

$K \}$ , then  $N^*$  is a subgroup of  $D^*$  containing  $K^*$  as a normal subgroup. Since  $K^*$  is a cyclic group, we can choose a generator  $g$  for  $K^*$ . Then, for  $x \in D^*$ ,  $x \in N^*$  is equivalent to  $xgx^{-1} \in K$  and, for  $y \in N^*$ , we will have  $y \in Kx$  if and only if  $xgx^{-1} = ygy^{-1}$ . We let  $q = |N^*/K^*|$ .

Now  $D$  is a (left) vector space over  $K$  and, for  $a \in K$ ,  $D_a$  is a  $K$ -linear operator on  $D$ . From the above, we have

$$D_a^{p^m} = D_a^{p^m} = D_a.$$

Thus the minimal polynomial for  $D_a$  is a divisor of

$$\lambda^{p^m} - \lambda$$

and necessarily splits in  $K[\lambda]$  into distinct first degree factors so that  $D_a$  is diagonalizable. Applying this to  $D_g$ , for  $E_k = \{x: [g, x] = kx\}$ , we have  $E_0 = K$  and  $D = \sum E_k$ , where the sum is direct and taken over all  $k \in K$  with  $E_k \neq 0$ . Now, if  $x \in D^*$ ,  $gx - xg = kx$  for some  $k \in K$  is equivalent to requiring that  $x$  belong to  $N^*$ . Moreover,  $y \in E_k$  is equivalent to  $y \in Kx$ . Then each  $E_k$  is a  $K$ -subspace of dimension 1 and  $E_k^*$  is the coset  $K^*x$  in  $N^*$ . Hence,  $\dim_K D = q$ .

From the structure of finite fields it follows readily that  $K$  is a Galois extension of  $Z$ . We can identify  $N^*/K^*$  with a subgroup of  $G(K/Z)$  and, if  $J$  is the fixed field for  $N^*/K^*$ ,  $a \in J$  implies  $xax^{-1} = a$  for all  $x \in N^*$ . Then  $D_a$  is zero on each  $E_k$  so that  $D_a = 0$  and  $a \in Z$ . Hence  $N^*/K^* = G(K/Z)$ , implying  $\dim_Z K = |N^*/K^*| = q$ .

Combining the results above leads to  $\dim_Z D = (\dim_K D)(\dim_Z K) = q^2$  and  $\dim_Z C(b) = q$  for all  $b \in D$ ,  $b \notin Z$ . If  $r = |Z|$  then

$$|D^*| = r^{q^2} - 1$$

and  $|C(b)^*| = r^q - 1$  for  $b \notin Z$ . Thus, if  $s$  is the number of conjugacy classes containing more than one element, the class equation applied to  $D^*$  gives

$$r^{q^2} - 1 = (r - 1) + s(r^{q^2} - 1/r^q - 1),$$

which implies  $r^{q(q-1)} + \dots + r^q + 1$  must be a divisor of  $r - 1$ , giving the desired contradiction.

REFERENCES

1. I. N. Herstein Topics in Algebra, 2nd ed., Xerox, 1975.
2. \_\_\_\_\_, Wedderburn's theorem and a theorem of Jacobson, this MONTHLY, 68 (1961) 249-251.
3. \_\_\_\_\_, Non-Commutative Rings, MAA Carus Monograph No. 15, John Wiley and Sons, (1968).
4. N. Jacobson, Basic Algebra II, W. S. Freeman, 1980.
5. T. J. Kaczynski, Another proof of Wedderburn's theorem, this MONTHLY, 71 (1964) 652-653.

## Common Hyperplane Medians for Random Vectors

THEODORE P. HILL\*

*School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332*

Suppose some black and some white points are sprinkled randomly on the plane (e.g., salt and pepper on a table). Is there always a straight line so that at least half

---

\*Partially supported by NSF Grants DMS-84-01604 and DMS-86-01608.

of the black points and at least half of the white points lie on one side of the line (including the line), and at least half of each color lies on the other side (again including the line)? Steinhaus' "Ham Sandwich Theorem" (Corollary 1 below) says there is always a line bisecting the total *areas* occupied by both the black and the white spots, but does not answer the question for true (volumeless) points. The purpose of this note is to prove a generalization of the Ham Sandwich Theorem which includes *all* distributions: continuous, purely atomic, and mixed. If the black and white points represent statistical data or observations, then the positions where they fall are *random vectors*, and it is in this terminology that the theorem will be stated. (The two-dimensional example of black and white points on a plane was used only for illustrative purposes; what is essential is that there are no more than  $n$  colors in Euclidean  $n$ -space.)

A *median* of a random variable is a real number  $\alpha$  with the property that the probability the random variable is less than or equal to  $\alpha$ , and the probability that it is bigger than or equal to  $\alpha$ , are both at least one-half. One way of generalizing the notion of a median of a random variable to that of a median of a random vector is via hyperplanes.

DEFINITION. A *hyperplane median* for an  $n$ -dimensional random vector  $\vec{X} = (X_1, \dots, X_n)$  is an  $(n - 1)$ -dimensional hyperplane

$$\sum_{j=1}^n a_j x_j = b \text{ in } \mathbb{R}^n \text{ satisfying}$$

$$P\left(\sum_{j=1}^n a_j X_j \geq b\right) \geq 1/2 \text{ and } P\left(\sum_{j=1}^n a_j X_j \leq b\right) \geq 1/2.$$

Clearly every finite-dimensional random vector has a hyperplane median, and the main purpose of this note is to prove the following generalization of this fact.

THEOREM 1. *Every collection of  $m \leq n$  random  $n$ -dimensional vectors has a common hyperplane median.*

An immediate corollary is the Ham Sandwich Theorem, apparently first proved by Steinhaus (cf. [2], p. 291).

COROLLARY 1. (Ham Sandwich Theorem) *If  $B_1, \dots, B_n$  are bounded Lebesgue measurable sets in  $\mathbb{R}^n$ , then there is a hyperplane which simultaneously bisects the measure ( $n$ -dimensional volume) of each  $B_i$ .*

(The picturesque name Ham Sandwich Problem is also attributed to Steinhaus, who viewed it as the problem of simultaneously bisecting the ham, cheese, and bread in an ordinary ham sandwich by a single planar cut of a knife.)

*Proof of Corollary.* Without loss of generality, assume each  $B_i$  has strictly positive measure, and let  $\vec{Y}_i$  be the  $n$ -dimensional random vector which is uniformly distributed on  $B_i$ . Then apply Theorem 1.1 to  $\vec{Y}_1, \dots, \vec{Y}_n$ .  $\square$

The key idea in the proof of Theorem 1 is an application of the following theorem of Borsuk and Ulam [1]; unlike the case in the Ham Sandwich Theorem, however, the measures (distributions) involved are not assumed to be continuous (or even nonatomic), and the usual continuity argument fails.

**BORSUK-ULAM THEOREM.** *If  $f$  is a continuous map of the surface of the unit ball in  $(n + 1)$ -dimensional space into  $n$ -dimensional space such that  $f(-\vec{x}) = -f(\vec{x})$  for every  $\vec{x}$ , then there is some point on the ball mapped into the origin.*

Let  $\vec{X}$  be an  $n$ -dimensional random vector with distribution  $\mu$  (i.e.,  $\mu(A) = P(\vec{X} \in A)$  for every  $n$ -dimensional Borel set  $A$ ), let  $\vec{x} \cdot \vec{y}$  denote the usual inner product of  $\vec{x}$  and  $\vec{y}$ ; and let  $\|\vec{x}\|$  be the norm of  $\vec{x}$ .

**DEFINITION.** For  $\vec{v} \in \mathbb{R}^n$ ,  $m_*(\vec{v})$ ,  $m^*(\vec{v})$ , and  $m(\vec{v})$  are the lower, upper, and midpoint (respectively) medians for  $\vec{X}$  orthogonal to  $\vec{v}$ . That is,  $m_*$ ,  $m^*$ , and  $m$  are the functions from  $\mathbb{R}^n \rightarrow \mathbb{R}$  given by:

$$m_*(\vec{v}) = \inf \left\{ a : P(\vec{X} \cdot \vec{v} \leq a) \geq 1/2 \text{ and } P(\vec{X} \cdot \vec{v} \geq a) \geq 1/2 \right\};$$

$$m^*(\vec{v}) = \sup \left\{ a : P(\vec{X} \cdot \vec{v} \leq a) \geq 1/2 \text{ and } P(\vec{X} \cdot \vec{v} \geq a) \geq 1/2 \right\};$$

and

$$m(\vec{v}) = (m_*(\vec{v}) + m^*(\vec{v}))/2.$$

The proof of Theorem 1 will be based on two lemmas: the first lists some easy facts concerning  $m_*$ ,  $m^*$  and  $m$ ; and the second is the key continuity result.

**LEMMA 1.**

- (i)  $m_*(\vec{v})$ ,  $m^*(\vec{v})$ , and  $m(\vec{v})$  are finite and attained for all  $\vec{v}$ ;
- (ii)  $m_*(\vec{0}) = m^*(\vec{0}) = m(\vec{0}) = 0$ ;
- (iii)  $m(-\vec{v}) = -m(\vec{v})$ ; and
- (iv) if  $\vec{v} \neq \vec{0}$  and  $m(\vec{v}) = \alpha$ , then  $\vec{x} \cdot \vec{v} = \alpha$  is a hyperplane median for  $\vec{X}$ .

**LEMMA 2.** *If the support of  $\mu$  is compact, then  $m_*$ ,  $m^*$ , and  $m$  are continuous.*

*Proof.* Only the argument for the continuity of  $m_*$  will be given; that for  $m^*$  follows similarly, and together they imply the continuity of  $m$ .

Let  $K = \text{support of } \mu$ ; since  $K$  is compact there exists an  $M > 0$  such that  $\|\vec{x}\| \leq M$  for all  $\vec{x}$  in  $K$ .

Fix  $\vec{u} \in \mathbb{R}^n$  and  $\epsilon > 0$ ; then for every  $\vec{v}$  in  $\mathbb{R}^n$  satisfying  $\|\vec{u} - \vec{v}\| < \epsilon/M$ ,

$$|\vec{x} \cdot (\vec{u} - \vec{v})| < \epsilon \text{ for all } \vec{x} \text{ in } K. \tag{1}$$

Let  $\alpha = m_*(\vec{u})$ . Then

$$\begin{aligned} \{ \vec{x} \in K : \vec{x} \cdot \vec{v} \leq \alpha - 2\epsilon \} &\subseteq \{ \vec{x} \in K : \vec{x} \cdot \vec{u} \leq \alpha - \epsilon \} \\ &\subseteq \{ \vec{x} \in K : \vec{x} \cdot \vec{u} \leq \alpha \} \\ &\subseteq \{ \vec{x} \in K : \vec{x} \cdot \vec{v} \leq \alpha + \epsilon \}. \end{aligned} \tag{2}$$

(For the first inclusion in (2), observe that  $\vec{x} \cdot \vec{v} \leq \alpha - 2\epsilon$  implies  $\vec{x} \cdot \vec{u} \leq \alpha - 2\epsilon + \vec{x} \cdot (\vec{u} - \vec{v})$  which by (1) implies  $\vec{x} \cdot \vec{u} \leq \alpha - \epsilon$ . The second inclusion is trivial and the third is analogous to the first.)

Since  $\mu(K) = 1$ , it follows from (2) and the definition of  $\alpha$  that

$$\begin{aligned} P(\vec{X} \cdot \vec{v} \leq \alpha - 2\epsilon) &\leq P(\vec{X} \cdot \vec{u} \leq \alpha - \epsilon) < 1/2 \\ &\leq P(\vec{X} \cdot \vec{u} \leq \alpha) \leq P(\vec{X} \cdot \vec{v} \leq \alpha + \epsilon). \end{aligned} \tag{3}$$

From (3) it follows that  $\alpha - 2\epsilon \leq m_*(\vec{v}) \leq \alpha + \epsilon$ , so since  $\epsilon$  was arbitrary and  $\alpha = m_*(\vec{u})$ , this establishes the continuity of  $m_*$ .  $\square$

The conclusion of Lemma 2 may fail if the support of  $\mu$  is not compact, even if  $\mu$  is continuous (i.e., absolutely continuous with respect to Lebesgue measure).

*Example 1.* Suppose  $\vec{X}$  is a continuous 2-dimensional random vector with strictly positive density of total mass  $1/2$  on  $A = \{(x, y): 1 \leq x \leq 2, y \in \mathbb{R}\}$  and uniformly distributed with total mass  $1/2$  on  $B = \{(x, y): 3 \leq x \leq 4, -1 \leq y \leq 1\}$ . It is easy to check that  $m((1,0)) = 5/2$ , but that  $\lim_{n \rightarrow \infty} m((1, n^{-1})) = 3$ .

*Proof of Theorem 1.* Fix  $m \leq n$ , let  $\vec{X}_1, \dots, \vec{X}_m$  be  $n$ -dimensional random vectors with distributions  $\mu_1, \dots, \mu_m$  respectively, and let  $m_i$  denote the midpoint-median function for  $\vec{X}_i$  (as in Definition 2.1). Without loss of generality,  $m = n$ .

*Case 1.* The support of each  $\mu_i, i = 1, \dots, n$ , is compact.

Let  $f$  be the function from the surface of the unit ball in  $\mathbb{R}^{n+1}$  into  $\mathbb{R}^n$  given by  $f = (f_1, \dots, f_n)$ , where  $f_i(\vec{v}, \beta) = m_i(\vec{v}) - \beta$  for all  $i = 1, \dots, n$ .

From Lemma 2 it follows easily that  $f$  is continuous, and from Lemma 1(iii) that  $f(-\vec{v}, -\beta) = -f(\vec{v}, \beta)$ , so the Borsuk-Ulam theorem implies there exists a point  $(\vec{v}_0, \beta_0)$  on the surface of the unit ball in  $\mathbb{R}^{n+1}$  satisfying  $f(\vec{v}_0, \beta_0) = \vec{0}$ , that is,

$$m_i(\vec{v}_0) = \beta_0 \text{ for all } i = 1, \dots, n. \tag{4}$$

By Lemma 1(ii) and the fact that  $\|(\vec{v}_0, \beta_0)\| = 1$  it follows that  $\vec{v}_0 \neq 0$ , so the existence of a common median for  $\vec{X}_1, \dots, \vec{X}_n$  follows by (4) and Lemma 1(iv). This completes the argument for Case 1.

*General Case.* Let  $r > 0$  be such that  $\mu_i(\{\vec{x}: \|\vec{x}\| \leq r\}) > 0$  for all  $i = 1, \dots, n$ . For  $j = 1, 2, \dots$ , let  $K_j$  denote the closed sphere  $\{\vec{x}: \|\vec{x}\| \leq r + j\}$  in  $\mathbb{R}^n$ , and let  $\mu_{i,j}$  denote the (re-normalized) restriction of  $\mu_i$  to  $K_j$ , for  $i = 1, \dots, n$ . By Case 1, for each  $j \geq 1$  there is a common median hyperplane  $H_j = \{\vec{x} \in \mathbb{R}^n: \vec{a}_j \cdot \vec{x} = b_j\}$  for  $\mu_{1,j}, \dots, \mu_{n,j}$ . Since  $\mu_i(K_j) \rightarrow 1$  as  $j \rightarrow \infty$ , for all  $i = 1, \dots, n$ , it follows easily that a subsequence  $\{H_{j_k}\}$  of  $\{H_j\}$  converges (that is, there is an  $\vec{a} \in \mathbb{R}^n$  and  $b \in \mathbb{R}$  so that  $\vec{a}_{j_k} \rightarrow \vec{a}$  and  $b_{j_k} \rightarrow b$ ), and that the limiting hyperplane  $H = \{\vec{x}: \vec{a} \cdot \vec{x} = b\}$  is a common median for  $\vec{X}_1, \dots, \vec{X}_n$ .  $\square$

A generalization of the notion of “median” that is often of use in statistics is that of a  $\beta$ -quantile, which is a number  $\alpha$  with the property that the probability the random variable is  $\leq \alpha$  is at least  $\beta$ , and the probability that it is  $\geq \alpha$  is at least  $1 - \beta$ . Clearly every random variable has at least one  $\beta$ -quantile for each  $\beta \in (0,1)$ , and it is natural to ask whether Theorem 1 can be generalized to conclude that every collection of  $m \leq n$  random variables has a common hyperplane  $\beta$ -quantile. The next example shows that in general the *only* common hyperplane  $\beta$ -quantiles which exist are for  $\beta = 1/2$ , i.e., medians.

*Example 2.* Let  $\vec{X}_1$  be the 2-dimensional random vector which is uniformly distributed on the boundary of the unit circle in  $\mathbb{R}^2$ , and let  $\vec{X}_2$  be the constant 2-dimensional random vector which is always the origin. It is easy to see that there are no common hyperplane  $\beta$ -quantiles for  $\vec{X}_1$  and  $\vec{X}_2$  for *any*  $\beta \in (0,1)$  except  $\beta = 1/2$ .

(Note that for a fixed  $\beta \in (0,1)$ ,  $\beta \neq 1/2$ , the above example can be modified slightly to yield *continuous* distributions with no common hyperplane  $\beta$ -quantile).

As pointed out by Carl Spruill, Theorem 1 can also be proved from the Ham Sandwich Theorem by adding  $N(0, n^{-1})$  random variables to each component to

make each  $\vec{X}_i$  (and hence each  $\mu_i$ ) continuous, applying for each  $n$  a slightly generalized version of the Ham Sandwich Theorem (for arbitrary continuous probability distributions), and then appealing to standard results for weak convergence of probability measures. The proof presented above has the advantage that it is more elementary (since it relies essentially only on the Borsuk-Ulam Theorem), and it also shows how to modify directly the classical proof of the Ham Sandwich Theorem to include noncontinuous distributions.

It should also be emphasized that both proofs of Theorem 1 (the given proof, and Spruill's proof) are highly nonconstructive, and it would perhaps be of interest to devise a practical, constructive method for determining the common hyperplane medians guaranteed by the theorem. Spruill has also raised the question of which curves or manifolds other than straight lines or hyperplanes can serve as common medians.

#### REFERENCES

1. K. Borsuk, Drei Sätze über die  $n$ -dimensionale euklidische Sphäre, *Fund. Math.*, 20 (1933) 177–190.
2. M. Agoston, *Algebraic Topology*, Marcel Dekker, New York, 1976.