Z. Wahrscheinlichkeitstheorie verw. Geb. 18, 281–297 (1971)

Limiting Convex Hulls of Samples: Theory and Function Space Examples*

Lloyd Fisher

Introduction

Suppose that $X_1, X_2, \ldots$ is a sequence of independent identically distributed random variables. Let $m_n = \min\{X_1, \ldots, X_n\}$ and $M_n = \max\{X_1, \ldots, X_n\}$. Let $R_n = [m_n, M_n]$, the smallest interval containing the sample $\{X_1, \ldots, X_n\}$, be called the range of the sample. The subject of the range of the sample (or equivalently the extreme values of the sample) has been extensively studied, the most fundamental paper being that of Gnedenko [8].

The motivation of the present paper was to find a suitable generalization of the range to random variables taking values in a higher dimensional space. If the X_i's are random vectors in a linear space one possible generalization is the convex hull of the sample $\{X_1, \ldots, X_n\}$ which reduces to R_n if the space is the real line. It is this quantity that we shall consider here. This generalization was suggested by Professor Lamperti.

The convex hull of a sample has been the object of previous investigations. Rényi and Sulanke [11] have considered the problem of the asymptotic behavior of the expected area, perimeter, and number of vertices of the convex hull of i.i.d. points in the plane. Efron [3] has considered similar problems in two and three dimensions, but has emphasized fixed values of the sample size. He also considers the probability content of the sample. Geffroy [6, 7] shows that if the sample points come from a k dimensional normal distribution that the convex hull is "almost ellipsoidal" in shape as the sample size becomes large.

In this paper the question of whether or not the convex hull has a limiting shape shall be considered. The mathematical formulation is given below.

1. Definitions

Let $X_1, \ldots, X_n, \ldots$ be a sequence of independent, identically distributed Borel random vectors taking values in a separable Banach space B. Let $S_n = \{X_1, \ldots, X_n\}$ be the first n sample points. For any bounded set $A \subseteq B$ let $|A| = \max\{\|X\| : X \in A\}$. Let $A^\varepsilon = \{X\colon$ there exists $Y \in A$, $\|Y - X\| < \varepsilon\}$. For two bounded subsets A and C of B let $d(A, C) = \inf\{\varepsilon > 0\colon A^\varepsilon \supseteq C$ and $C^\varepsilon \supseteq A\}$. Let R_n be the convex hull of S_n (the range of the sample). For any set $A \subseteq B$ and real number c let $cA = \{cX\colon X \in A\}$.

Since we will be interested only in the shape of R_n and not its size we shall change the scale until R_n touches the surface of the unit ball, but is contained in the

* This research was supported by an NSF co-operative fellowship at Dartmouth College and NSF Grant No. GP-7519 at the University of Washington.

unit ball. Thus, let

$$NR_n = R_n/|R_n| \quad \text{if } |R_n|>0 \quad \text{otherwise let } NR_n = \mathbf{0}.$$

The N denotes normalized. Let $NS_n = S_n/|S_n|$ if $|S_n|>0$ otherwise $NS_n = \mathbf{0}$.

Let T be a closed bounded subset of B and Y_n a sequence of random subsets of B that are bounded with probability one. Then we write:

$$\lim_{n\to\infty} Y_n = T \text{ i.p.} \quad \text{or} \quad Y_n \to T \text{ i.p.}$$

if for all $\varepsilon>0$, $P(d(Y_n, T)>\varepsilon)\to 0$. Similarly, $\lim_{n\to\infty} Y_n = T$ a.s. or $Y_n\to T$ a.s. if $P(d(Y_n, T)\to 0)=1$. In the sequel Y_n shall be NR_n or NS_n. The possible limits T are required to be closed in order to give a unique limit.

For each $X\in B$ let $X^\varepsilon = \{X\}^\varepsilon$, that is, the ε-neighborhood of X. In the following w denotes weak since a weaker type of convergence is being considered. We write

$$\operatorname*{limit}_{n\to\infty} Y_n = T \text{ w.i.p.} \quad \text{or} \quad Y_n \to T \text{ w.i.p.}$$

if $X\in T$, $\varepsilon>0$ implies $P(Y_n \cap X^\varepsilon \neq \emptyset)\to 1$ and $X\notin T$ implies there exists $\varepsilon>0$ such that $P(Y_n\cap X^\varepsilon = \emptyset)\to 1$. Similarly,

$$\operatorname*{limit}_{n\to\infty} Y_n = T \text{ w.a.s.} \quad \text{or} \quad Y_n \to T \text{ w.a.s.}$$

if $X\in T$, $\varepsilon>0$ implies $\lim_{n\to\infty} P(Y_m \cap X^\varepsilon \neq \emptyset, m\geqq n)=1$ and $X\notin T$ implies there exists $\varepsilon>0$ such that $\lim_{n\to\infty} P(Y_m\cap X^\varepsilon = \emptyset, m\geqq n)=1$.

2. Elementary Consequences of the Definitions

The proof the following easy lemma will be omitted.

Lemma 2.1. *If $Z_n = NR_n$ or NS_n then*

(a) *$Z_n \to T$ a.s. implies $Z_n\to T$ i.p., w.a.s.*

(d) *$Z_n\to T$ i.p. implies $Z_n\to T$ w.i.p.*

(c) *$Z_n\to T$ w.a.s. implies $Z_n \to T$ w.i.p.*

(d) *$NS_n\to T$ i.p. (a.s.) and T closed and convex implies $NR_n\to T$ i.p. (a.s.)*

Lemma 2.2. *Let $Z_n = NR_n$ or NS_n then if $Z_n\to T$ i.p. or a.s., T is a compact subset of B.*

Proof. Since the limit is a closed set by definition to prove the lemma it is necessary to show that the set is totally bounded. Let $\varepsilon>0$ be given, choose an $N(\varepsilon)$ such that $n\geqq N(\varepsilon)$ implies $P(d(Z_n, T)<\varepsilon/2)\geqq \frac{1}{2}$. Choose any configuration of Z_n such that $Z_n^{\varepsilon/2}\geqq T$. The convex hull of Z_n is a polytope with n-vertices and since it is a bounded, closed subset of a finite dimensional subspace it is compact. Cover the convex hull with a finite number of $\varepsilon/2$ spheres. Using the same centers but taking ε spheres we have a finite ε covering of T.

Corollary 2.3. *If $T\subseteq B$ has an interior and B is infinite dimensional then T cannot be a limit i.p. or a.s. of NR_n or NS_n.*

Example. Pick a countable dense subset Y_i of the unit ball of l_2. Let T be the unit ball and let each X_i have a discrete distribution which puts positive measure on each Y_i and is concentrated on Y_i. Then $NR_n \to T$ w.a.s. and $NS_n \to T$ w.a.s. but no limit exists i.p. or a.s.

Theorem 2.4. *In order that $NR_n \to T$ i.p. (a.s.) it is necessary and sufficient that:*

(a) *T is a compact, convex subset of B.*

(b) *$P(T^\varepsilon \supseteq NR_n) \to 1$ for each $\varepsilon > 0$. $\big(\lim_{n\to\infty} P(T^\varepsilon \supseteq NR_m, m \geqq n) = 1.\big)$*

(c) *If X is an extreme point of T then for each $\varepsilon > 0$, $P(NS_n \cap X^\varepsilon \neq \emptyset) \to 1$. $\big(\lim_{n\to\infty} P(NS_m \cap X^\varepsilon \neq \emptyset, m \geqq n) \to 1.\big)$*

Proof. Let CH be the operator that takes bounded subsets of B into the closure of their convex hulls.

Since NR_n is convex it is clear by Lemma 2.2 that if a limit is to exist (a) must be satisfied. Further, for any limit T we have $P\big(d(T, NS_n) < \varepsilon\big) \to 1$ so that (b) must also hold.

Only if: Let $NS_n \to T$ i.p. Let X be an extreme point of T. Then $X \notin CH(T - X^\varepsilon)$ for each $\varepsilon > 0$ ([10], p. 132). Let $f(\delta) = d\big(CH(T^\delta - X^\varepsilon), X\big)$. We show that $f(\delta) > 0$ for some $\delta > 0$. Suppose not then we may find a sequence of triples $(X_1^n, X_2^n, \lambda_n)$ where X_1^n and $X_2^n \in T - X^\varepsilon$, $0 \leqq \lambda_n \leqq 1$ and $\|X - [\lambda_n X_1^n + (1-\lambda_n) X_2^n]\| < 1/n$. Since $T - X^\varepsilon$ is compact (being a closed subset of T) we may without loss of generality (by taking appropriate subsequences) assume that $X_1^n \to X_1 \in T - X^\varepsilon$, $X_2^n \to X_2 \in T - X^\varepsilon$ and $\lambda_n \to \lambda$. Then,

$$\begin{aligned} \|X - [\lambda X_1 + (1-\lambda) X_2]\| &\leqq \|X - [\lambda_n X_1^n + (1-\lambda_n) X_2^n]\| \\ &\quad + \|[\lambda_n - \lambda]\, X_1^n\| + \|[\lambda_n - \lambda]\, X_2^n\| \\ &\quad + \|\lambda [X_1^n - X_1]\| + \|(1-\lambda)[X_2^n - X_2]\| \end{aligned}$$

and as the right hand side may be made as small as desired, $X = X_1 + (1-\lambda) X_2$ contradicting $X \notin CH(T - X^\varepsilon)$. Let $f(\delta) > 0$. Since $P(NS_n \subseteq T^\delta) \to 1$, if $P(NS_n \cap X^\varepsilon \neq \emptyset) \not\to 1$ then by the above $P(\{X\} \subseteq NR_n) \not\to 1$ since whenever $NR_n = CH(NS_n) \subseteq CH(T^\delta - X^\varepsilon)$ then $0 < f(\delta) < d(X, NR_n) \leqq d(T, NR_n)$.

If: Assume (a), (b) and (c) note that $P\big(d(NR_n, T) \leqq \varepsilon\big) \to 1$ iff $P(NR_n \subseteq T^\varepsilon) \to 1$ and $P(T \subseteq NR_n^\varepsilon) \to 1$. Since $P(NR_n \subseteq T^\varepsilon) \to 1$ by (b) to show that $NR_n \to T$ i.p. we need only show $P(T \subseteq NR_n^\varepsilon) \to 1$. Cover the closure of the extreme points of T by a finite number of $\varepsilon/3$ spheres (which may be done since a closed subset of a compact set is compact). The probability that all the spheres have points of NS_n approaches one by (c). By the Krein-Millman Theorem ([10], p. 131), NR_n has points within $2\varepsilon/3$ of each element of T when the above holds. Thus, $P(NR_n^\varepsilon \supseteq T) \to 1$. The theorem for $NR_n \to T$ a.s. is proved similarly concluding the proof.

Let μ be the distribution of the X_i, that is, for each Borel set $A \subseteq B$ we define $\mu(A) = P(X_i \in A)$. It is clear that if $\mu(\{0\}) \neq 1$ then $NR_n \to T$ i.p. or a.s. or $NS_n \to T$ i.p. or a.s. implies $|T| = 1$. However, in the case of weak limits this need not be true.

Example. Let $B = l_2$. Let $\{U_i\}_{i=1}^\infty$ be an orthonormal basis for l_2. Let μ be discrete on $c_i U_i$ where c_i is a sequence of real numbers. Let $p_i = \mu(\{c_i U_i\})$.

If $c_i = i$ and $p_1 = 1 - e^{-1}$ and for $i > 1$, $p_i = e^{-i^2} - e^{-(i-1)^2}$ then $|S_n|/(\log n)^{\frac{1}{2}} \to 1$ a.s. which implies $NR_n \to \{\mathbf{0}\}$ w.a.s.

If B is finite dimensional then the following lemma shows that if $NR_n \to T$ w.i.p. then $|T| = 1$.

Lemma 2.5. *Let B be finite dimensional then*

(a) *$NR_n \to T$ i.p. (a.s.) iff $NR_n \to T$ w.i.p. (w.a.s.).*

(b) *$NS_n \to T$ i.p. (a.s.) iff $NS_n \to T$ w.i.p. (w.a.s.).*

Proof. The only if portion of the lemma is contained in Lemma 2.1. We prove only (a) the proof of (b) being similar.

Suppose that $NR_n \to T$ w.i.p. Let $X, Y \in T$. If NR_n has points within ε of X and Y then NR_n (being convex) has points within ε of $\lambda X + (1-\lambda) Y$ where $0 \leqq \lambda \leqq 1$. Thus, T is convex. T is closed since if every neighborhood of X has points of T this implies that the probability that each neighborhood has points of NR_n approaches one and $X \in T$. Since B is finite dimensional the unit sphere is compact and thus T is a convex compact set (satisfying (a) of Theorem 2.4).

If $X \notin T$ there exists a sphere $S(X)$ about X such that $P(NR_n \cap S(X) = \emptyset) \to 1$. Let S be the unit ball. For a given $\varepsilon > 0$ cover $S - T^\varepsilon$ with a finite number of $S(X)$'s. Since $|NR_n| \leqq 1$ it follows that $P(NR_n \subseteq T^\varepsilon) \to 1$ and condition (b) of Theorem 2.4 is satisfied.

As in the proof of Theorem 2.4 NR_n has points in each neighborhood of each extreme point X of T with a probability approaching one and (b) of Theorem 2.4 is satisfied so that it follows that (c) of Theorem 2.4 is satisfied and by Theorem 2.4 $NR_n \to T$ i.p. the a.s. statement follows similarly concluding the proof.

The following proposition shows that if μ has bounded support the problem under consideration is trivial.

Proposition 2.6. *Let μ have bounded support.*

(a) *If $\mu(\{\mathbf{0}\}) = 1$, $NS_n \to \{\mathbf{0}\}$ a.s.*

(b) *Let $\mu(\{\mathbf{0}\}) \neq 1$, if μ has compact support C then*

$$NS_n \to C/|C| \text{ a.s.}$$

(c) *If μ has noncompact support C then*

$$NS_n \to C/|C| \text{ w.a.s.}$$

but NR_n does not have a limit i.p.

Proof. (a) is clear. Under either (b) or (c) it is clear that $|S_n| \to |C|$ a.s. From this it is clear that $P((C/|C|)^\varepsilon \supseteq NS_m, m \geqq n) \to 1$ for each $\varepsilon > 0$. Further if $X \in C/|C|$ then for each $\varepsilon > 0$, $\mu((|C| X)^\varepsilon) > 0$ and thus $P(X^\varepsilon \cap NS_n \neq \emptyset) \to 1$ giving the first part of (C). If $NR_n \to T$ i.p. then by Lemma 2.1, $T = C/|C|$. But $C/|C|$ is not compact which would contradict Lemma 2.2 so that (C) is proved.

To prove (b) cover $C/|C|$ with a finite number of ε spheres then $P(NS_m$ interesects each sphere, $m \geqq n) \to 1$ implying $P(NS_m^\varepsilon \supseteq T, m \geqq n) \to 1$ which completes the proof of (b) and the proposition.

3. Rate of Growth of S_n

Since Proposition 2.6 disposes of the problem when μ has bounded support we *assume throughout the remainder of this section that μ has unbounded support.*

The following theorem shows that S_n must grow in an "orderly fashion." Let $M_n=|S_n|$, the radius of the smallest sphere containing $\{X_1, \ldots, X_n\}$.

Theorem 3.1. *Let T be a convex set with at least two nonzero extreme points.*

(a) *If $NR_n \to T$ i.p. then M_n is relatively stable in probability.*

(b) *If $NR_n \to T$ a.s. then M_n is relatively stable almost surely.*

Proof. Pick X and $Y \in T$, $X \neq \lambda Y$ for any λ. $X \neq \mathbf{0}$, $Y \neq \mathbf{0}$. We can find a continuous linear functional f on B such that $f(X)>0$ and $f(Y)<0$.

(a) Suppose the $NR_n \to T$ i.p. then T is compact. Let $f_M=\max\{f(X)\colon X \in T\}$ and $f_m=\min\{f(X)\colon X \in T\}$. By the Krein-Millman Theorem we may find points X_M and X_m of T such that $f_M=f(X_M)$ and $f_m=f(X_m)$ and X_M and X_m are extreme points of T.

Consider the sequence of independent identically distribution random variables $Z_i=f(X_i)$. Let $Q_n=\max\{f(X_i), i=1, \ldots, n\}$ $m_n=\min\{f(X_i), i=1, \ldots, n\}$.

Since X_M is an extreme point of T by Theorem 2.4(c) and the fact that f is continuous,

$$P(Q_n/M_n \geqq f_M-\varepsilon) \to 1 \quad \text{for each } \varepsilon>0. \tag{3.1}$$

By Theorem 2.4(b), $P(NS_n \subseteq T^{\varepsilon}) \to 1$ and using the continuity of f,

$$P(Q_n/M_n \leqq f_M+\varepsilon) \to 1 \quad \text{for each } \varepsilon>0. \tag{3.2}$$

By (3.1) and (3.2) $Q_n/M_n \to F_M$ i.p. In the same manner one shows that $m_n/M_n \to f_m$ i.p. Thus, $Q_n/m_n \to f_M/f_m \neq 0$ i.p. By Theorem A of the appendix Q_n and m_n are relatively stable i.p. Let $Q_n/a_n \to 1$ i.p., a_n a sequence of real numbers, $M_n/a_n \to 1/F_M$ i.p. so that M_n is relatively stable i.p.

Clearly (b) can be proved the same way concluding the proof.

Corollary 3.2. *If $NS_n \to T$ i.p. (a.s.) where T is not a line segment then M_n is relatively stable i.p. (a.s.).*

The restrictions upon the rate of growth of the sample may be extended to show that in certain directions the sample must grow at the same rate as M_n, while in other directions the rate must be less then or equal to the rate of M_n. These ideas are formulated below.

The set of points on the surface of the unit sphere shall be indexed by Θ, that is, for each Θ there is an $X(\Theta)$ such that $\|X(\Theta)\|=1$ and if $\|X\|=1$ then $X=X(\Theta)$ for some Θ. We topologize the index set by the metric $d(\Theta, \Theta')=\|X(\Theta)-X(\Theta')\|$. $c(\Theta, \varepsilon)$ will denote the cone in the Θ-direction generated by the ε-neighborhood of $x(\Theta)$ on the surface of the unit sphere, that is, $c(\Theta, \varepsilon)=\{Y\colon$ there exists Z with $\|Z-X(\Theta)\|<\varepsilon$, $\|Z\|=1$ and a $c>0$ such that $Y=cZ\}$.

We shall also want to truncate cones. For $\varepsilon>0$, $\delta>0$, and $Y \in B$, $\|Y\| \neq 0$ define:

$$c(Y, \varepsilon, \delta)=\{Z\colon\ Y/\|Y\|=X(\Theta),\ Z \in c(\Theta, \varepsilon) \text{ and } \|Z\| \leqq \|Y\|+\delta\}.$$

$M(n, \Theta, \varepsilon)$ denotes the random variable giving the maximum norm of the points of S_n in $c(\Theta, \varepsilon)$: $M(n, \Theta, \varepsilon) = \max(\{\|X_i\| : X_i \in C(\Theta, \varepsilon), i = 1, 2, \ldots, n\}, 0)$. Let Y_n be a sequence of random variables and a_n a sequence of constants, we write $Y_n \leqq a_n$ i.p. if $P(Y_n \leqq a_n) \to 1$ and $Y_n \leqq a_n$ a.s. if $\lim_{n\to\infty} P(Y_m \leqq a_m;\, m \geqq n) = 1$. Similar definitions hold for $Y_n \geqq a_n$ i.p. (a.s.).

The following definitions give the maximum and minimum radius of growth in a given direction for the normalized sample.

$$\mathrm{RSIP}(\Theta) = \lim_{\varepsilon\to 0} \inf_{r\geqq 0} \{M(n, \Theta, \varepsilon)/M_n \leqq r \text{ i.p.}\},$$

$$\mathrm{RIIP}(\Theta) = \lim_{\varepsilon\to 0} \sup_{r\geqq 0} \{M(n, \Theta, \varepsilon)/M_n \geqq r \text{ i.p.}\},$$

$$\mathrm{RSAS}(\Theta) = \lim_{\varepsilon\to 0} \inf_{r\geqq 0} \{M(n, \Theta, \varepsilon)/M_n \leqq r \text{ a.s.}\},$$

$$\mathrm{RIAS}(\Theta) = \lim_{\varepsilon\to 0} \inf_{r\geqq 0} \{M(n, \Theta, \varepsilon)/M_n \geqq r \text{ a.s.}\}.$$

(Recall that since μ is unbounded $M_n = \mathbf{0}$ only finitely often with probability one.) The following lemma is clear.

Lemma 3.3. $1 \geqq \mathrm{RSAS}(\Theta) \geqq \mathrm{RSIP}(\Theta) \geqq \mathrm{RIIP}(\Theta) \geqq \mathrm{RIAS}(\Theta) \geqq 0$.

Theorem 3.4. *Let $NR_n \to T$ i.p. (a.s.).*

(a) *Let $X = c \cdot X(\Theta)$, $c \geqq 0$ be an extreme point of T then* $\mathrm{RSIP}(\Theta) = \mathrm{RIIP}(\Theta) = \|X\|$ $(\mathrm{RSAS}(\Theta) = \mathrm{RIAS}(\Theta) = \|X\|)$.

(b) *Let $T(\Theta) = \max\{d \colon dX(\Theta) \in T\} = c$. If $cX(\Theta)$ is not an extreme point of T then* $\mathrm{RSIP}(\Theta) \leqq c$ $(\mathrm{RSAS}(\Theta) \leqq c)$.

Proof. (a) We consider two cases, $c > 0$ and $c = 0$.

(i) Suppose that $X = cX(\Theta) \neq \vec{0}$ is an extreme point of T. First note that for each $\varepsilon > 0$ we may choose a cone $c(\Theta, \delta)$ such that $(c(\Theta, \delta) - c(cX(\Theta), \delta, \varepsilon)) \cap T = \emptyset$. Geometrically the assertion is as follows: By taking a small enough cone with its apex at the origin and central ray through X we find that the intersection of the cone with T does not extend very far beyond X (in norm), (although the ray between $\mathbf{0}$ and X may lie on the surface of T).

Suppose the assertion is not true then there exists an $\varepsilon > 0$ such that $c(\Theta, \delta) - c(X, \delta, \varepsilon)$ contains points of T for each $\delta > 0$. Let Y_n be such a point for $\delta = 1/n$. $\|Y_n\| \leqq 1$ since $Y_n \in T$ and $|T| = 1$. Note that $\{Y_n, n = 1, 2, \ldots\} \cup \{Y \colon Y = rX, \|X\| + \varepsilon \leqq r \leqq 1\}$ is a compact set since it is closed as any limit point of the Y_n's must lie on the segment of the ray generated by X in the set given above and the set is totally bounded since any covering of the line segment by ε spheres contains all but a finite number of Y_n's. Thus, choose Y a limit point of Y_n, $n = 1, 2, \ldots$, then $\|Y\| \geqq \|X\| + \varepsilon$ and $Y \in T$ since T is closed. Since μ is unbounded $\mathbf{0} \in T$ and as Y is on the ray through X, X is not an extreme point which is a contradiction.

Let $\varepsilon > 0$ be given choose δ as above. For $0 < \beta < \delta$ if A is the closure of $c(\Theta, \beta) - c(X, \beta, 2\varepsilon)$ then A is contained in $c(\Theta, \delta) - c(X, \delta, \varepsilon)$ and $A \cap T$ is empty. Also $d(A, T) = a > 0$ and $A \cap T^b$ is empty where $b = a/2$. By Theorem 2.4, (b), $d(NS_n, T) \leqq b$ i.p. (a.s.) from which it follows that:

$$M(n, \Theta, \beta)/M_n \leqq \|X\| + 2\varepsilon \text{ i.p. (a.s.).} \tag{3.3}$$

From Theorem 2.4(c) it is immediate that:

$$M(n,\Theta,\beta)/M_n \geqq \|X\| - \varepsilon \text{ i.p. (a.s.).} \tag{3.4}$$

By combining (3.3) and (3.4)(a) is proved in the case where the extreme point is not the origin.

(ii) Suppose that $X=\mathbf{0}$, then there exists an infinity of rays originating at the origin which intersect T only there (i.e. every straight line through the origin has at least one half line satisfying the above condition or $\mathbf{0}$ is not an extreme point). Let one such ray be in the Θ-direction. Consider the line segment $L=\{r\cdot X(\Theta): \varepsilon \leqq r \leqq 1\}$. $T\cap L=0$. Let $d(L,T)=2a>0$, then it follows that:

$$M(n,\Theta,\min(a,\varepsilon))/M_n \leqq \varepsilon \text{ i.p. (a.s.)}$$

since $c(\Theta,\min(a,\varepsilon))-c(\varepsilon X(\Theta)/2, a, \varepsilon/2)$ is disjoint from $T^{a/2}$. Thus, $\mathrm{RSIP}(\Theta)\leqq 0$ and by Lemma 3.3 equal to zero $(\mathrm{RSAS}(\Theta)=0)$ completing the proof of (a).

(b) The proof of (b) is similar to that of part (a) so we sketch the proof. For each $\varepsilon>0$ we can find a $\delta>0$ such that the distance from T to $c(\Theta,\delta)-c(T(\Theta)X(\Theta),\delta,\varepsilon)$ is positive. Using Theorem 2.4(b) we find that

$$M(n,\Theta,\delta)/M_n \leqq T(\Theta)+\varepsilon \text{ i.p. (a.s.)}$$

from which (b) follows.

In the case that B is finite dimensional the type of behavior exhibited in the last theorem also gives a sufficient condition for a limit to exist.

Theorem 3.5. *Let B be finite dimensional.*

(a) *Let*

$$A=\{\mathrm{RSIP}(\Theta)\cdot X(\Theta):\ RSIP(\Theta)=\mathrm{RIIP}(\Theta)\}$$
$$C=\{\mathrm{RSIP}(\Theta)\cdot X(\Theta),\ \textit{all } \Theta\}.$$

Then $\lim_{n\to\infty} NR_n = T$ *i.p. for some T if and only if C is contained in the closure of the convex hull of A.*

(b) (a) *holds with IP replaced by AS and i.p. replaced by a.s.*

Proof. We only prove (a) since the proof (b) is essentially the same. Theorem 3.4 implies that if a limit exists A contains all the extreme points of T and is contained in T. The Krein-Millman theorem gives the only if portion of the theorem.

Thus, suppose that C is contained in the closure of the convex hull of A, which we will call T. Choose $\varepsilon(\Theta)$ such that:

$$M(n,\Theta,\varepsilon(\Theta))/M_n \leqq \mathrm{RSIP}(\Theta)+\varepsilon/2 \text{ i.p.} \tag{3.5}$$

The cones $c(\Theta,\varepsilon(\Theta))$ along with an $\varepsilon/2$-sphere about $\mathbf{0}$ cover the compact set T. Choose a finite subcovering. Then we have by (3.5) $P(NS_n\subseteq T^{\varepsilon})\to 1$ since T^{ε} contains

$$\{\mathbf{0}\}^{\varepsilon/2}\cup\{\bigcup_{\Theta} c(T(\Theta)\,X(\Theta),\varepsilon(\Theta),\varepsilon/2)\}.$$

Let X be an extreme point of T. Then by the theorem mentioned previously, [10] p. 132, X is in the closure of A, thus if $X = T(\Theta)\, X(\Theta)$,

$$M(n, \Theta, \varepsilon)/M_n \geqq \mathrm{RIIP}(\Theta) - \varepsilon \ \text{ i.p.} \tag{3.6}$$

As in the proof of Theorem 3.4 for each $\varepsilon > 0$ we can find a $\delta > 0$ such that $c(\Theta, \delta) - c(T(\Theta)\, X(\Theta), \delta, \varepsilon)$ is disjoint from A. Since C is contained in A

$$M(n, \Theta, \varepsilon)/M_n \leqq \mathrm{RSIP}(\Theta) + \varepsilon \ \text{ i.p.} \tag{3.7}$$

By (3.6) and (3.7), $X \in A$. Thus, for each $\varepsilon > 0$, $P(NS_n \cap X^\varepsilon \neq \emptyset) \to 1$. Since T is convex and compact by Theorem 2.4, $NR_n \to T$ i.p. concluding the proof.

As an application of this theorem the following result holds:

Theorem 3.6. *Let B be finite dimensional and μ a product of Borel probability measure on the set of angles and of Borel probability measure on the radial distance from* $\mathbf{0}$, *then:*

(a) *If there exists a Θ_0 such that $\mu(\{X\colon X = r X(\Theta_0), r \geqq 0\}) = 1$ then $NR_n \to T$ a.s. where $T = \{r X(\Theta_0)\colon 0 \leqq r \leqq 1\}$.*

(b) *Suppose that the radical measure does not put mass one on one angle. Let H be the support of the angular measure and $A = \{X(\Theta)\colon \Theta \in H\}$. Let T be the closure of the convex hull of A. Let F be the distribution function of the radical distribution.*

(i) $\lim_{n\to\infty} NR_n = T'$ *i.p. for some T' iff* $\lim_{r\to\infty} (1 - F(kr))/(1 - F(r)) = 0$ *for each* $k > 1$. *In this case $T' = T$.*

(ii) $\lim_{n\to\infty} NR_n = T'$ *a.s. for some T' iff* $\int_0^\infty (1 - F(kr))^{-1}\, dF(r) < \infty$ *for all* $k < 1$. *In this case $T' = T$.*

Proof. (a) is clear. Assume that the hypotheses of (b) are true. Suppose that $NS_n \to T'$ i.p. where T' has only one nonzero extreme point, that is, T' is a line segment from zero to $X(\Theta_0)$. Since the support of the angular measure is not a point we may find a closed set of angles A such that $\mu(A\, x\, [0, \infty)) = p > 0$ and $\Theta_0 \notin A$. Let $M(A, n)$ be the maximum norm of the sample points whose direction is in A. Let $M(A^c, n)$ bear a similar relation to A^c. Then $M(A, n)/M_n \to 0$ i.p. which contradicts Theorem B of the appendix. Thus, T' must have at least two nonzero extreme points and M_n is relatively stable i.p. if (i) is true and a.s. if (ii) holds by Theorem 3.1. The conditions on the distribution function F are those that M_n be relatively stable i.p. (a.s.), Barndorff-Nielsen [1], and the only if portion is proved.

Let M_n be relatively stable i.p. (a.s.) we show that $NR_n \to T$ i.p. (a.s.) completing the proof. If $\Theta \notin \mathrm{H}$ it is clear that $\mathrm{RSAS}(\Theta) = \mathrm{RIAS}(\Theta) = 0$. Let $\Theta \in H$. We may find $L(n)$, a slowly varying function, such that $M_n/L(n) \to 1$ i.p. (a.s.) [5, 8]. Since $\Theta \in H$ for each $\varepsilon > 0$, $p = \mu(c(\Theta, \varepsilon)) > 0$. By the above and the strong law of large numbers,

$$M(n, \Theta, \varepsilon)/M(n) \sim L(n\,p)/L(n) \to 1 \ \text{ i.p. (a.s.)}.$$

Thus, $\mathrm{RSIP}(\Theta) = \mathrm{RIIP}(\Theta) = 1$ $(\mathrm{RSAS}(\Theta) = \mathrm{RIIP}(\Theta) = 1)$. The proof is completed by referring to Theorem 3.5.

4. Dense Convergence

In general $NR_n \to T$ i.p. or a.s. does not imply that NS_n converges to a limit since NS_n converging is a more stringent condition requiring normalized sample points close to all points of the limit and not just the extreme points. The following theorem gives a case where this implication can be made.

Theorem 4.1. *Let $T(\Theta)=\max\{c: cX(\Theta)\in T\}$ where T is a compact convex set with at least two nonzero extreme points which contains the origin. Let the distribution of the X_i have unbounded support and let $T(\Theta)\,X(\Theta)$ be an extreme point of T for each Θ then*

(i) *$NR_n \to T$ i.p. iff $NS_n \to T$ i.p.*

(ii) *$NR_n \to T$ a.s. iff $NS_n \to T$ a.s.*

Proof. If $NS_n \to T$ i.p. or a.s. the corresponding result for NR_n holds by Theorem 2.4.

Suppose that $NR_n \to T$ i.p. By Theorem 3.1 M_n is relatively stable i.p. thus we may choose a sequence $L(n)$ such that $M_n/L(n) \to 1$ i.p. Let $T(\Theta)>0$ and $\varepsilon>0$ be given. As in equations (3.3) and (3.4) we may choose a $\delta>0$ such that

$$P\big(T(\Theta)(1-\varepsilon)\leqq M(n,\Theta,\delta)/L(n)\leqq T(\Theta)(1+\varepsilon)\big)\to 1. \tag{4.1}$$

Select an $x\,\varepsilon\,(0,1)$. It is easy to see there exists an N such that for $n\geqq N$ there exists a sequence $n(m)$ with the property that $|L(n)\,x/L(n(m))-1|<\varepsilon$.

Consider the sequence of random variables,

$$Z_i=\begin{cases}\|X_i\| & \text{if } X_i\in c(\Theta,\delta)\\ 0 & \text{otherwise.}\end{cases}$$

Let $Q_n=\{Z_1/L(n),\ldots,Z_n/L(n)\}$. Then letting

$$\begin{aligned} I=[T(\Theta)\,x(1-3\varepsilon),\,T(\Theta)\,x(1+3\varepsilon)],\qquad & P(Q_n\cap I\neq\emptyset)\geqq P\big(M(n(m),\Theta)/L(n)\in I\big)\\ &=P\big(x\,T(\Theta)(1-3\varepsilon)\,L(n)/L(n(m))\big)\leqq M\big(n(m),\Theta,\delta\big)/L\big(n(m)\big)\\ &\leqq x\,T(\Theta)(1+3\varepsilon)\,L(n)/L(n(m))\big)\geqq P\big(T(\Theta)(1-4\varepsilon)\\ &\leqq M\big(n(m),\Theta,\delta\big)/L\big(n(m)\big)\leqq T(\Theta)(1+4\varepsilon)\big)\end{aligned}$$

which approaches one as $m\to\infty$. Thus, we conclude that

$$P\big(\max\{d(x,Q_n)\colon x\in[\Theta,T(\Theta)]\}>\varepsilon\big)\to 0 \tag{4.2}$$

for each $\varepsilon>0$ where $d(x,Q_n)=\min\limits_i|x-Z_i/L(n)|$.

Let X and $Y\in c(\Theta,\delta)$ and $\big|\|X\|-\|Y\|\big|<\varepsilon$ and $\max(\|Y\|,\|X\|)<T(\Theta)+\beta$, $\beta>0$, then it is easy to see that $\|X-Y\|<2\big(T(\Theta)+\beta\big)+\varepsilon$.

Choose δ so small that $c(\Theta,\delta)\cap T\subseteq c\big(\pi(\Theta)\,X(\Theta),\delta,\varepsilon\big)$ (see the proof Theorem 3.4) and that $4\delta<\varepsilon$. Then by the preceeding remarks and (4.2) we see that

$$P\big(\max\{d(x,NS_n)\colon x\in c(\Theta,\delta)\}>2\varepsilon\big)\to 0. \tag{4.3}$$

Cover T with an open 2ε sphere about $\mathbf{0}$ and $c(\Theta,\delta)$ cones as outlined above for all angles where $T(\Theta)>0$. Since T is compact choose a finite subcovering. By (4.3) we see that $P(NS_n^{2\varepsilon}\supset T)\to 1$. Since $NR_n\to T$ i.p. we know that $P(T^{2\varepsilon}\subset NS_n)\to 1$, thus $P\big(d(NS_n\,T)<2\varepsilon\big)\to 1$ and (i) is proved. (ii) follows in the same manner.

5. Examples of Limiting Shapes

In this section three examples of limiting convex hulls are given. The following two results will be needed.

Theorem 5.1. (a) *Let μ be a measure on n-dimensional Euclidean space, E^n, which is the product measure of $N(\mu_i, \sigma_i^2)$ measures along an orthogonal set of axes. Let $\sigma_1^2 = \max\{\sigma_i^2, i=1, \ldots, n\}$. In the framework of section* 1, *let $B=E^n$ and let each X_i have distribution μ. Then*

$$\lim_{n\to\infty} NS_n = \left\{(x_1, \ldots, x_n): \sum_{i=1}^{n} (x_i \sigma_1/\sigma_i)^2 \leqq 1\right\} \text{ a.s.}$$

The same limit is obtained if instead of NS_n we use $S_n/\sigma_1(2\log n)^{\frac{1}{2}}$.

(b) *Let μ be a measure on E^n which is a product measure along an orthogonal set of axes of Poisson distributions with parameters λ_i, $i=1, \ldots, n$. Let $B=E^n$ and each X_i have distribution μ. Then,*

$$\lim_{n\to\infty} NS_n = \left\{(x_1, \ldots, x_n): \sum_{i=1}^{n} x_i \leqq 1, x_i \geqq 0, i=1, 2, \ldots, n\right\} \text{ a.s.}$$

Proof. (a) follows from the stronger results of Geffroy [7] or from Fisher [5].

(b) Let $G(x) = 1 - F(x) = \exp(-\lambda) \sum_{i=[x]} \lambda^i/i!$ for $x \geqq 0$ and $G(x)=1$ for $x<0$. That is, $G(x)$ is the tail of a Poisson distribution function with parameter λ. Let $L(y) = \log(y)/\log(\log y)$. We now show that for $x>0$ and $y>0$,

$$\log G(xL(y)) = -x\log y + o(\log y)\ y\to\infty. \tag{5.1}$$

Using the fact that $G(x) \approx \exp(-\lambda)\,\lambda^{[x]}/[x]!$ as $x\to\infty$ (where $[x]$ is the integer part of x) we see that for fixed x, $y\to\infty$

$$\begin{aligned}\log G(xL(y)) &= \log\big((\exp(-\lambda)\,\lambda^{[xL(y)]}/[xL(y)]!)\,(1+o(1))\big) \qquad \text{as } y\to\infty\\ &= [xL(y)]\log\lambda + o(\log y) - \log[xL(y)]!.\end{aligned}$$

Now $[xL(y)] = [x\log y/\log(\log y)] = o(\log y)$ and using Stirling's formula we find

$$\log G(xL(y)) = -[xL(y)]\log[xL(y)] + o(\log y).$$

Note that as $Z\to\infty$,

$$|Z\log Z - [Z]\log[Z]| \leqq |Z\log Z - [Z]\log Z| + |[Z](\log Z - \log[Z])| = o(Z).$$

Thus,

$$\begin{aligned}\log G(xL(y)) &= -(xL(y))\log(xL(y)) + o(\log y)\\ &= -xL(y)\log L(y) + o(\log y)\\ &= -x\log y(\log_2 y - \log_3 y)/\log_2 y + o(\log y)\\ &= -x\log y + x\log y\log_3 y/\log_2 y + o(\log y)\end{aligned}$$

which proves (5.1).

Let F be as above. For $x>1$ define $N(x)$ as $N(x)=\min\{y\colon F(y-0)\leqq 1-1/x\leqq F(y)\}$. Note that

$$yG(N(y))\leqq 1 \quad \text{and} \quad 1\leqq yG(N(y)-0) \quad \text{by definition of } N. \tag{5.2}$$

By (5.1), $0<\varepsilon<1$, $\log y\, G((1\pm\varepsilon)\, L(y))=\pm\varepsilon \log y+o(\log y)$ so that $y\, G((1+\varepsilon)\, L(y))\to 0$ and $y\, G((1-1)\, L(y))\to\infty$. Since G is nonincreasing, by (5.2) for large y, $(1+\varepsilon)\, L(y)>N(y)>(1-\varepsilon)\, L(y)$ so that $N(y)\cong L(y)$. The proof of (b) now follows from Theorem 3.4 of [5].

A. Normal Distributions in l_2

For a discussion of normal distributions in Hilbert space see Grenander [9], pp. 140–143. For our purposes we summarize the results by noting that a normal distribution is a product measure of $N(\mu_i, \sigma_i^2)$ measures along an orthogonal set of axes where $\sum_{i=1} \sigma_i^2<\infty$.

Theorem 5.2. *Let $B=l_2$ and X_i have distribution μ which is normal. Let $X_i=(x_1, x_2, \ldots)$ where x_i is $N(\mu_i, \sigma_i^2)$ independently of the other x_j's. Let $\sigma_1^2=\max\{\sigma_i^2, i=1, 2, 3, \ldots\}$, then*

$$\lim_{n\to\infty} NS_n=\{(x_1, \ldots)\colon \sum_{i=1} (x_i\, \sigma_1/\sigma_i)^2\leqq 1\} \quad a.s.$$

Proof. Without loss of generality assume that $\mu_i=0$ for all i and $\sigma_1=1$. By Theorem 5.1(a) if we "observe" only the first m coordinates then $\lim_{n\to\infty} NS_n^m= \left\{(x_1, \ldots, x_m)\colon \sum_{i=1}^{m} (x_i/\sigma_i)^2\leqq 1\right\}$ where NS_n^m is NS_n projected onto the m dimensional subspace of vectors whose only nonzero coordinates are in the first m coordinates. The same limit occurs when considering $S_n/(2\log n)^{\frac{1}{2}}$.

We now proceed to show that in the full Hilbert space the limit is $A=\{(x_1, \ldots)\colon \sum_{i=1} (x_i/\sigma_i)^2\leqq 1\}$. We prove this result by showing that "most" of the coordinates are "negligible."

If $|u|<\frac{1}{2}\sigma_i^{-2}$ then $E(\exp(u x_i^2))=(1-2u\sigma_i^2)^{-\frac{1}{2}}$. Since $\sigma_i^2\to 0$ we see that

$$(1-2u\sigma_i^2)^{-\frac{1}{2}}\cong 1+u\sigma_i^2 \qquad \text{as } i\to\infty.$$

Let $\|X(n)\|^2=\sum_{i=n}^{\infty} x_i^2$ then $E(\exp(u\|X(n)\|^2))=\prod_{i=n}^{\infty}(1-2u\sigma_i^2)^{-\frac{1}{2}}$ which converges if $|u|<\frac{1}{2}\sigma_i^{-2}$ for each $i\geqq n$ since $(1-2u_i^2)^{-\frac{1}{2}}\cong 1+u_i^2$ and $\sum |u|\, \sigma_i^2<\infty$.

By the general Chebyshev inequality if $a, b, >0$ we have $P(\|X(n)\|^2>b)\leqq E(\exp(a\|X(n)\|^2))/\exp(a b)$. Let $0<u<1/(2\max_{i\geqq k}\sigma_i^2)$. Then for any $\varepsilon>0$, $P(\|X(k)\|^2> \varepsilon\, 2\log n)\leqq \prod_{i=k}^{\infty}(1-2\, u_i^2)^{-\frac{1}{2}}/\exp(u\,\varepsilon\, 2\log n)=f(u)/n^{2\varepsilon u}$. Let $1/\varepsilon<u<1/(2\max_{i\geqq k}\sigma_i^2)$. (Note that this implies $2\max_{i\geqq k}\sigma_i^2<\varepsilon$ so that k must be fairly large.) Let $A(k, n)$ be the event that a sample point has $\|X(k)\|^2>\varepsilon\, 2\log n$. Note that $\sum_n P(A(k, n))\leqq c(k, \varepsilon)+\sum_n f(u)/n^2<\infty$ when $\varepsilon u>1$ and where $c(k, c)$ is a fixed constant.

Let $\varepsilon>0$ be given. Choose k as above so that $2\max_{i\geqq k}\sigma_i^2<\varepsilon$. By the result of Theorem 5.1(a), $\sum_{i=1}^{m-1}(x(i,n)/\sigma_i)^2\leqq(1+\varepsilon)(2\log n)$ for all large n with probability one where $x(i,n)$ is the i^{th} coordinate of the n^{th} sample point. Recalling the interpretation of $A(k,n)$ we use the fact that $\sum_n P(A(k,n))<\infty$ (for each k) along with the Borel-Cantelli lemma to see that with probability one $\sum_{i=k}^{\infty}x(i,n)^2\leqq\varepsilon\, 2\log n$ for all large n. Thus, $\operatorname{limit}_{n\to\infty} P(S_m/(2\log m)^{\frac{1}{2}}\subseteq A^{\varepsilon}, m\geqq n)=1$.

Noting that for large k the distance between A and the set A projected onto its first k coordinates is less than ε the theorem follows from Theorem 5.1 (a) (since then $P(NS_m^{\varepsilon}\supseteq A, m\geqq n)\to 1$) and the previous remarks.

B. Poisson Process on K

Let K be the space of real-valued functions on $[0,1]$ that are continuous from the right and have a limit from the left. K is topologized with the $J1$-metric (e.g. Skorokhod [12]). Let μ be the measure on K associated with a Poisson process with parameter λ.

Theorem 5.3. *Let $B=K$ and each X_i have distribution μ. Then $\lim_{n\to\infty}NS_n=\{f\colon f$ is continuous and nondecreasing on $[0,1]$. $f(0)=0$. $f(1)\leqq 1\}$ w.a.s.*

Note. Let W be the limit described above. W is convex, closed and noncompact. Thus, NR_n has the same limit and the limit cannot exist as a strong limit since the set W is not compact.

Proof. Let $\varepsilon>0$ be given. Each sample point $X_n\equiv f_n\in K$ has only a finite number of jumps of height one in $[0,1]$ (Doob [2], p. 401) a.s. Let $L(y)=\log y/\log(\log y)$.

The distance between f and g in the $J1$-metric is less than or equal to the distance between f and g in the sup metric. Let $f(n)$ denote the n^{th} sample point chosen. Let W^{ε} be the ε-neighborhood of W in the sup norm. We will show that $P(f(n)/L(n)\in W^{\varepsilon}, n\geqq N)\to 1$ as $N\to\infty$. Let $\|\cdot\|$ denote the sup norm of functions in K. Let $f(n)$ have jumps at $z(1),\dots,z(s)$. Define $\underset{\sim}{f}$ as follows: $\underset{\sim}{f}(0)=(n)(0)$, $\underset{\sim}{f}(z(i))=(f(n)(z(i)))+f(n)(z(i)-(0))/2$, $\underset{\sim}{f}(1)=f(1)$ and $\underset{\sim}{f}$ is found by linear interpolation in between the above points. Then $\|f(n)-\underset{\sim}{f}\|=\frac{1}{2}$ if we have a jump, $=0$ if there are no jumps. Thus, $\|\underset{\sim}{f}/L(n)-f(n)/L(n)\|\leqq\frac{1}{2}L(n)$. By the proof of Theorem 5.1 (b), (5.2) $P(f(n)(1)/L(n)\leqq 1+\varepsilon/2, n\geqq N)\to 1$ as $N\to\infty$. Define $h^n(x)=\min(\underset{\sim}{f}(x)/L(n),1)$. Then,

$$
\begin{aligned}
&P(f(n)/L(n)\in W^{\varepsilon}, n\geqq N)\\
&\quad\geqq P(\|h^n-f(n)/L(n)\|\leqq\varepsilon, n\geqq N)\\
&\quad\geqq P(\|h^n-\underset{\sim}{f}/L(n)\|+\|\underset{\sim}{f}/L(n)-f(n)/L(n)\|\leqq\varepsilon, n\geqq N)\\
&\quad\geqq P(\|h^n-\underset{\sim}{f}/L(n)\|<\varepsilon/2, n\geqq N \text{ and } \tfrac{1}{2}L(n)<\varepsilon/2).
\end{aligned}
$$

By (5.2) this quantity approaches 1 as $N\to\infty$. Let $f\in W$. Let $\varepsilon>0$ be given. Since f is uniformly continuous choose $\delta>0$ such that $|x-y|<\delta$ implies $|f(x)-f(y)|<\varepsilon$.

Consider the points $x(i)=i\,\delta/2$ where $i=0,\dots,k-1$ and $0<1-(k-1)\,\delta/2\leqq\delta/2$ and $x(k)=1$. Define $df(n,x(i))=f(n)(x(i+1))-f(n)(x(i))$ and $dx(i)=x(i+1)-x(i)$. The joint distribution of $df(n,x(i))$, $i=1,\dots,k-1$ is the product measure of Poisson distributions with parameters $dx(i)$. Since $f(1)-f(o)\leqq 1$ we have $\sum_{j=0}^{k-1} df(x(j))\leqq 1$. From the proof of Theorem 5.1 (b) with a probability approaching one as $N\to\infty$ for $n\geqq N$ we may find a sample point $f(i)$ with $|f(i)(x(j))/L(n)-f(x(j))|\leqq\varepsilon$ for $j=0,1,2,\dots,k$ and $i\leqq n$. Let $x\in[x(j),\,x(j+1)]$ and

$$\begin{aligned} f(i)(x)\equiv f(i,x), \quad & |f(i,x)|L(n)-f(x)|\leqq|f(i,x)/L(n)-f(x(j))| \\ & +|f(x)-f(x(j))|\leqq\varepsilon \\ & +|f(i,x(j))/L(n)-f(i,x)/L(n)|+|f(i,x(j))/L(n)-f(x(j))| \\ & \leqq 2\varepsilon+|f(i,x(j))/L(n)-f(i,x(j+1))/L(n)|\leqq 5\,\varepsilon. \end{aligned}$$

Thus, $f\in W$ and $\varepsilon>0$ implies $\lim\limits_{n\to\infty} P(S_n/L(n)\cap f^\varepsilon\neq\emptyset,\, n\geqq N)=1$ where the ε neighborhood is in the sup norm. Since f is continuous convergence to f is equivalent for the sup and J 1-metrics, thus the result also holds for the ε neighborhood of f in the J 1-metric.

C. Wiener Measure on $C_k([0,1])$

Let $C_k([0,1])$ be the space of k-dimensional real-valued continuous functions defined on $[0,1]$. Let $K=\{f: f\in C_k([0,1]), f(0)=0, f$ is absolutely continuous and $\int_0^1 (\dot f(t))^2\,dt\leqq 1\}$ where $(\dot f(t))^2$ is the usual Euclidean inner product of $f(t)$ with itself. (The sup norm is used on $C_k([0,1])$ with respect to the Euclidean metric in E^k.)

The set K was used by Strassen [13] who proved the following theorem: Let x be Brownian motion in E^k. Define $x(n,t)=(2n\log(\log n))^{-\frac{1}{2}}\,x(n\,t)$ for $t\in(0,1)$. With probability one the sequence $(x(n))\; n\geqq 3$ is relatively norm compact and the set of its limit points coincides with K. Following Strassen we prove:

Theorem 5.4. *Let μ be Wiener measure on $C_k([0,1])$. Let each X_i have distribution μ. Then $\lim\limits_{n\to\infty} NS_n=K$ a.s.*

Proof. The set K is compact (Strassen [13]) and convex. Let $L(n)=(2\log n)^{\frac{1}{2}}$. We show that $\lim\limits_{n\to\infty} S_n/L(n)=K$ a.s. which gives the result of the theorem. Let $f(n)=X_n$. The first portion of proof involves showing that $\lim\limits_{N\to\infty} P(f(n)/L(n)\in K^\varepsilon,\, n\geqq N)=1$. The proof follows that of Strassen, pp. 212–214, [13] with a different normalizing factor $(2\log n)^{\frac{1}{2}}$ instead of $(2n\log\log n)^{\frac{1}{2}}$ but goes through in the same manner and will not be repeated here.

For $f\in C_k([0,1])$ let $|f(x)|$ denote the Euclidean length of the vector $f(x)$. Now let $f\in K$ we show that

$$\lim_{N\to\infty} P((S_n/L(n))\cap f^\varepsilon\neq\emptyset,\, n\geqq N)=1 \qquad \text{for each } \varepsilon>0. \tag{5.3}$$

Choose a finite increasing set of points $t(i)$, $i=1,2,\dots$, such that for all t in $[0,1]$ there exists an i such that $|f(t)-f(t(i))|\leqq \varepsilon/3$ and also such that $t(i+1)-t(i)\leqq\varepsilon/6$. Let $f(j,t)$ be the j^{th} component of $f(t)$.

Now

$$|f(t(i))-f(t(i-1))|\leqq \int_{t(i-1)}^{t(i)} |\dot f(t)|\,dt \leqq \left(\int_{t(i-1)}^{t(i)} (\dot f(t))^2\,dt\right)^{\frac12} (t(i)-t(i-1))^{\frac12}.$$

Thus,

$$d(i)\equiv (f(t(i))-f(t(i-1)))^2/(t(i)-t(i-1)) \leqq \int_{t(i-1)}^{t(i)} (\dot f(t))^2\,dt$$

and finally

$$\sum_{i,j} (f(j,t(i))-f(j,t(i-1)))^2/(t(i)-t(i-1))\leqq 1 \tag{5.4}$$

since the first sum is $=\sum_i d(i)\leqq \int_0^1 (\dot f(t))^2\,dt\leqq 1$ by definition of K.

Let B be a $k\,l$-dimensional Euclidean space and m a measure on B which is a product measure of one dimensional normal distributions where l of the distributions have variance $t(i)-t(i-1)$. Let $N(x,\varepsilon,p)$ be the open square box with sides of length ε and center at x in Euclidean p-space. The probability that $f(j,t(i))/L(n)-f(t(i))\in N(\mathbf{0},\varepsilon,k)$ for $i=0,1,\dots,l$ and some $j\leqq n$ for all $n\geqq N$ is the same as the probability that if we choose a sequence Y_n of independent random vectors from E^k with distribution m that for $n\geqq N$, $S_n/L(n)\cap N((t(1),\dots,t(l)),\varepsilon,k)$ is not empty. By Theorem 5.1 (a) this probability approaches one as $N\to\infty$.

Now with probability one for all large n the $f(n)/L(n)$ are within $\varepsilon/6$ of K. Now if $f-\underline{f}$ is in $N(0,\varepsilon/6,k)$ at each $t(i)$, $\underline{f}(0)=0$ and $\underline{f}$ is within $\varepsilon/6$ of K we see that

$$\begin{aligned}|f(t)-\underline{f}(t)|&\leqq|f(t)-f(t(j))|+|f(t(j))-\underline{f}(t(j))|\\&\quad+|\underline{f}(t(j))-\underline{f}(t)|\leqq\varepsilon/6+|\underline{f}(t(j))-g(t(j))|+|g(t(j))-g((t))|\\&\quad+|g(t)-\underline{f}(t)|\leqq 4\varepsilon/6+\left|\int_{t(j)}^{t}|\dot g(t)|\,dt\right|\leqq 4\varepsilon/6+|t(j)-t|^{\frac12}<\varepsilon\end{aligned}$$

where $\underline{f}$ is within $\varepsilon/6$ of g which is in K and $t\in[t(i),t(i+1)]$ and the last step used Schwarz's inequality. Since K is compact, $\lim_{N\to\infty} P((S_n/L(n))^{\varepsilon}\supseteq K, n\geqq N)=1$ and the proof is complete.

Acknowledgment. The author wishes to express his deepest thanks to Professor John Lamperti of Dartmouth College who directed the research presented in this paper. Many of these results were announced in [4].

Appendix. Relative Stability of the Extreme Values of a Sample

Theorem A. *Let* $X_1,\dots,X_n,\dots$ *be a sequence of i.i.d. r.v.'s with distribution function F such that* $0<F(x)<1$ *for all x. Let*

$$M_n=\max\{X_1,\dots,X_n\},$$
$$m_n=\min\{X_1,\dots,X_n\}.$$

Then (A) *if* $M_n/m_n \to l \neq 0$ *i.p. then* M_n *and* m_n *are relatively stable in probability;* (B) *if* $M_n/m_n \to l \neq 0$ *a.s. then* M_n *and* m_n *are relatively stable almost surely.*

Proof. (A) Geffroy [6] proves that if $M_n + m_n$ is stable i.p. then M_n and m_n are stable i.p.

Let $\bar{M}_n$ and $\bar{m}_n$ be the maximum and minimum of a sequence of random variables with distribution function

$$\bar{F}(x) = \begin{cases} F(e^x) & \text{for } x \geqq 1 \\ F(0) & \text{for } -1 < x < 1 \\ F(-e^{|x|}) & \text{for } x \leqq -1. \end{cases}$$

Assuming that $\bar{X}_1, \dots, \bar{X}_n$ has points >1 and points <1 we see using $\max(\log a_1, \dots, \log a_m) = \log(\max(a_1, \dots, a_m))$, that we can define (for large n) $\bar{M}_n = \log M_n$ and $\bar{m}_n = -\log|m_n|$. Since $M_n/|m_n| \to c = |l| \neq 0$ i.p. $\log M_n - \log|m_n| - \log c \to 0$ i.p. Thus, $\bar{M}_n + \bar{m}_n$ is stable i.p. By the results of Geffroy, $\bar{M}_n$ and $\bar{m}$ are stable i.p. Thus, M_n and m_n are relatively stable i.p. proving (A).

(B) Suppose that $M_n/m_n \to l \neq 0$ a.s., then M_n and m_n are relatively stable i.p. by (A). Let $L(n) = \min\{x \mid F(x-0) \leqq 1 - 1/n \leqq F(x)\}$ and $G(n) = \min\{x \mid F(x-0) \leqq 1/n \leqq F(x)\}$. Then by the results of Gnedenko [8] $M_n/L(n) \to 1$ i.p. and $m_n/G(n) \to 1$ i.p. where L and G are slowly varying functions.

Suppose that $M_n/L(n) \nrightarrow 1$ a.s. Then there exists an $\varepsilon > 0$ such that

$$P(M_n > L(n)(1+\varepsilon) \text{ i.o.}) > 0$$

or

$$P(M_n < L(n)(1-\varepsilon) \text{ i.o.}) > 0.$$

We will assume that the first inequality holds (the proof for the other inequality proceeds in an analogous manner). In this case

(A.1) $\quad P(M_n > L(n)(1+\varepsilon) \text{ i.o.}) = 1$ since the event is in the tail field of the i.i.d. r.v.'s X_i.

Let $p = F(0)$, $q = 1 - p = 1 - F(0)$. Let $A(n, c)$ be the event that $m \geqq n$ implies that the number of nonpositive X_i among $X_1, \dots, X_m$ lies in the interval $[pm(1-c), pm(1+c)] = I_{m,c}$. For each $c > 0$ the strong law of large numbers gives,

$$\lim_{n \to \infty} P(A(n, c)) = 1. \tag{A.2}$$

Let $B(n, k, l, r)$ be the event $\{|m_n/G(n) - 1| < \varepsilon/2$, l is the number of nonnegative X_i's among $X_1, \dots, X_n$ and r is the number of nonnegative X_i's among $X_1, \dots, X_{n+k}\}$. Let $c(m)$ be the event $|m_m/G(m) - 1| < \varepsilon/2$. By (A.2), the fact that G is a slowly varying function and $m_n/G(n) \to 1$ i.p. we have:

$$\lim_{n \to \infty} \min_{*} P(c(n+k) \mid B(n, k, l, r)) = 1 \tag{A.3}$$

where $*$ is the condition $k = 0, 1, 2, \dots$, $l \in I_{n,1}$, $r \in I_{n+k,1}$.

Let $E(n, k)$ be the event $\{|m_n/G(n) - 1| < \varepsilon/2$, $M_{n+i} < L(n+i)(1+\varepsilon)$, $i = 0, 1, 2, \dots, k-1$, $M_{n+k} > L(n+k)(1+\varepsilon)\}$. Noting that $A(N, 1) \subseteq \bigcup_{l \in I_{N,1},\, p \in I_{N+k,1}} B(N, k, l, p)$ for each $k \geqq 0$, it is probabilistically clear that

$$\begin{aligned} P(c(N+k) \mid E(N, k) \cap A(N, 1)) &= P(c(N+k) \mid E(N, k) \cap A(N, 1) \cap B(N, k, l, p)) \\ &\geqq \min_{l \in I_{N,1},\, p \in I_{N+k,1}} P(c(N+K) \mid A(N, 1) \cap B(N, k, l, p)) \\ &= \min_{l \in I_{N,1},\, p \in I_{N+k,1}} P(c(N+k) \mid B(N, k, l, p)). \end{aligned} \tag{A.4}$$

(Since $P(c(N+k) \mid A(N, 1) \cap B(N, k, l, p)) = P(c(N+k) \mid B(N, k, l, p))$.)

Let

$$D(N) = \bigcup_{k=1}^{\infty} (E(N, k) \cap c(N+k)).$$

Choose N so large that $P(A(N,1))\geqq\frac{3}{4}$ by (A.2) and $\min_{l\in I_{N,1},\, p\in I_{N+k,1}} P(c(N+k)|B(N,k,l,p))\geqq\frac{1}{2}$ by (A.3). Then,

$$\begin{aligned}
P(D(N)) &= \sum_{k=1}^{\infty} P(E(N,k)\cap c(N+k)) \geqq \sum_{k=1}^{\infty} P(E(N,k)\cap c(N+k)\cap A(N,1))\\
&= \sum_{k=1} P(E(N,k)\, A(N,1))\, P(c(N+k)|E(N,k)\, A(N,1))\\
&\geqq \tfrac{1}{2}\sum_{k=1}^{\infty} P(E(N,k)\cap A(N,1))
\end{aligned}$$

by the selection of N and (A.4). Thus,

$$P(D(N))\geqq\tfrac{1}{2}\left(P\left(\bigcup_{k=1}^{\infty} E(N,k)\right)-P(A(N,1)^c)\right)=\tfrac{1}{2}\left(P\left(\bigcup_{k=1}^{\infty} E(N,k)\right)-\tfrac{1}{4}\right).$$

By $M_n/L(n)\to 1$ i.p. and (A.1) we see that

$$\lim_{n\to\infty} p\left(\bigcup_{k=1}^{\infty} E(N,K)\right)=1.$$

Thus,

$$\text{(A.5)}\qquad \liminf_{N\to\infty} P(D(N))\geqq\tfrac{3}{8}.$$

Since the event $\{D(n)$ happens i.o.$\}$ lies in the tail field of $X_1, X_2, \ldots$ the probability must be 0 or 1. By (A.5), $P(D(n)$ happens i.o.$)=1$. But $D(n)$ implies that for some $k\geqq 0$.

$$|M_{n+k}/G(n+k)-1|<\varepsilon/2 \quad\text{and}\quad M_{n+k}/L(n+k)>(1+\varepsilon).$$

Since $M_n/m_n\to l\neq 0$ i.p. $G(n+k)/L(n+k)\to l$. Thus,

$$\left|\frac{M_{n+k}}{m_{n+k}}\right|>\frac{|G(n+k)|(1+\varepsilon)}{|L(n+k)|(1+\varepsilon/2)}>|l|(1+3\varepsilon/4)$$

for small ε, large $n+K$, when $D(n)$ happens. But $D(n)$ occurs i.o. a.s. contradicting $M_n/m_n\to$ a.s. End of proof.

Theorem B. *Let $0<p<1$ and $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables. Let Y_i be a sequence of Bernoulli trials with probability p of being one (independent of the X_i's). Let*

$$m_n=\max\{Y_i X_i,\ i=1,\ldots,n\},$$
$$M_n=\max\{(1-Y_i) X_i,\ i=1,\ldots,n\}$$

then it cannot happen that $m_n/M_n\to 0$ i.p.

Proof. Let F be the d.f. of the X_i. Let $N(n)=\sum_{i=1}^{n} Y_i$ and $Q(n)$ be the event $N(n)>\frac{np}{2}$. By the weak law of large numbers, $P(Q(n))\to 1$ as $n\to\infty$. It is clear that:

$$\text{(A.6)}\qquad P(m_n<\varepsilon M_n|N(n)=j)\leqq P(m_n<\varepsilon M_n|N(n)=k)$$

for any $k<j$ and $\varepsilon>0$. Now,

$$\begin{aligned}
P(m_n<\varepsilon M_n\, Q(n)) &= \sum_{j=\left[\frac{np}{2}\right]+1}^{n} P(m_n<\varepsilon M_n|N(n)=j)\cdot P(N(n)=j)\\
&\leqq P\left(m_n<\varepsilon M_n|N(n)=\left[\frac{np}{2}\right]\right)\cdot\sum p(N(n)=j)\\
&= P\left(m_n<\varepsilon M_n|N(n)=\left[\frac{np}{2}\right]\right)\cdot P(Q(n))
\end{aligned}$$

where we have used (A.6). Suppose that $m_n/M_n \to 0$ i.p. then since $p(m_n < \varepsilon M_n \cap Q(n)) \to 1$ as does $p(Q(n))$ we must have $P\left(m_n < \varepsilon M_n \,|\, N(n) = \left[\frac{n p}{2}\right]\right) \to 1$ for each $\varepsilon > 0$.

The following probabilities P will all be conditioned upon the event $N(n) = [n p/2]$. Let $x_n = \inf\{y\colon F^{n-[np/2]}(y) \geqq \frac{1}{2}\}$. If $F(x) = 1$ for some finite x, the theorem is clear so we assume that $F(x) < 1$ for all x. Then $x_n \to \infty$. If $m_n \geqq x_n - 1$ and $M_n \leqq x_n$ then $m_n/M_n \geqq 1 - 1/x_n$ so that

$$P(m_n \geqq x_n - 1,\, M_n \leqq x_n) = F^{n-[np/2]}(x_n) \cdot \left(1 - F^{[np/2]}(x_n - 1)\right) \to 0 \tag{A.7}$$

which implies $F^{[np/2]}(x_n - 1) \to 1$. Now by definition of x_n:

$$F^{n-[np/2]}(x_n - 1) \leqq \tfrac{1}{2}$$

so that

$$F^{[np/2]}(x_n - 1) \leqq F^{np/2}(x_n - 1)/F(x_n - 1) \leqq \left(F^{n-[np/2]}(x_n - 1)\right)^{p/2}/F(x_n - 1) \leqq (\tfrac{1}{2})^{p/2}/F(x_n - 1)$$

which approaches $(\frac{1}{2})^{p/2} < 1$ since $p > 0$ contradicting the implication of (A.7) and giving the desired result.

References

1. Barndorff-Nielsen, O.: On the limit behavior of extreme order statistics. Ann. math. Statistics **34**, 992–1002 (1963).
2. Doob, J.L.: Stochastic processes. New York: Wiley 1953.
3. Efron, B.: The convex hull of a random set of points. Biometrika **52**, 331–343 (1965).
4. Fisher, L.: The convex hull of a sample. Bull. Amer. math. Soc. **72**, 555–558 (1966).
5. – Limiting sets and convex hulls of samples from product measures. Ann. math. Statistics **40**, 1824–1832 (1969).
6. Geffroy, J.: Contribution à la théorie des valeurs extrèmes. Publ. Inst. Statist. Univ. Paris **7**, 36–123 and **8**, 3–52 (1958).
7. – Localisation asymptotique du polyèdre d'appui d'un échantillon Laplacian à k dimensions. Publ. Inst. Statist. Univ. Paris **10**, 213–228 (1961).
8. Gnedenko, B.V.: Sur la distribution limite du terme maximum d'une serie aléatoire. Ann. of Math. II. Ser., **44**, 423–453 (1943).
9. Grenander, U.: Probabilities on algebraic structures. New York: Wiley 1963.
10. Kelley, J.L., Namioka, I., Donoghne, W.F., Jr., Lucas, K.R., Pettis, B.J., Paulsen, E.T., Price, G.B., Robertson, W., Scott, W.R., Smith, K.T.: Linear topological spaces. New Jersey: Van Nostrand 1963.
11. Rényi, A., Sulanke, R.: Über die konvexe Hülle von n zufällig gewählten Punkten I und II. Z. Wahrscheinlichkeitstheorie verw. Geb. **2**, 75–84 (1963) and **3**, 138–148 (1964).
12. Skorohod: Limit theorems for stochastic processes. Theor. Probab. Appl. **1**, 261–290 (1956).
13. Strassen, V.: An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie verw. Geb. **3**, 211–226 (1964).

Lloyd Fisher
Department of Mathematics
University of Washington
Seattle, Washington 98105
USA

(Received February 15, 1969)