A Sharp Partitioning-Inequality for Non-Atomic Probability Measures Based on the Mass of the Infimum of the Measures

Theodore P. Hill

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

Summary. If $\mu_1, \ldots, \mu_n$ are non-atomic probability measures on the same measurable space $(S, \mathscr{F})$, then there is an $\mathscr{F}$-measurable partition $\{A_i\}_{i=1}^n$ of S so that $\mu_i(A_i) \geqq (n-1+m)^{-1}$ for all $i=1, \ldots, n$, where $m = \left\| \bigwedge_{i=1}^n \mu_i \right\|$ is the total mass of the largest measure dominated by each of the μ_i's; moreover, this bound is attained for all $n \geqq 1$ and all m in $[0,1]$. This result is an analog of the bound $(n+1-M)^{-1}$ of Elton et al. [5] based on the mass M of the supremum of the measures; each gives a quantative generalization of a well-known cake-cutting inequality of Urbanik [10] and of Dubins and Spanier [2].

§ 1. Introduction

Suppose $\mu_1, \mu_2, \ldots, \mu_n$ are non-atomic probability measures on the same measurable space $(S, \mathscr{F})$, and let

$$m = \left\| \bigwedge_{i=1}^n \mu_i \right\|$$

denote the total mass of the sub-probability measure $\bigwedge_{i=1}^n \mu_i$, the largest measure dominated by each of the measures μ_i. The main purpose of this note is to prove the following result.

Theorem 1. *If $\mu_1, \ldots, \mu_n$ are non-atomic probability measures on the same measurable space $(S, \mathscr{F})$, then there is an $\mathscr{F}$-measurable partition $\{A_i\}_{i=1}^n$ of S satisfying*

$$\mu_i(A_i) \geqq (n-1+m)^{-1} \quad \textit{for all} \quad i=1, \ldots, n; \tag{1}$$

moreover, this bound is attained for all positive integers n and all $m \in [0, 1]$.

Theorem 1 is a direct analog of a sharp partitioning result of Elton et al. [5] based on the total mass $M=\left\|\bigvee_{i=1}^{n} \mu_i\right\|$ of the smallest measure $\bigvee_{i=1}^{n} \mu_i$ dominating each of the μ_i's; namely, the existence of a measurable partition $\{A_i\}_{i=1}^{n}$ of S satisfying

$$\mu_i(A_i) \geqq (n+1-M)^{-1} \quad \text{for all } i=1, \ldots, n. \tag{2}$$

It is easy to see that for $n=2$, both inequalities (1) and (2) are identical (since in that case $m+M=2$), but that for $n>2$ neither implies the other. Since

$$m=1 \Leftrightarrow M=1 \Leftrightarrow \mu_i=\mu_j \quad \text{for all } i, j=1, \ldots, n$$

both inequalities (1) and (2) give *quantitative* generalizations of a well-known "cake-cutting" result of Urbanik [10] and of Dubins and Spanier [2] which state that if $\mu_i \neq \mu_j$ for some $i \neq j$, then there is a measurable partition $\{A_i\}_{i=1}^{n}$ of S satisfying

$$\mu_i(A_i) > n^{-1} \quad \text{for all } i=1, \ldots, n. \tag{3}$$

(In the cake-cutting interpretation of these inequalities, S represents a cake which must be divided among n people, and $\mu_i(A)$ represents the value of piece A to person i; the reader is referred to [2] or [5] for more details.)

§ 2. Proof of Main Theorem

Since the conclusion of Theorem 1 is trivial if $n=1$, assume $n>1$. Throughout this section, Π_k denotes the collection of $\mathscr{F}$-measurable k-partitions of S, that is

$$\Pi_k=\{\{A_i\}_{i=1}^{k}: A_i \cap A_j=\emptyset \text{ if } i \neq j, \text{ and } A_i \in \mathscr{F}\ \forall\, i=1, \ldots, k\},$$

and $\vec{\mu}=(\mu_1, \ldots, \mu_n)$ is an n-dimensional vector-valued measure each of whose coordinates is a non-atomic (non-negative, countably additive) finite measure.

Let $PR(\vec{\mu})$ denote the *partition-range* of $\vec{\mu}$, that is,

$$PR(\vec{\mu})=\{(a_1, \ldots, a_n) \in \mathbb{R}^n: \exists\, \{A_i\}_{i=1}^{n} \in \Pi_n \text{ with } \mu_i(A_i)=a_i\ \forall\, i=1, \ldots, n\}.$$

Two of the tools in the proof of Theorem 1 are a generalization of Lyapounov's Convexity Theorem due to Dvoretzky et al. [3] and an application of the convexity theorem by Neyman which solved Fisher's "Problem of the Nile"; both results are recorded here for convenience, and the reader is referred to [2] for more details concerning these and related results.

Lemma 2.1 ([3]). *$PR(\vec{\mu})$ is convex and compact.*

Lemma 2.2 ([9]). *For each positive integer k, there exists a measurable partition $\{E_i\}_{i=1}^{k}$ of S satisfying*

$$\mu_j(E_i)=k^{-1} \mu_j(S) \quad \textit{for all } j=1, \ldots, n \textit{ and } i=1, \ldots, k.$$

The other main tool in the proof is an "inversion principle", which allows any small equipartition value t to be transformed into a new large value t', and vice versa. For the remainder of this paper $\mu_1, \ldots, \mu_n$ are probability measures, and $\vec{1}=(1,1,\ldots,1)$.

Proposition 2.3 (Inversion Principle).

$$\vec{a}\in PR(\mu)\Rightarrow \vec{a}'=(\vec{1}-\vec{a})/(n-1)\in PR(\vec{\mu}).$$

Proof. Fix $\vec{a}=(a_1, \ldots, a_n)\in PR(\vec{\mu})$, and let $\{A_i\}_{i=1}^n$ be any element (partition) in Π_n with $\vec{a}=(\mu_1(A_1), \ldots, \mu_n(A_n))$.

For each $i=1, \ldots, n$, Lemma 2.2 (with $k=n-1$ and $S=A_i$) implies the existence of an $\mathscr{F}$-measurable $(n-1)$-partition $\{A_{i,k}\}_{k=1, k\neq i}^n$ of A_i satisfying

$$\mu_j(A_{i,k})=(n-1)^{-1}\mu_j(A_i) \qquad \forall j=1, \ldots, n,\ \forall k\neq i,\ k=1, \ldots, n.$$

Letting $B_j=\bigcup\{A_{i,j}: i\neq j,\ 1\leqq i\leqq n\}$, it follows easily that

$$\mu_j(B_j)=(n-1)^{-1}(1-a_j) \qquad \text{for each } j=1, \ldots, n.$$

Since $\{B_j\}_{j=1}^n\in\Pi_n$, this implies that $(\vec{1}-\vec{a})/(n-1)\in PR(\vec{\mu})$. □

Note that $\vec{a}$ small implies $\vec{a}'$ is large (and vice versa) and that $\vec{a}''\neq\vec{a}$. The useful aspect in this paper is that in general $\vec{a}'$ lies outside the convex hull of $\vec{a}$ and the unit coordinate vectors.

Proof of Theorem 1
Letting $\vec{e}_i$ denote the i^{th} unit coordinate vector $(0, \ldots, 0, 1, 0, \ldots, 0)$ of $\mathbb{R}^n$, it is clear (taking $A_i=S$, $A_j=\emptyset$ for $j\neq i$) that $\vec{e}_i\in PR(\vec{\mu})$ for all $i=1, \ldots, n$.

Fix $\vec{a}=(a_1, \ldots, a_n)\in PR(\vec{\mu})$, and let

$$\beta_i=a_i\left(n-1+\sum_{j=1}^n a_j\right)^{-1} \qquad \text{for } i=1, \ldots, n;$$

and

$$\beta_{n+1}=(n-1)\left(n-1+\sum_{i=1}^n a_i\right)^{-1}.$$

Together, Lemma 2.1 (convexity) and Proposition 2.3 imply that

$$\vec{v}=\sum_{i=1}^n \beta_i\vec{e}_i+\beta_{n+1}(\vec{1}-\vec{a})/(n-1)\in PR(\vec{\mu}),$$

and an easy calculation shows that $\vec{v}=(\alpha, \alpha, \ldots, \alpha)\in\mathbb{R}^n$, where $\alpha=\left(n-1+\sum_{i=1}^n a_i\right)^{-1}$.

Choosing $\{A_i\}_{i=1}^n$ so that $\sum_{i=1}^n a_i=m+\varepsilon$ (using the compactness conclusion of Lemma 2.1, ε may even be taken to be zero), establishes the inequality (1). That this bound is attained follows from the next example. □

Example 2.4. For $n>1$ and $m\in[0,1]$, let $(S,\mathscr{F})=([0,1]$, Borels), let $f_1\colon[0,1]\to\mathbb{R}$ be $2I_{[0,1/2)}$ and $f_i=2mI_{[0,1/2)}+2(1-m)I_{[1/2,1]}$ for $i=2,\ldots,n$, and define $\mu_1,\ldots,\mu_n$ on $(S,\mathscr{F})$ by $\mu_i(A)=\int_A f_i\,d\lambda$. Then $\mu_1,\ldots,\mu_n$ are non-atomic probability measures on $(S,\mathscr{F})$ with $\left\|\bigwedge_{i=1}^n \mu_i\right\|=m$, and an easy calculation shows that for *every* $\mathscr{F}$-partition $\{A_i\}_{i=1}^n$ of $[0,1]$, $\min_{i\leq n}\mu_i(A_i)\leq(n-1+m)^{-1}$, and that in fact this bound is attained.

Remarks. The idea to use the Dvoretzky-Wald-Wolfowitz result to establish (1) was triggered by a recent proof of (2) by Legut [8] using that same result; Dubins and Spanier [2] used a similar generalization (matrix convexity) to establish (3).

The extremal case $m=0$ in (1) is not completely analogous to the extremal case $M=n$ in (2), since $M=n$ implies the measures have essentially disjoint support and that the optimal-partitioning constant is *always* 1 in that case, whereas in the $m=0$ case the optimal-partitioning constant may be strictly bigger than $(n-1)^{-1}$ for some $\vec{\mu}$.

The inversion principle and the bound $(n-1+\sum a_i)^{-1}$ in the proof may be of some use when m is not known or easy to calculate, but instead only several partition-vectors $(a_1,\ldots,a_n)\in PR(\vec{\mu})$ are known.

If the measures have atoms, convexity and all the inequalities (1)-(3) may fail; analogs of the convexity theorem and (3) based on the maximum atom mass are contained in [4] and [7]. Similarly, if the measures are no longer assumed to be probability measures, again (1)-(3) may fail; [6] contains an analog of (3) based on the total masses of the measures (the constant n^{-1} is replaced by n^{-1} times the harmonic mean of the total masses of the measures).

§3. Applications

In the classification problem of statistical decision theory, the *minimax risk* $R(\mu_1,\ldots,\mu_n)$ of probability distributions $\mu_1,\ldots,\mu_n$ can also be expressed (see [5]) as

$$R(\mu_1,\ldots,\mu_n)=1-\sup\{\min_{i\leq n}\mu_i(A_i)\colon\{A_i\}_{i=1}^n\in\Pi_n\},$$

so Theorem 1 has the following immediate consequence.

Corollary 1. *If $\mu_1,\ldots,\mu_n$ are non-atomic probability distributions, then the minimax risk in the corresponding classification problem is at most $(n-2+m)/(n-1+m)$, and this bound is best possible.*

In [1] and [5], an application of (2) was made to the problem of distributing k indivisible objects to n people via lotteries, and a similar application can be made of (1).

Acknowledgement. The author is grateful to the referee for suggesting emphasis of the inversion principle, and for several other suggestions and comments.

References

1. Demko, S., Hill, T.: Equitable distribution of indivisible objects. Preprint (1985)
2. Dubins, L., Spanier, E.: How to cut a cake fairly. Am. Math. Monthly **68**, 1–17 (1961)
3. Dvoretzky, A., Wald, A., Wolfowitz, J.: Relations among certain ranges of vector measures. Pacific J. Math. **1**, 59–74 (1951)
4. Elton, J., Hill, T.: A generalization of Lyapounov's convexity theorem to measures with atoms. Proc. Am. Math. Soc., **99**, 297–304 (1987)
5. Elton, J., Hill, T., Kertz, R.: Optimal-partitioning inequalities for non-atomic probability measures. Trans. Am. Math. Soc. **296**, 703–725 (1986)
6. Hill, T.: Equipartitioning the common domain of non-atomic measures. Math. Z. **189**, 415–419 (1985)
7. Hill, T.: Partitioning general probability measures. Ann. Prob., to appear (1987)
8. Legut, J.: Inequalities for α-optimal partitioning of measurable space. Math. Z., to appear (1987)
9. Neyman, J.: Un théorème d'existence. C.R. Acad. Sci. Paris 222, 843–845 (1946)
10. Urbanik, K.: Quelques théorèmes sur les mesures. Fund. Math. **41**, 150–162 (1955)