

THE CONVEX HULL OF A SAMPLE¹

BY LLOYD D. FISHER, JR.

Communicated by D. A. Darling, December 16, 1965

1. The convex hull of a random sample may be considered as one possible analogue of the range of a one-dimensional sample. Recent work along this line has dealt with the expected number of vertices, faces, surface area and other quantities connected with the convex hull of n independently and identically distributed random points in the plane and in higher dimensions. See Renyi and Salanke [6] and Efron [1]. Geffroy [2], [3] has shown that if the random points are normally distributed then the convex hull "grows like" an ellipsoid (for a concise statement of these results see Math. Rev. 25 #4559). Some features of the limiting geometric behavior of the convex hull will be described here.

2. Let B be a Banach space. Let $x(1), x(2), \dots$ be a sequence of independent, identically distributed, Borel measurable, B -valued random variables defined on a probability space (Ω, \mathcal{F}, P) . Let m be the measure defined on all Borel sets A in B by $m(A) = P(\{\omega: x(1)(\omega) \in A\})$. Let $H(n)$ be the convex hull of $\{x(1), \dots, x(n)\}$. Let $S(n) = H(n) / \max_{i=1, \dots, n} \|x(i)\|$, (where $\{0\}/0 = \{0\}$). In other words, $S(n)$ is the random convex hull normalized to have norm one. Let A and C be closed, bounded, convex sets contained in B . We define $d(A, C) = \inf \epsilon$ such that $\epsilon > 0$, $A + \epsilon \supseteq C$ and $C + \epsilon \supseteq A$, where $A + \epsilon = \{y \mid \exists x \in A, z \in B \text{ with } 0 \leq \|z\| < \epsilon \text{ and } y = x + z\}$. We say that $C = \text{LIP}$ iff $d(C, S(n)) \rightarrow 0$ in probability and that $C = \text{LAS}$ iff $d(C, S(n)) \rightarrow 0$ almost surely.

It is easy to see that if LAS or LIP exists then it is a compact set. We are thus led to a second, weaker type of limit. Let C be a closed, convex set contained in B . We say that $C = \text{WLIP}$ if for all $x \notin C$, $\exists \epsilon > 0$ such that $P(x \in S(n) + \epsilon) \rightarrow 0$ and if for all $x \in C$ then $P(x \in S(n) + \epsilon) \rightarrow 1$ for each $\epsilon > 0$. We shall say that $C = \text{WLAS}$ if for all $x \notin C$, $\exists \epsilon > 0$ such that $x \notin S(n) + \epsilon$ almost surely for all large n , and also if for all $x \in C$ then $x \in S(n) + \epsilon$ almost surely for large n and each $\epsilon > 0$. It is easy to prove that if B is a finite dimensional space then LIP exists iff WLIP exists and the two sets are identical. A

¹ This work was supported by an NSF Fellowship. The research is part of a doctoral thesis written under the direction of Professor John Lamperti at Dartmouth College.

similar remark holds for LAS and WLAS. We now turn to some results involving the above types of limits.

3. For the first result we need some additional terminology. Let $y(1), y(2), \dots$ be independent, identically distributed random variables. Let $m(n) = \max(y(1), \dots, y(n))$. Then we say that $m(n)$ is relatively stable (a) in probability if there exists a sequence $B(n) > 0$ such that $P(|m(n)/B(n) - 1| < \epsilon) \rightarrow 1$ for each $\epsilon > 0$; (b) almost surely if there exists a sequence $B(n) > 0$ such that $m(n)/B(n) \rightarrow 1$ almost surely. The study of the relative stability of the maximum of a sample was begun by Gnedenko [4] and has undergone development since that time.

THEOREM 1. *Let LIP(LAS) exist and contain at least two nonzero extreme points. Let $y(i) = \|x(i)\|$. Then $m(n)$ is relatively stable in probability (almost surely).*

The proof depends upon the fact that normalized sample points must be close to each extreme point of LIP(LAS). Whether or not this occurs will depend upon how the maxima in small cones about the extreme points grow. The growth in two disjoint cones is almost independent because the number of sample points in each cone is approximately the number of sample points times the probability that a given sample point falls into the given cone.

Let F be a distribution function. For each $x > 1$ let

$$L(x) = \min\{y \mid F(y - 0) \leq 1 - \frac{1}{x} \leq F(y + 0)\}.$$

If the maximum associated with a series of independent, identically distributed random variables with distribution function F , $F(0) < 1$, is relatively stable then we may choose $B(n) = L(n)$ and it also follows that L is a slowly varying function; that is, $L(cx)/L(x) \rightarrow 1$ as $x \rightarrow \infty$ for each $c > 0$.

THEOREM 2. *Let $B = E^n \equiv n$ -dimensional Euclidean space. Let m be the product measure of a radial measure and an angular measure. Let $\theta(m)$ be the support of the angular measure. If $\theta(m)$ contains at least two angles then LIP(LAS) exists iff $m(n)$ (defined in Theorem 1) is relatively stable in probability (almost surely). The limit is the convex hull of the set consisting of the origin and $\{(1, \theta) \mid \theta \in \theta(m)\}$.*

The proof uses Theorem 1 and the fact that L is a slowly varying function which allows us to show that the maximum radius in each open cone containing angles in $\theta(m)$ grows at the same rate.

Let $N(p) = \{(x, y) \mid x^p + y^p \leq 1 \text{ and } x \geq 0, y \geq 0\}$ for $1 \leq p < \infty$. Let

$$N(\infty) = \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

THEOREM 3. Let $B = E^2 = \{(x, y) \mid x, y \in R\}$. Let m be the product measure along the x and y axes of two identically distributed measures each with distribution function F . Let $F(x) = 0$ for $x < 0$ and $F(0) < 1$. Then if LAS or LIP exists it is equal to $N(p)$ for some $p \in [1, \infty]$ (up to normalization). We then have:

- (a) If LIP exists but not LAS, $p = 1$.
- (b) In order that $1 < p < \infty$ correspond to LAS it is necessary and sufficient that $L(y^{x^p}) \cong xL(y)$ for large y and each $x \in (0, 1)$.
- (c) In order that $p = \infty$ correspond to LAS it is necessary and sufficient that $L(y^{1/2}) > (1 - \epsilon)L(y)$ for large y and each $\epsilon > 0$.

The proof is somewhat involved although elementary methods are used. It involves showing that we may assume that $F(x)$ is continuous and strictly increasing for large x . We then can express the existence of the limit strictly in terms of L and manipulations with L complete the proof. From this result it is possible to find the most general limits in E^n which can arise from a product measure of one-dimensional measures along an orthogonal set of axes.

4. We now give three specific examples of limiting convex hulls.

EXAMPLE 1. Let m be a normal distribution in l_2 which is the product measure of $N(0, \sigma(i))$ measures along the orthogonal set of axes $\{x(i)\}$. (See Grenander [5] for more information on normal distributions in l_2 .) Then $\text{LAS} = \{(x(1), x(2), \dots) \mid \sum (x(i)/\sigma(i))^2 \leq 1\}$.

EXAMPLE 2. Let m be Poisson measure on K with the J_1 -topology (see Skorohod, [7] for definitions of K and the J_1 -topology). Then $\text{WLAS} = \{f \mid f(0) = 0, f(1) \leq 1, f \text{ is nondecreasing and continuous}\}$.

EXAMPLE 3. Let m be Wiener measure on $C_k[0, 1] \equiv$ the space of k -dimensional continuous functions defined on $[0, 1]$. Then $\text{LAS} = \{f \mid f(0) = 0, f \text{ is absolutely continuous and } \int_0^1 \dot{f}(t)^2 dt \leq 1\}$ where \dot{f}^2 is the inner product of \dot{f} with itself.

The first example is essentially Geffroy's result, but with a weaker type of convergence and an infinite dimensional space. The third example is inspired by a result of Strassen [8] and the proof is closely related to the proof that he gave for his somewhat similar result.

5. These results with detailed proofs will be published elsewhere at a later date. I wish to express my deep gratitude to my advisor, Professor Lamperti.

REFERENCES

1. Bradley Efron, *The convex hull of a random set of points*, Technical Report No. 103, Dept. of Statistics, Stanford University, Stanford, California, 1965.

2. Jean Geffroy, *Contribution à la théorie des valeurs extrêmes*, Publ. Inst. Statist. Univ. Paris 7 (1958), 36–123; 8 (1959) 3–52.
3. ———, *Localisation asymptotique du polyèdre d'appui d'un échantillon Laplacien à k dimensions*, Publ. Inst. Statist. Univ. Paris 10 (1961), 213–228.
4. B. V. Gnedenko, *Sur la distribution limite du terme maximum d'une série aléatoire*, Ann. of Math. 44 (1943), 423–453.
5. Ulf Grenander, *Probabilities on algebraic structures*, Wiley, New York, 1963.
6. A. Rényi and R. Sulanke, *Über die konvexe Hülle von n zufällig gewählten Punkten I and II*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 2 (1963), 75–84; 3 (1964), 138–148.
7. A. V. Skorohod, *Limit theorems for stochastic processes*, Theor. Probability Appl. 1 (1956), 261–290. (Engl. transl.)
8. V. Strassen, *An invariance principle for the law of the iterated logarithm*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 3 (1964), 211–226.

DARTMOUTH COLLEGE

A RELATION BETWEEN A THEOREM OF BOHR AND SIDON SETS

BY DANIEL RIDER¹

Communicated by W. Rudin, December 17, 1965

1. **Introduction.** In 1913, Bohr [1] proved the following theorem for Dirichlet series: if

$$(1) \quad f(\sigma + it) = \sum_{n=1}^{\infty} c(n)n^{-\sigma-it}$$

and if $|f(\sigma + it)| \leq 1$ for all $\sigma > 0$, then

$$(2) \quad \sum_p |c(p)| \leq 1,$$

the sum in (2) extending over all primes.

A set of positive integers E will be called a *Bohr set* if there is a finite constant B such that for every function f as in (1)

$$(3) \quad \sum_{n \in E} |c(n)| \leq B.$$

¹ This research was supported in part by Air Force Office of Scientific Research Grant AF-AOSR 335-63.