HOW MUCH ARE INCREASING SETS POSITIVELY CORRELATED?

MICHEL TALAGRAND*

Received July 18, 1994

Two upward directed sets of sequences of zeroes and ones are positively correlated. We provide a lower bound on the correlation, in function of how much the two sets simultaneously depend on the same coordinates.

I. Introduction

Consider a set $A$ of sequences of length $n$ of zeroes and ones, that is, $A \subseteq \{0,1\}^n$. We say that $A$ is increasing (or upward directed) if

$$x \in A, \forall i \leq n, y_i \geq x_i \Rightarrow y \in A.$$ 

Given two increasing subsets $A, B$ of $\{0,1\}^n$, a well known result asserts that

$$2^n \text{card } (A \cap B) \geq \text{card } A \text{card } B$$

or, equivalently

$$\mu(A \cap B) \geq \mu(A) \mu(B)$$

(1.1)

where $\mu(A) = 2^{-n} \text{card } A$ is the normalized counting measure on $\{0,1\}^n$.

For $x \in \{0,1\}^n$, and $i \leq n$, consider the sequence $T_i x$ obtained from $x$ by changing the $i^{th}$-coordinate. For a set $A \subseteq \{0,1\}^n$, we set

$$A_i = \{x \in A; T_i x \notin A \}.$$ 

The quantity $\mu(A_i)$ expresses "how much $A$ depends on the $i^{th}$ coordinate." In particular if $\mu(A_i) = 0$, then $A$ does not depend on the $i^{th}$ coordinate in the sense

* Work partially supported by an NSF grant.
that \( x \in A \) if and only if \( T_i x \in A \). We should also observe that if \( A \) is increasing then \( x \in A_i \Rightarrow x_i = 1 \).

The usual proof of (1.1) by induction upon the number of coordinates shows that equality holds in (1.1) if and only if for each coordinate \( i \), either \( A \) or \( B \) does not depend on \( i \). A natural question, which is the object of the present paper, is to find a quantitative version of this fact, that is to find a lower bound of \( \mu (A \cap B) - \mu (A) \mu (B) \) in function of "how much \( A \) and \( B \) simultaneously depend on the same coordinates." Certainly there are many conceivable ways to quantify this, but, since \( A \) and \( B \) simultaneously depend on coordinate \( i \) if and only if \( \mu (A_i) \mu (B_i) > 0 \), it is natural to introduce these quantities. We will prove the following.

**Theorem 1.1.** Consider, for \( 0 \leq x \leq 1 \), the function \( \varphi (x) = x / \log (e / x) \). Then, for some universal constant \( K \), for all \( n \) and all increasing sets \( A, B \subset \{0,1\}^n \), we have

\[
\mu (A \cap B) - \mu (A) \mu (B) \geq \frac{1}{K} \varphi \left( \sum_{i \leq n} \mu (A_i) \mu (B_i) \right).
\]

It would be of course very difficult to have an exact expression for the difference \( \mu (A \cap B) - \mu (A) \mu (B) \). But we will show that in a case of special importance, the lower bound given by (1.2) is of correct order. Consider an integer \( k \geq n/2 \), and set

\[
A = \left\{ (x_i); \sum_{i \leq n} x_i \geq n - k \right\}; \quad B = \left\{ (x_i); \sum_{i \leq n} x_i > k \right\}.
\]

Thus \( \mu (A) + \mu (B) = 1 \). Set \( \varepsilon = \mu (B) \). Since \( B \subset A \), we have

\[
\mu (A \cap B) - \mu (A) \mu (B) = \varepsilon - \varepsilon (1 - \varepsilon) = \varepsilon^2.
\]

On the other hand computation (or the arguments of Section 2) show that \( \sum_{i \leq n} \mu (A_i) \mu (B_i) \) is of order \( \varepsilon^2 \log (1 / \varepsilon) \). And \( \varphi (\varepsilon^2 \log (1 / \varepsilon)) \) is of order \( \varepsilon^2 \), so that (1.2) is sharp in this case. This should be the moment to point out that the weaker inequality

\[
\mu (A \cap B) - \mu (A) \mu (B) \geq \sum_{i \leq n} \varphi (\mu (A_i) \mu (B_i))
\]

that is considerably easier to prove than (1.2), would not be sharp on the example above.

The proof of (1.2) will be by induction over the number of coordinates. However, before that can be done a new fact has to be proved about increasing sets. While somewhat technical, this new fact is the heart of the paper, and is better explained as part of a circle of ideas.