z}_n = \mathbf{v}w_n - \varphi_n k_n^{\frac{1}{2}}$ and $\mathbf{v}_n = \mathbf{B}_n\mathbf{u}_n$. Now by (5.2), hypothesis (3), and Taylor's theorem $(L_n - \mathbf{t}_n'\varphi_n + \frac{1}{2}\mathbf{t}_n'\mathbf{B}_n\mathbf{t}_n) = o(k_n^{-1})$ as $n \to \infty$. Moreover, it follows from (5.3) that

$$(5.5\text{a})\qquad \int \cdots \int_{\mathbf{v}w_n}^{\infty} d\Phi(\mathbf{x}:\mathbf{A}_n) \sim C(\mathbf{v}, \mathbf{A}_0) w_n^{-p} \exp((-w_n^2/2)(\mathbf{v}'\mathbf{A}_n^{-1}\mathbf{v})),$$

$$(5.5\text{b})\qquad \int \cdots \int_{\mathbf{v}_n w_n}^{\infty} d\Phi(\mathbf{x}:\mathbf{B}_n) \sim C(\mathbf{v}, \mathbf{A}_0) w_n^{-p} \exp((-w_n^2/2)(\mathbf{v}_n'\mathbf{B}_n^{-1}\mathbf{v}_n))$$

as $n \to \infty$ with $C(\mathbf{v}, \mathbf{A}_0)$ equal to the right side of (5.3). Since finally $|\mathbf{v}'\mathbf{A}_n^{-1}\mathbf{v} - \mathbf{v}_n'\mathbf{B}_n^{-1}\mathbf{v}_n| = |\mathbf{u}'(\mathbf{A}_n - \mathbf{B}_n)\mathbf{u}_n| = o(w_n^{-2})$ as $n \to \infty$, the theorem follows. QED

We remark that in the presence of bounded higher moments or moment generating functions, Theorem 5.1 may be used to obtain large deviation results for unbounded random vectors as may be seen by a truncation argument. The details are omitted.

The first assertion in Lemma 2.2 (with $\epsilon = +1$) follows from an application of Theorem 5.1 to the triangular array $\{Z_{n,k}(x_n): k = 1, \cdots, n, n = 1, 2, \cdots\}$ where x_n is chosen to maximize $|P(X_n(x) \geqq w) - (1 - \Phi(w))|$ for $-1 \leqq x \leqq \eta$. Since $(1, 1)'$ is clearly admissible with respect to the identity matrix, the second part follows similarly. For the third, one also needs to use (5.5a) and the discussion immediately following (2.5).

REFERENCES

[1] Bergström, H. (1945). On the central limit theorem in the space R_k, $k > 1$. *Skand. Acturietidskr.* **28** 106–127.

[2] Berman, S. (1962). A law of large numbers for the maximum of a stationary Gaussian sequence. *Ann. Math. Statist.* **33** 93–98.

[3] Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilitiés. *Actualités Sci. Indust.* No. 736.

[4] Cramér, H. (1962). On the maximum of a normal stationary stochastic process. *Bull. Amer. Math Soc.* **68** 512–517.

[5] Craswell, K. J. (1965). Density estimation in a topological group. *Ann. Math. Statist.* **36** 1047–1048.

[6] Feller, W. (1959). *An Introduction to Probability Theory and its Applications*, **1.** Wiley, New York.

[7] Nadarya, E. A. (1965). On non-parameteric estimates of density functions and regression curves. *Theor. Prob. Appl.* **10** 186–190.

[8] PARZEN, E. (1962). On estimation of a probability density and mode. *Ann. Math. Statist.* **33** 1065–1076.

[9] PROKHOROV, YU. V. (1956). Convergence of random processes and limit theorems. *Theor. Prob. Appl.* **1** 157–222.

[10] ROSENBLATT, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.* **27** 832–835.

[11] WATSON, G. S. (1954). Extreme values in samples from m-dependent, stationary, stochastic processes. *Ann. Math. Statist.* **25** 798–800.