Internal. J. Math. & Math. Sci.
VOL. 13 NO. 4 (1990) 825-827

ON THE DISCREPANCY OF COLORING FINITE SETS

D. HAJELA

Bellcore
445 South Street
Morristown, New Jersey 07960

(Received October 11, 1989)

ABSTRACT. Given a subset S of $\{1,\ldots,n\}$ and a map $X:\{1,\ldots,n\} \to \{-1,1\}$, (i.e. a coloring of $\{1,\ldots,n\}$ with two colors, say red and blue) define the discrepancy of S with respect to X to be $d_X(S) = \left|\sum_{i\varepsilon S} X(i)\right|$ (the difference between the reds and blues on S). Given n subsets of $\{1,\ldots,n\}$, a question of Erdos was to find a coloring of $\{1,\ldots,n\}$ which simultaneously minimized the discrepancy of the n subsets. We give new and simple proofs of some of the results obtained previously on this problem via an inequality for vectors.

KEY WORDS AND PHRASES. Discrepancy problems, colorings of finite sets.
1980 AMS SUBJECT CLASSIFICATION CODES. Primary 05A05, Secondary 05C35.

1. INTRODUCTION.

The problem of Erdos that we are interested in this note can be formulated as follows: For a subset S of $\{1,\ldots,n\}$ and a map $X:\{1,\ldots,n\} \to \{-1,1\}$, i.e. a coloring of $\{1,\ldots,n\}$ with two colors, say red and blue, define the discrepancy of S with respect to X to be $d_X(S) = \left|\sum_{i\varepsilon S} X(i)\right|$, which is the difference between the reds and blues on S. Given subsets $S_1,\ldots,S_k$ of $\{1,\ldots,n\}$, the problem is to find a coloring X which minimizes $d_X(S_i)$ simultaneously for $i=1,\ldots,k$.

A simple reduction (see [1], [3]) shows that we may assume $k \leq n$ in the above, and so for the rest of the note we shall consider the case of k=n, i.e. n subsets of $\{1,\ldots,n\}$. It is easy to see [3] that there are subsets $S_1,\ldots,S_n$ such that for every coloring X, $\max_{i\leq n} d_X(S_i) \geq c\sqrt{n}$ for some constant $c > 0$. In the other direction, there have been a number of results [1], [2], [3], with the best possible result being obtained in [4] (recently, a generalization and an improvement in some cases of the results in [4] has been obtained in [6]):

$$\min_X \max_{i\leq n} d_X(S_i) \leq c\sqrt{n} \tag{1.1}$$

for some $c > 0$. Most of these results are obtained in a simple fashion from certain probabilistic inequalities, with the exception of the result (1.1) in [4] (and the results in [6]), which are more involved.

We give a very simple proof of some of the results obtained prior to (1.1), in [1], [2], [3]. The proof is based on a geometrical inequality for vectors. The use of this inequality in the context of combinatorial problems appears to be new. This, and the fact that the inequality also seems to be of use for other combinatorial problems, serves as the motivation for this note.

2. MAIN RESULTS.

In order to state the inequality, recall that for a vector $x = (x_1,\dots,x_n)$ in R^n:

$$||x||_p = \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} & \text{for } 1 \leq p < \infty \\ \max_{i \leq n} |x_i| & \text{for } p = \infty \end{cases}$$

The desired inequality, is a special case of a general class of inequalities which can be found in [5]. For any vectors $v_1,\dots,v_n$ in R^n and any $2 \leq p < \infty$,

$$\left(\frac{1}{2^n} \sum_X ||X(1)v_1 + \dots + X(n)v_n||_p^p\right)^{1/p} \leq \sqrt{p} \left(\sum_{i=1}^{n} ||v_i||_p^2\right)^{1/2} \tag{2.1}$$

Here the sum over is over all colorings $X: \{1,\dots,n\} \to \{-1,1\}$ (in (2.1), for a vector x, $||x||_p^p$ is of course simply $\sum_{i=1}^{n} |x_i|^p$). The theorem we prove is:

THEOREM. (a) Let $v_1,\dots,v_n$ be vectors in R^n with $||v_i||_\infty \leq 1$ for $i=1,\dots,n$. Then for any $2 \leq p < \infty$,

$$\left(\frac{1}{2^n} \sum_X ||X(1)v_1+\dots+X(n)v_n||_p^p\right)^{1/p} \leq \sqrt{p}\; n^{1/2+1/p} \tag{2.2}$$

where the sum over X is over all colorings X.

(b) For any $S_1,\dots,S_n$, which are subsets of $\{1,\dots,n\}$ and any $2 \leq p < \infty$

$$\left(\frac{1}{2^n} \sum_X \sum_{i=1}^{n} d_X(S_i)^p\right)^{1/p} \leq \sqrt{p}\; n^{1/2+1/p} \tag{2.3}$$

(c) For any $S_1,\dots,S_n$, which are subsets of $\{1,\dots,n\}$,

$$\left(\frac{1}{2^n} \sum_X (\max_{i \leq n} d_X(S_i))^p\right)^{1/p} \leq \sqrt{2\, en \log n} \tag{2.4}$$

for $1 \leq p \leq 2 \log n$.

In particular (2.3) in the above theorem implies that there is a X (since (2.3) averages over all X) such that

$$\left(\sum_{i=1}^{n} d_X(S_i)^p\right)^{1/p} < \sqrt{p}\ n^{1/2+1/p} \tag{2.5}$$

Inequality (2.5) may be found in [3] with $\sqrt{p}$ replaced by a non-explicit function of p. Also (2.4), with p = 1 implies that there is a coloring X so that

$$\max_{i<n} d_X(S_i) < (2en \log n)^{1/2} \tag{2.6}$$

(2.6) can be found in [1] and [2] with a different constant instead of $\sqrt{2e}$. The proof of the above theorem follows at once from (2.1).

PROOF. (a) Since $v_i = (v_{i1},\ldots,v_{in})$ has $||v_i||_\infty < 1$, it follows that $|v_{ij}| < 1$ for all $1 < i,\ j < n$. Thus, $||v_i||_p < n^{1/p}$ and the result follows at once from (2.1).

(b) To the sets S_i, $i=1,\ldots,n$ we associate the incidence vectors v_j, $j=1,\ldots,n$ where $v_j(i) = 1$ if $j \in S_i$ and 0 otherwise for $i=1,\ldots,n$. Note that for any X,

$$||X(1)v_1 + \ldots + X(n)v_n||_p^p = \sum_{i=1}^{n} d_X(S_i)^p \tag{2.7}$$

and that $||v_i||_\infty < 1$ for all i. The result now follows from (2.2).

(c) From (2.3) we have

$$\left(\frac{1}{2^n} \sum_X (\max_{i<n} d_X(S_i))^p\right)^{1/p} < \sqrt{p}\ n^{1/2+1/p} \tag{2.8}$$

Notice that the left hand side of (2.8) is an increasing function of p. Hence letting $p_o(n) = 2 \log n$, which minimizes $\sqrt{p}\ n^{1/2+1/p}$, we have for $p < p_o(n)$ that

$$\left(\frac{1}{2^n} \sum_X (\max_{i<n} d_X(S_i))^p\right)^{1/p} < \sqrt{p_o}\, n^{1/2+1/p_o} = \sqrt{2e}\ \sqrt{n \log n}.$$

REMARK. Notice that the results in (b) and (c) of the above theorem can be imporved if we have some idea of the structure of the S_i, since then we can estimate their incidence vectors v_i better.

REFERENCES

1. BECK, J. and FIALA, T. Integer Making Theorems, Discr. Appl. Math. 3 (1981), 1-8.

2. OLSON, J. and SPENCER, J. Balancing Familites of Sets, J. Combinatorial Theory, Ser. A. 25 (1978), 29-37.

3. SPENCER, J. Discrete Ham Sandwich Theorems, Europ. J. Combinatorics 2 (1981), 291-298.

4. SPENCER, J. Six Standard Deviations Suffice, Trans. Amer. Math. Soc. 289 (1985), 679-706.

5. WOYCZINSKI, W. Geometry and Martingales in Banach Spaces; Part II: Independent Increments, in Probabilitity on Banach Spaces, Marcel Dekker, 1978.