The Annals of Statistics
1975, Vol. 3, No. 4, 897–905

MULTIVARIATE PROBABILITIES OF LARGE DEVIATIONS

BY GERALD L. SIEVERS
Western Michigan University

Let $\{P_n\}_{n=1}^{\infty}$ denote a sequence of probability measures on (R^k, B^k), where R^k is k-dimensional Euclidean space and B^k the Borel subsets. For $A \in B^k$, let $e(A) = \lim_{n\to\infty} (1/n) \log P_n(A)$, if the limit exists. Sufficient conditions are given for expressing $e(A)$ as the supremum of $e(B)$ for certain "rectangular" sets $B = \times_{j=1}^{k} (\alpha_j, \beta_j)$ with either $\alpha_j = -\infty$ or $\beta_j = +\infty$ for each $j = 1, \cdots, k$. Also, some k-dimensional generalizations of the density theorem of Killeen, *et al.* (1972) are given for expressing $e(A)$ in terms of certain limits of the sequence of density (or probability) functions. Finally, an example is considered where P_n is the distribution of k order statistics from a sample of size n.

1. Notation and main results. For $n = 1, 2, \cdots$ let P_n denote a probability measure on (R^k, B^k) where R^k is k-dimensional Euclidean space and B^k the Borel subsets. For $A \in B^k$ define the extended real-valued function $(\log 0 = -\infty)$ $e(A) = \lim_{n\to\infty} (1/n) \log P_n(A)$, if the limit exists. Similarly define $\underline{e}(A)$ and $\bar{e}(A)$ using lim inf and lim sup, respectively.

This paper is concerned with the determination of $e(A)$, the exponential rate of convergence to zero of a sequence of probabilities $P_n(\mathbf{X} \in A)$, where $\mathbf{X} = (X_1, \cdots, X_k)$ is a random vector with distribution P_n. A special case of interest is $A = \{T \geqq a\}$, where $T = T(\mathbf{X})$ is some statistic. In this section some methods are developed for determining $e(A)$ from $e(B)$ for simpler events B of the form $B = \{X_1 * x_1, \cdots X_k * x_k\}$ where the $*$'s are either $\geqq$ or $\leqq$. In many cases it should be easier, using standard methods, to determine the rates $e(B)$ than to determine $e(A)$ directly. This point is discussed again after the theorems of this section.

First we introduce some notation.

Typical points in R^k are denoted $\mathbf{x} = (x_1, \cdots, x_k)$ and the closure (interior) of $A \in B^k$ is denoted $c\ell(A)(A^\circ)$.

With $A(1) = \{t : t \geqq 0\}$ and $A(2) = \{t : t \leqq 0\}$, define $M = 2^k$ subsets of R^k of the form $\times_{j=1}^{k} A(\alpha_j)$ where $\alpha_j = 1$ or 2 for $j = 1, \cdots, k$. These sets are denoted $Q_1(\mathbf{0}) = \times_{j=1}^{k} A(1), Q_2(\mathbf{0}), \cdots, Q_M(\mathbf{0})$.

For $\mathbf{x} \in R^k$ and $i = 1, \cdots, M$ let $Q_i(\mathbf{x}) = \{\mathbf{y} \in R^k : \mathbf{y} = \mathbf{x} + \mathbf{z}, \mathbf{z} \in Q_i(\mathbf{0})\}$ and $N_i(\mathbf{x}, \varepsilon) = \{\mathbf{y} \in R^k : \mathbf{y} = \mathbf{x} + \mathbf{z}, \mathbf{z} \in Q_i(\mathbf{0}), |z_j| \leqq \varepsilon \text{ for } j = 1, \cdots, k\}$ for $\varepsilon > 0$. Also let $m_i(\mathbf{x}) = e(Q_i(\mathbf{x}))$ if the limit exists, $\underline{m}_i(\mathbf{x}) = \underline{e}(Q_i(\mathbf{x}))$ and $\bar{m}_i(\mathbf{x}) = \bar{e}(Q_i(\mathbf{x}))$.

For all integers i, $1 \leqq i \leqq M$; r, $1 \leqq r \leqq k - 1$; $j_1, \cdots, j_r$, $1 \leqq j_1 < \cdots < j_r \leqq k$ and points $\mathbf{y} = (y_1, \cdots, y_r) \in R^r$ define the extended real-valued function

Received July 1973; revised October 1974.
AMS 1970 *subject classifications.* Primary 60F10; Secondary 62G20, 62G30.
Key words and phrases. Large deviations, density limit, order statistics, exact slope.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.*
www.jstor.org

$g_i(\mathbf{y}; j_1, \cdots, j_r) = \inf\{m_i(\mathbf{x}) : \mathbf{x} \in Q_i(\mathbf{0}),\ x_j = y_j \text{ for } j = j_1, \cdots, j_r\}$. The function is not defined for $\mathbf{y}$ at which the infimum of an empty set arises.

Consider the following conditions:

(A) For each $i = 1, \cdots, M$, $m_i(\mathbf{x})$ exists and is a continuous function of $\mathbf{x} \in Q_i(\mathbf{0})$,

(B) There exists a point in R^k which we take to be $\mathbf{0}$ (by suitable translation, if necessary) such that for all $\varepsilon > 0$, integers i, $1 \leqq i \leqq M$, and $\mathbf{x} \in Q_i(\mathbf{0})$ we have $m_i(\mathbf{x}) = e(N_i(\mathbf{x}, \varepsilon))$, and

(C) For all integers i, $1 \leqq i \leqq M$; r, $1 \leqq r \leqq k-1$, and $j_1, \cdots, j_r$, $1 \leqq j_1 < \cdots < j_r \leqq k$, the function $g_i(\mathbf{y}; j_1, \cdots, j_r)$ is a continuous function of $\mathbf{y}$ on its domain of definition.

THEOREM 1.1. *Assume conditions* (A) *and* (B) *hold. If* $A \in B^k$, $\emptyset \neq A \subset c\ell(A^\circ)$ *and* A *is bounded, then*

$$e(A) = \max_{1 \leqq i \leqq M} \sup\{m_i(\mathbf{x}) : \mathbf{x} \in A^\circ \cap Q_i(\mathbf{0})\}. \tag{1.1}$$

THEOREM 1.2. *Assume conditions* (A), (B) *and* (C) *hold. If* $A \in B^k$ *and* $\emptyset \neq A \subset c\ell(A^\circ)$, *then* (1.1) *holds.*

The following corollary is immediate.

COROLLARY 1.1. *For integers* i, j *with* $1 \leqq i \leqq j \leqq M$, *let* $m_i(\mathbf{x}) = m_j(\mathbf{x})$ *for all* $\mathbf{x} \in Q_i(\mathbf{0}) \cap Q_j(\mathbf{0})$. *Then the definition* $m(\mathbf{x}) = m_i(\mathbf{x})$ *for* $\mathbf{x} \in Q_i(\mathbf{0})$, *defines* $m(\mathbf{x})$ *unambiguously on* R^k. *Let the hypothesis of Theorem* 1.1 *or* 1.2 *hold. Then* $e(A) = \sup\{m(\mathbf{x}) : \mathbf{x} \in A\}$.

The following lemmas give sufficient conditions for conditions (B) and (C) that are frequently easy to verify in particular cases.

LEMMA 1.1. *The following condition is sufficient for condition* (B): *There exists a point in* R^k *which we take to be* $\mathbf{0}$ *such that for all integers* i, $1 \leqq i \leqq M$, $m_i(\mathbf{x})$ *is strictly decreasing on* $Q_i(\mathbf{0})$; *that is, if* $\mathbf{x} \in Q_i(\mathbf{0})$ *and* $m_i(\mathbf{x}) > -\infty$, *then* $m_i(\mathbf{x}') < m_i(\mathbf{x})$ *for all* $\mathbf{x}' \in Q_i(\mathbf{x})$ *with* $\mathbf{x}' \neq \mathbf{x}$.

LEMMA 1.2. *The following condition is sufficient for condition* (C): *For all integers* i, $1 \leqq i \leqq M$, *and* j, $1 \leqq j \leqq k$, $g_i(y; j) = -\infty$ *for all* $y \in R^1$ *for which the function is defined.*

Without assuming the existence of the limits $m_i(\mathbf{x})$ in condition (A) slightly more general versions of the theorems can be formulated involving $\underline{m}_i(\mathbf{x})$ and $\bar{m}_i(\mathbf{x})$ with essentially the same proofs. The conditions that $m_i(\mathbf{x})$ be continuous and that $A \subset c\ell(A^\circ)$ are natural for a result like (1.1). Without either one it is easy to find examples when $k = 1$ where (1.1) fails. Lemma 2.5 shows that condition (B) is necessary for (1.1) when $A \subset c\ell(A^\circ)$ and condition (A) holds. Condition (C) does not appear to be necessary for (1.1); however a related condition is needed when A is unbounded. The example at the end of Section 2 illustrates the difficulty.

In particular cases, the theorems can be used when A depends on n and (possibly after suitable transformation) such a sequence of sets converges in some sense.

It should be possible to extend the theorems to more general spaces than R^k.

In particular cases various means can be used to evaluate the $m_i(\mathbf{x})$ functions. Since $Q_i(\mathbf{x})$ is a product set, if P_n is a product measure, $(1/n)\log P_n[Q_i(\mathbf{x})]$ is a sum of k terms involving univariate probabilities and the limits can be examined separately (see example (a) of Section 4). If P_n is not a product measure a similar representation can be made using conditional probabilities although this approach does not appear to have been exploited in the literature.

It seems clear that a k-dimensional version of the theorem in Sievers (1969) or Theorem 2.2 in Bahadur (1971) holds using joint moment-generating functions and, if this is so, would be useful in computing $m_i(\mathbf{x})$.

The first three theorems of Section 3 are k-dimensional generalizations of theorems in Killeen, *et al.* (1972) for evaluating $m_i(\mathbf{x})$ from a sequence of density or probability functions. The last two theorems of Section 3 are concerned with the evaluation of $e(A)$ directly from a sequence of density or probability functions. In particular, they are applicable in cases where condition (B) does not hold.

2. Proof of theorems. The following two lemmas are immediate.

LEMMA 2.1. *If $B_r \in B^k$ for $r = 1, \cdots, m$, then $\bar{e}(\bigcup_{r=1}^m B_r) = \max\{\bar{e}(B_r): r = 1, \cdots, m\}$ and a similar identity holds for $\underline{e}(\cdot)$ and $e(\cdot)$ if the limits exist.*

LEMMA 2.2. *If $B_1, B_2 \in B^k$ with $B_1 \subset B_2$, then $\bar{e}(B_1) \leqq \bar{e}(B_2)$ and a similar inequality holds for $\underline{e}(\cdot)$ and $e(\cdot)$ if the limits exist.*

LEMMA 2.3. *Assume conditions* (A) *and* (B) *hold and that $A \in B^k$ with $\emptyset \neq A \subset \mathscr{cl}(A^\circ)$ and $A \subset Q_i(\mathbf{0})$ for some integer i, $1 \leqq i \leqq M$. Then* $\sup\{m_i(\mathbf{x}): \mathbf{x} \in A\} \leqq \underline{e}(A)$.

PROOF. Let $\mathbf{x} \in A$. If $\mathbf{x} \in A^\circ$ there exists $\varepsilon > 0$ such that $N_i(\mathbf{x}, \varepsilon) \subset A$. Then from conditions (A) and (B) and Lemma 2.2, $m_i(\mathbf{x}) = e(N_i(\mathbf{x}, \varepsilon)) \leqq \underline{e}(A)$. If $\mathbf{x} \notin A^\circ$ there exists a sequence $\{\mathbf{x}_m\}_{m=1}^\infty$ in A° converging to $\mathbf{x}$. Then $m_i(\mathbf{x}_m) \leqq \underline{e}(A)$ and with continuity, $m_i(\mathbf{x}) \leqq \underline{e}(A)$.

LEMMA 2.4. *If $\emptyset \neq A \subset Q_i(\mathbf{0})$ for some integer i, $1 \leqq i \leqq M$, and $\varepsilon > 0$, then there exists a finite number of points $\mathbf{x}_1, \cdots, \mathbf{x}_r \in Q_i(\mathbf{0})$ and $\mathbf{y}_1, \cdots, \mathbf{y}_r \in A$ such that*

(i) $||\mathbf{x}_j - \mathbf{y}_j|| < \varepsilon$ *for all* $j = 1, \cdots, r$,
(ii) $A \subset \bigcup_{j=1}^r Q_i(\mathbf{x}_j)$ *and*
(iii) *For some* j, $1 \leqq j \leqq r$, $\bar{e}(A) \leqq \bar{m}_i(\mathbf{x}_j)$.

PROOF. For parts (i) and (ii) only a brief outline of the proof will be given. It is enough to consider the case $i = 1$ and with suitable translation, if necessary, assume $\inf\{x_s: \mathbf{x} = (x_1, \cdots, x_k) \in A\} = 0$ for each $s = 1, \cdots, k$. The proof can be made by induction on k. As a starting point choose $\Delta < \varepsilon/k^{\frac{1}{2}}$ and consider "slices" of $Q_1(\mathbf{0})$ of the form $\{\mathbf{x}: (m-1)\Delta \leqq x_s < m\Delta\}$ for $m = 1, 2, \cdots$ and $s = 1, \cdots, k$. With the first point $\mathbf{x}_1$, suitably chosen on the boundary of $Q_1(\mathbf{0})$, note that $Q_1(\mathbf{0}) - Q_1(\mathbf{x}_1)$ can be viewed as a finite union of "slices" and since these "slices" are essentially $k-1$ dimensional, the inductive hypothesis can be applied to them.

Now (iii) follows from (i) and (ii) by noting that $A = \bigcup_{j=1}^{r} (A \cap Q_i(\mathbf{x}_j))$ and with Lemmas 2.1 and 2.2, $\bar{e}(A) = \max\{\bar{e}(A \cap Q_i(\mathbf{x}_j)) : j = 1, \cdots, r\} \leqq \max\{\bar{m}_i(\mathbf{x}_j) : j = 1, \cdots, r\}$.

PROOF OF THEOREM 1.1. $A \subset cl(A^\circ)$ implies $A \subset \bigcup_{i=1}^{M} cl(A^\circ \cap Q_i(\mathbf{0}))$ and with Lemmas 2.1 and 2.2

$$\text{(2.1)} \qquad \begin{aligned} &\max\{\underline{e}(A^\circ \cap Q_i(\mathbf{0})) : 1 \leqq i \leqq M\} \\ &\qquad \leqq \underline{e}(A) \leqq \bar{e}(A) \leqq \max\{\bar{e}(cl(A^\circ \cap Q_i(\mathbf{0}))) : 1 \leqq i \leqq M\}\,. \end{aligned}$$

Since $A^\circ \cap Q_i(\mathbf{0}) \subset cl(A^\circ \cap Q_i(\mathbf{0}))$, Lemma 2.3 implies

$$\text{(2.2)} \qquad \sup\{m_i(\mathbf{x}) : \mathbf{x} \in A^\circ \cap Q_i(\mathbf{0})\} \leqq \underline{e}(A^\circ \cap Q_i(\mathbf{0}))$$

for all $i = 1, \cdots, M$ for which $A^\circ \cap Q_i(\mathbf{0}) \neq \emptyset$.

For each fixed integer i, $1 \leqq i \leqq M$, for which $A^\circ \cap Q_i(\mathbf{0}) \neq \emptyset$, Lemma 2.4 implies that for all $m = 1, 2, \cdots$ there exists $\mathbf{x}_m \in Q_i(\mathbf{0})$ and $\mathbf{y}_m \in cl(A^\circ \cap Q_i(\mathbf{0}))$ such that $||\mathbf{x}_m - \mathbf{y}_m|| < 1/m$ and

$$\text{(2.3)} \qquad \bar{e}(cl(A^\circ \cap Q_i(\mathbf{0})) \leqq m_i(\mathbf{x}_m)\,.$$

The hypothesis implies that the sequence $\{\mathbf{y}_m\}$, and hence $\{\mathbf{x}_m\}$, is bounded and so there exists a subsequence (which we take to be the original sequence) such that $\mathbf{x}_m \to \mathbf{z}$ as $m \to \infty$ for some $\mathbf{z} \in cl(A^\circ \cap Q_i(\mathbf{0}))$.

With (2.3) and the continuity of $m_i(\mathbf{x})$,

$$\text{(2.4)} \qquad \begin{aligned} \bar{e}(cl(A^\circ \cap Q_i(\mathbf{0}))) &\leqq m_i(\mathbf{z}) \\ &\leqq \sup\{m_i(\mathbf{x}) : \mathbf{x} \in cl(A^\circ \cap Q_i(\mathbf{0}))\} \\ &= \sup\{m_i(\mathbf{x}) : \mathbf{x} \in A^\circ \cap Q_i(\mathbf{0})\}\,. \end{aligned}$$

The theorem then follows from (2.1), (2.2) and (2.4).

PROOF OF THEOREM 1.2. As in the previous proof we have (2.1), (2.2), and (2.3). For a given integer i, $1 \leqq i \leqq M$, if the sequence $\{\mathbf{x}_m\}$ is bounded then (2.4) follows as before. So assume that $\{\mathbf{x}_m\}$ is unbounded. Without loss of generality, assume that $i = 1$ and that the coordinate sequences $\{\mathbf{x}_{mj}\}$ are bounded for $1 \leqq j \leqq r$ and unbounded for $r + 1 \leqq j \leqq k$ for some integer r, $1 \leqq r \leqq k - 1$. Note that all k coordinate sequences of $\{\mathbf{x}_m\}$ cannot be unbounded, for then (2.2) and (2.3) contradict condition (B) if $A^\circ \cap Q_i(\mathbf{0}) \neq \emptyset$.

Let $a_j = \sup\{x_{mj} : m \geqq 1\}$ for $1 \leqq j \leqq r$ and $\mathbf{a} = (a_1, \cdots, a_r)$.

Assume that (2.4) is false; then there exists $\varepsilon > 0$ such that $\sup\{m_1(\mathbf{x}) : \mathbf{x} \in cl(A^\circ \cap Q_1(\mathbf{0}))\} \leqq \bar{e}(cl(A^\circ \cap Q_1(\mathbf{0}))) - \varepsilon$. Then $m_1(\mathbf{y}_m) \leqq \bar{e}(cl(A^\circ \cap Q_1(\mathbf{0}))) - \varepsilon$ and with (2.3) we have $m_1(\mathbf{x}_m) - m_1(\mathbf{y}_m) \geqq \varepsilon$ for all $m = 1, 2, \cdots$. Then for all

$\delta > 0$ and m sufficiently large (depending on δ), $m_1(a_1 - \delta, \cdots, a_r - \delta, x_{mr+1}, \cdots, x_{mk}) - m_1(a_1 + \delta, \cdots, a_r + \delta, y_{mr+1}, \cdots, y_{mk}) \geqq \varepsilon$. Now let $m \to \infty$ and with the monotonicity of $m_1(\mathbf{x})$ we have $g_1(\mathbf{a} - \boldsymbol{\delta}; 1, \cdots, r) - g_1(\mathbf{a} + \boldsymbol{\delta}; 1, \cdots, r) \geqq \varepsilon$, which contradicts condition (C). Hence (2.4) holds.

The argument of the previous paragraph requires $g_1(\mathbf{a}; 1, \cdots, r) > -\infty$. If $g_1(\mathbf{a}; 1, \cdots, r) = -\infty$, then (2.3) implies $\bar{e}(c\ell(A^\circ \cap Q_1(\mathbf{0}))) \leqq m_1(\mathbf{x}_m) \leqq m_1(a_1 - \delta, \cdots, a_r - \delta, x_{mr+1}, \cdots, x_{mk})$ for all $\delta > 0$ and m sufficiently large (depending on δ). Then $\bar{e}(c\ell(A^\circ \cap Q_1(\mathbf{0}))) \leqq g_1(\mathbf{a} - \boldsymbol{\delta}; 1, \cdots, r)$ and with continuity, $\bar{e}(c\ell(A^\circ \cap Q_1(\mathbf{0}))) = -\infty$. This with (2.2) implies $\sup\{m_1(\mathbf{x}) : \mathbf{x} \in A^\circ \cap Q_1(\mathbf{0})\} = -\infty$ and so (2.4) holds.

The theorem follows from (2.1), (2.2) and (2.4).

Proof of Lemma 1.1. For a given integer i, $1 \leqq i \leqq M$, and $\varepsilon > 0$ let $\mathbf{x} \in Q_i(\mathbf{0})$ with $m_i(\mathbf{x}) > -\infty$. Let $\mathbf{x}, \mathbf{x}_1, \cdots, \mathbf{x}_{M-1}$ denote the "corner points" of $N_i(\mathbf{x}, \varepsilon)$. For all $j = 1, \cdots, M-1$, $\mathbf{x}_j \in Q_i(\mathbf{x})$ and $\mathbf{x}_j \neq \mathbf{x}$ so by hypothesis $m_i(\mathbf{x}_j) < m_i(\mathbf{x})$. Now $Q_i(\mathbf{x}) = N_i(\mathbf{x}, \varepsilon) \cup (\bigcup_{j=1}^{M-1} Q_i(\mathbf{x}_j))$ and by Lemma 2.1, $m_i(\mathbf{x}) = \max\{e(N_i(\mathbf{x}, \varepsilon)), m_i(\mathbf{x}_1), \cdots, m_i(\mathbf{x}_{M-1})\} = e(N_i(\mathbf{x}, \varepsilon))$.

Proof of Lemma 1.2. The hypothesis and the monotonicity of $m_i(\mathbf{x})$ imply that the functions $g_i(\mathbf{y}; j_1, \cdots, j_r) = -\infty$ and hence are continuous.

Lemma 2.5. *Assume for all $A \in B^k$ with $\varnothing \neq A \subset c\ell(A^\circ)$ and $A \subset Q_i(\mathbf{0})$ for some integer i, $1 \leqq i \leqq M$, that $\sup\{\underline{m}_i(\mathbf{x}); \mathbf{x} \in A\} \leqq \underline{e}(A)$. For all integers i, $1 \leqq i \leqq M$, and $\varepsilon > 0$ if $\mathbf{x} \in Q_i(\mathbf{0})$ then $\underline{m}_i(\mathbf{x}) = \underline{e}(N_i(\mathbf{x}, \varepsilon))$.*

Proof. If we assume the conclusion false, there exists i, $1 \leqq i \leqq M$, $\varepsilon > 0$ and $\mathbf{x} \in Q_i(\mathbf{0})$ such that $\underline{m}_i(\mathbf{x}) > \underline{e}(N_i(\mathbf{x}, \varepsilon))$. But this, with $A = N_i(\mathbf{x}, \varepsilon)$, contradicts the hypotheis.

Example. Suppose for each $n = 1, 2, \cdots$ that P_n is absolutely continuous bivariate distribution with density function

$$f_n(\mathbf{x}) = n^2 x_1 \exp\{-nx_1(1 + x_2)\}, \qquad x_1, x_2 \geqq 0$$

and zero otherwise. Then for $\mathbf{x} \in Q_1(\mathbf{0})$, $m_1(\mathbf{x}) = -x_1(1 + x_2)$. Conditions (A) and (B) hold but not condition (C) since $g_1(y; 1)$ is discontinuous at $y = 0$. However, with event $A = \{\mathbf{x} \in R^2 : x_1 x_2 \geqq 1\}$, straightforward calculations show $e(A) = -1 = \sup\{m_1(\mathbf{x}) : \mathbf{x} \in A\}$ so that the conclusion of Theorem 1.2 holds.

Suppose we modify the distribution P_n by removing a probability mass $\exp\{-n/2\}$ from near $\mathbf{0}$ and placing it at the point $(1/n, n)$. Then $m_1(\mathbf{x})$ remains unchanged and $e(A) = \max\{-1, -\frac{1}{2}\} = -\frac{1}{2}$. Hence conditions (A) and (B) hold but not condition (C) and the conclusion of Theorem 1.2 is false.

3. Density and probability function results. The first three theorems of this section are k-dimensional versions of theorems in Killeen, *et al.* (1972) and since the proofs are somewhat similar to those in this reference they will be omitted.

Theorem 3.1. *Suppose P_n is an absolutely continuous distribution on R^k with* *density function* $f_n(\mathbf{x})$ *for* $n = 1, 2, \cdots$. *For a given integer* i, $1 \leqq i \leqq M$, *let* $\mathbf{a} \in Q_i(\mathbf{0})$ *and assume*:

(i) *there exists an integer* N *such that for* $n \geqq N$, $f_n(\mathbf{x})$ *is non increasing on* $Q_i(\mathbf{a})$; *that is*, $\mathbf{x} \in Q_i(\mathbf{a})$ *and* $\mathbf{x}' \in Q_i(\mathbf{x})$ *imply* $f_n(\mathbf{x}') \leqq f_n(\mathbf{x})$,

(ii) *there exists a sequence* $\{\boldsymbol{\delta}_n\}$ *in* $Q_i(\mathbf{0})$ *such that for* $n = 1, 2, \cdots, \boldsymbol{\delta}_n$ *has coordinates identical in absolute value, say* δ_n, *with* $(1/n) \log \delta_n = o(1)$ *and*

$$(1/n) \log [f_n(\mathbf{a} + \boldsymbol{\delta}_n)/f_n(\mathbf{a})] = o(1) \qquad \text{as } n \to \infty, \text{ and}$$

(iii) *there exists a non negative sequence* $\{\gamma_n\}$ *such that* $(1/n) \log \gamma_n = o(1)$ *as* $n \to \infty$ *and*

$$\limsup_{n\to\infty} (1/n) \log [P_n(Q_i(\mathbf{a}) - N_i(\mathbf{a}, \gamma_n))/f_n(\mathbf{a})] \leqq 0 .$$

Then

$$(3.1) \qquad (1/n) \log P_n[Q_i(\mathbf{a})] - (1/n) \log f_n(\mathbf{a}) = o(1) \qquad \text{as } n \to \infty .$$

THEOREM 3.2. *For each* $n = 1, 2, \cdots$ *suppose that* P_n *is a probability distribution on* N^k, *the points in* R^k *with integer coordinates, and has probability function* $f_n(\mathbf{x})$. *If for a given integer* i, $1 \leqq i \leqq M$, $\mathbf{a} \in Q_i(\mathbf{0}) \cap N^k$ *and conditions* (i) *and* (iii) *of Theorem* 3.1 *hold* ((i) *holding on* $Q_i(\mathbf{a}) \cap N^k$), *then* (3.1) *holds.*

Condition (iii) is readily seen to be necessary for (3.1) but is sometimes awkward to verify in particular cases. The following theorem gives a sufficient condition in terms of moment-generating functions.

THEOREM 3.3. *For* $n = 1, 2, \cdots$ *let* $\mathbf{X}_n = (X_{n1}, \cdots, X_{nk})$ *denote a random vector with distribution* P_n. *For a given integer* i, $1 \leqq i \leqq M$, *suppose there exists* $\mathbf{t} = (t_1, \cdots, t_k) \in (Q_i(\mathbf{0}))^0$ *and constant* $R < \infty$ *such that the joint moment-generating function* $E[\exp\{\sum_{j=1}^k t_j X_{nj}\}] < R$ *for all* $n = 1, 2, \cdots$. *If* $\{\gamma_n\}$ *is a nonnegative sequence with* $\gamma_n/n \to \infty$, *then for* $\mathbf{a} \in Q_i(\mathbf{0})$,

$$\lim_{n\to\infty} (1/n) \log P_n[Q_i(\mathbf{a}) - N_i(\mathbf{a}, \gamma_n)] = -\infty .$$

If, in addition, $(1/n) \log \gamma_n = o(1)$ *as* $n \to \infty$ *and* $\lim_{n\to\infty} (1/n) \log f_n(\mathbf{a}) = c > -\infty$, *then condition* (iii) *holds.*

For $\mathbf{x} \in R^k$ and $\varepsilon > 0$ let $N(\mathbf{x}, \varepsilon) = \{\mathbf{y} \in R^k : |x_j - y_j| < \varepsilon, j = 1, \cdots, k\}$.

LEMMA 3.1. *Assume* $B \subset R^k$ *has compact closure. Then for all* $\varepsilon > 0$ *there exists* $\mathbf{z} \in B$ *such that* $\bar{e}(B) = \bar{e}(B \cap N(\mathbf{z}, \varepsilon))$.

PROOF. $\{N(\mathbf{y}, \varepsilon) : \mathbf{y} \in B\}$ is an open cover of $c\ell(B)$. With compactness there is a finite subcover of $c\ell(B)$ and hence of B, say $\{N(\mathbf{y}_j, \varepsilon) : \mathbf{y}_j \in B, j = 1, \cdots, r\}$. Then from Lemma 2.1, $\bar{e}(B) = \bar{e}(\bigcup_{j=1}^r (B \cap N(\mathbf{y}_j, \varepsilon))) = \bar{e}(B \cap N(\mathbf{z}, \varepsilon))$ where $\mathbf{z} = \mathbf{y}_j$ for some $j = 1, \cdots, r$.

THEOREM 3.4. *Suppose* P_n *is an absolutely continuous distribution on* R^k *with density function* $f_n(\mathbf{x})$ *for* $n = 1, 2, \cdots$. *Suppose* $A \in B^k$ *with* $\emptyset \neq A \subset c\ell(A^\circ)$ *and the following conditions hold*:

(i) $\lim_{n\to\infty}(1/n)\log f_n(\mathbf{x}) = g(\mathbf{x})$ *exists and is continuous on* $c\ell(A)$,

(ii) *For each* $\mathbf{x} \in A^\circ$, *there exists sequences of nonnegative reals* $\{\delta_n(\mathbf{x})\}$ *and* $\{r_n(\mathbf{x})\}$ *and a positive integer* $M(\mathbf{x})$ *suchthat* $\lim_{n\to\infty}\delta_n(\mathbf{x}) = 0$, $\liminf_{n\to\infty}(1/n)\log r_n(\mathbf{x}) \geqq 0$ *and* $P_n[N(\mathbf{x}, \delta_n(\mathbf{x}))] \geqq f_n(\mathbf{x})r_n(\mathbf{x})$ *for all* $n \geqq M(\mathbf{x})$,

(iii) *For each* $\mathbf{x} \in c\ell(A)$ *and* $\varepsilon > 0$ *sufficiently small there exists* $\mathbf{y}(\mathbf{x}, \varepsilon) \in c\ell(N(\mathbf{x}, \varepsilon))$, *a sequence of nonnegative reals* $\{s_n(\mathbf{x}, \varepsilon)\}$ *and a positive integer* $M(\mathbf{x}, \varepsilon)$ *such that* $\limsup_{n\to\infty}(1/n)\log s_n(\mathbf{x}, \varepsilon) \leqq 0$ *and* $P_n[N(\mathbf{x}, \varepsilon)] \leqq f_n(\mathbf{y}(\mathbf{x}, \varepsilon))s_n(\mathbf{x}, \varepsilon)$ *for all* $n \geqq M(\mathbf{x}, \varepsilon)$ *and*

(iv) $\inf\{\bar{e}(A - N(\mathbf{0}, \gamma)) : \gamma > 0\} < \sup\{g(\mathbf{x}) : \mathbf{x} \in A\}$. *Then* $e(A) = \sup\{g(\mathbf{x}) : \mathbf{x} \in A\}$.

Proof. Let $\mathbf{x} \in A^\circ$. From condition (ii), $P_n(A) \geqq P_n[N(\mathbf{x}, \delta_n(\mathbf{x}))] \geqq f_n(\mathbf{x})r_n(\mathbf{x})$ for n sufficiently large. Then $\underline{e}(A) \geqq g(\mathbf{x})$. Since $A \subset c\ell(A^\circ)$ and $g(\mathbf{x})$ is continuous

$$\underline{e}(A) \geqq \sup\{g(\mathbf{x}) : \mathbf{x} \in A\}. \tag{3.2}$$

From Lemma 2.1, for all $\gamma > 0$ $\bar{e}(A) = \max\{\bar{e}(A \cap N(\mathbf{0}, \gamma)), \bar{e}(A - N(\mathbf{0}, \gamma))\}$. From (3.2) $\bar{e}(A) \geqq \sup\{g(\mathbf{x}) : \mathbf{x} \in A\}$ and with condition (iv), there is a $\gamma_0 > 0$ sufficiently large so that

$$\bar{e}(A) = \bar{e}(A \cap N(\mathbf{0}, \gamma_0)). \tag{3.3}$$

Let $\{\varepsilon_m\}$ be a sequence of positive reals with $\lim_{m\to\infty}\varepsilon_m = 0$. Define a sequence $\{\mathbf{x}_m\}$ in A inductively as follows: for each $m = 1, 2, \cdots$ apply Lemma 3.1 with $\varepsilon = \varepsilon_m$, $B = A \cap N(\mathbf{x}_{m-1}, \varepsilon_{m-1})$ and let $\mathbf{x} = \mathbf{z}$ of the Lemma. Use $\mathbf{x}_0 = \mathbf{0}$ and $\varepsilon_0 = \gamma_0$. Then

$$\bar{e}(A) = \bar{e}(A \cap N(\mathbf{x}_m, \varepsilon_m)) \leqq \bar{e}(N(\mathbf{x}_m, \varepsilon_m)) \tag{3.4}$$

and $\mathbf{x}_m \in A \cap N(\mathbf{x}_{m-1}, \varepsilon_{m-1})$ for $m = 1, 2, \cdots$. The sequence $\{\mathbf{x}_m\}$ converges, say to $\mathbf{x}_0 \in c\ell(A)$.

Using (3.4) and condition (iii), for all $\varepsilon > 0$, $\bar{e}(A) \leqq \bar{e}(N(\mathbf{x}_0, \varepsilon)) \leqq g(\mathbf{y}(\mathbf{x}_0, \varepsilon))$. Then the continuity of $g(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}_0$ implies

$$\begin{aligned}\bar{e}(A) \leqq g(\mathbf{x}_0) &\leqq \sup\{g(\mathbf{x}) : \mathbf{x} \in c\ell(A)\} \\ &= \sup\{g(\mathbf{x}) : \mathbf{x} \in A\}.\end{aligned} \tag{3.5}$$

The theorem follows from (3.2) and (3.5).

Corollary 3.1. *Condition* (iv) *of Theorem* 3.4 *can be omitted if A is bounded.*

Proof. Condition (iv) was used in the proof of Theorem 3.4 to establish (3.4) and (3.4) follows directly if A is bounded.

4. Applications.

Example (a). Functions of sequences of independent random variables. For $n = 1, 2, \cdots$ let $\mathbf{X}_n = (X_{n1}, \cdots, X_{nk})$ be a random vector with distribution P_n and mutually independent components. For each $j = 1, \cdots, k$ let

$$\phi_j(x) = \lim_{n\to\infty}(1/n)\log P_n[X_{nj} \in A_x]$$

where $A_x = [x, \infty)$ $((-\infty, x])$ if $x \geqq 0$ $(x < 0)$, assuming the limits exist. Assume that $\phi_j(x)$ is continuous and strictly increasing (decreasing) on $(-\infty, 0]$ $([0, \infty))$. For $i = 1, \cdots, M$ it follows that if $\mathbf{x} = (x_1, \cdots, x_k) \in Q_i(\mathbf{0})$ then $m_i(\mathbf{x}) = \sum_{j=1}^k \phi_j(x_j)$. Conditions (A), (B) (from Lemma 1.1) and (C) hold as does the hypothesis of Corollary 1.1. Then

$$e(A) = \sup \{\textstyle\sum_{j=1}^k \phi_j(x_j) : \mathbf{x} \in A\}$$

for all $A \in B^k$ with $\emptyset \neq A \subset c\ell(A^\circ)$. This yields an expression for the probability of a large deviation for a sequence of random variables $\{T_n\}$ with $T_n = T(\mathbf{X}_n)$ and $A = \{\mathbf{x} \in R^k : T(\mathbf{x}) \geqq t\}$.

EXAMPLE (b). Functions of a finite number of order statistics. For each $n = 1, 2, \cdots$ let $Y_{n1} \leqq \cdots \leqq Y_{nn}$ denote the order statistics of an independent sample of a size n from a population with an absolutely continuous distribution with cdf $F(x)$ and density $f(x)$. Assume that $f(x) > 0$ for $a < x < b$ and zero otherwise for $-\infty \leqq a < b \leqq \infty$. Assume $F(x)$ is differentiable and $F'(x) = f(x)$ for all x.

Let an integer k and constants $0 = p_0 < p_1 < \cdots < p_k < p_{k+1} = 1$ be given. For all $n = 1, 2, \cdots$ let $n_j = [np_j] + 1$ and $X_{nj} = Y_{nn_j}$ for $j = 1, \cdots, k$. Let $n_0 = 0$ and $n_{k+1} = n + 1$. For n sufficiently large, $\mathbf{X}_n = (X_{n1}, \cdots, X_{nk})$ has the familiar order statistic density.

The sequence $\{\mathbf{X}_n\}$ satisfies the conditions of Theorem 3.4 (excepting the initial elements where the density may not exist) for any $A \in B^k$ with $\emptyset \neq A \subset c\ell(A^\circ)$. Condition (i) is clear with

$$g(\mathbf{x}) = \textstyle\sum_{j=1}^{k+1} (p_j - p_{j-1}) \log [F(x_j) - F(x_{j-1}))/(p_j - p_{j-1})] \,. \tag{4.1}$$

The verification of conditions (ii) and (iii) is somewhat routine and will be omitted. For condition (iv), if $\gamma > 0$, $P_n(A - N(\mathbf{0}, \gamma)) \leqq P_n(X_{n1} < -\gamma \text{ or } X_{nk} > \gamma)$. Using $P_n(X_{nk} > \gamma) \leqq \binom{n}{n_k}[1 - F(\gamma)]^{n-n_k}$ and a similar bound for $P_n(X_{n1} < -\gamma)$, it follows that $\inf \{\bar{e}(A - N(\mathbf{0}, \gamma)) : \gamma > 0\} = -\infty$ and hence condition (iv) holds.

Finally, with the conditions of Theorem 3.4 holding, for any $A \in B^k$ with $\emptyset \neq A \subset c\ell(A^\circ)$ we have

$$\lim_{n\to\infty} (1/n) \log P_n(A) = \sup \{g(\mathbf{x}) : \mathbf{x} \in A\} \tag{4.2}$$

where $g(\mathbf{x})$ is given by (4.1). This result can be applied to give expressions for probabilities of large deviations for random variables which are functions of a finite number of order statistics, such as linear combinations ($A = \{\mathbf{x} : a_1 x_1 + \cdots + a_k x_k \geqq c\}$), quadratic forms, etc.

Witting (1959) has considered a χ^2 goodness-of-fit test with cell boundaries equal to selected order statistics for testing the null hypothesis that a given $F(x)$ is the population cdf. The statistic for a sample of size n is T_n where

$$T_n/n = \delta(v(\mathbf{X}_n), \mathbf{q}_n)$$

where $v(\mathbf{x}) = (v_1(\mathbf{x}), \cdots, v_{k+1}(\mathbf{x}))$ with $v_j(\mathbf{x}) = F(x_j) - F(x_{j-1})$ for $j = 1, \cdots, k+1$ $(x_0 = a, x_{k+1} = b)$, $\mathbf{q}_n = (q_{n1}, \cdots, q_{nk+1})$ with $q_{nj} = (n_j - n_{j-1})/(n+1)$ for $j = 1, \cdots, k+1$ and $\delta(\mathbf{v}, \mathbf{w}) = \sum_{j=1}^{k+1} (v_j - w_j)^2/w_j$. Note that $\mathbf{q}_n \to \mathbf{q}$ as $n \to \infty$ where $q_j = p_j - p_{j-1}$.

To determine the exact slope of the sequence $\{T_n\}$ let $A(t) = \{\mathbf{v} : \delta(\mathbf{v}, \mathbf{q}) \geqq t\}$ and $A_n(t) = \{\mathbf{v} : \delta(\mathbf{v}, \mathbf{q}_n) \geqq t\}$ for $n = 1, 2, \cdots$. From (4.2),

$$\begin{aligned}\phi(t) &= \lim_{n\to\infty} (1/n) \log P_n[v(\mathbf{X}_n) \in A(t)] \\ &= \sup\{-K(\mathbf{q}, \mathbf{v}) : \mathbf{v} \in A(t)\}\end{aligned}$$

where $K(\mathbf{q}, \mathbf{v}) = \sum_{j=1}^{k+1} q_j \log(q_j/v_j)$ is the Kullback–Leibler information number. Since $\phi(t)$ is continuous and for all $\varepsilon > 0$, $A(t+\varepsilon) \subset A_n(t) \subset A(t-\varepsilon)$ for n sufficiently large,

$$\lim_{n\to\infty} (1/n) \log P_n[T_n \geqq nt] = \lim_{n\to\infty} (1/n) \log P_n[v(\mathbf{X}_n) \in A_n(t)] = \phi(t)\,.$$

For an alternative cdf $G(x) \neq F(x)$, $(T_n/n) \to \delta(\mathbf{s}, \mathbf{q})$ a.s. if $b_1, \cdots, b_k$ are unique, where $s_j = F(b_j) - F(b_{j-1})$ and $G(b_j) = p_j$ for $j = 1, \cdots, k+1$ $(b_0 = a, b_{k+1} = b)$. Hence from Theorem 7.2 of Bahadur (1971), the exact slope of the sequence $\{T_n\}$ is

$$-2\phi(\delta(\mathbf{s}, \mathbf{q})) = 2 \inf\{K(\mathbf{q}, \mathbf{v}) : \mathbf{v} \in A(\delta(\mathbf{s}, \mathbf{q}))\}\,. \tag{4.3}$$

The usual χ^2 goodness-of-fit statistic which is comparable to T_n has cell boundaries a_j with $F(a_j) = p_j$ for $j = 1, \cdots, k$ $(a_0 = a, a_{k+1} = b)$. The exact slope of this sequence is (see Bahadur (1971), page 31)

$$2 \inf\{K(\mathbf{v}, \mathbf{q}) : \mathbf{v} \in A(\delta(\mathbf{r}, \mathbf{q}))\} \tag{4.4}$$

where $\mathbf{r} = (r_1, \cdots, r_{k+1})$ with $r_j = G(a_j) - G(a_{j-1})$. Some rough calculations indicate that (4.3) is larger than (4.4) for some alternatives G and that the reverse can hold for other G. A precise comparison appears difficult.

REFERENCES

[1] Bahadur, R. R. (1971). *Some Limit Theorems in Statistics*. Society for Industrial and Applied Mathematics, Philadelphia.

[2] Killeen, T. J., Hettmansperger, T. P. and Sievers, G. L. (1972). An elementary theorem on the probability of large deviations. *Ann. Math. Statist.* **43** 181–192.

[3] Sievers, G. L. (1969). On the probability of large deviations and exact slopes. *Ann. Math. Statist.* **40** 1908–1921.

[4] Witting, H. (1959). On a χ^2-test with cells determined by order statistics. *Ann. Math. Statist.* **30** 1271.

Department of Mathematics
Western Michigan University
Kalamazoo, Michigan 49001