Transactions of the Fourth Prague Conference

on Information Theory, Statistical Decision Functions, Random Processes
Held at Prague, from August 31 to September 11, 1965

Behaviour of Sequential Predictors of Binary Sequences

T. M. Cover

Published by the Publishing House of the Czechoslovak Academy of Sciences, Prague
1967

1 Introduction

2 Deterministic Predictors

Consider the set of 2^n sequences $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n) \in \{0,1\}^n$.
At stage k, after the observation $\Theta_1, \Theta_2, \ldots, \Theta_{k-1}$, the prediction 1 or 0 will be made with probability p_k and $1 - p_k$ respectively.

Definition 2.1 (Sequential Predictor)
A sequential predictor on $\{0,1\}^n$ will be completely specified by the set of functions

$$p_1, p_2(\Theta_1), p_3(\Theta_1, \Theta_2), \ldots, p_n(\Theta_1, \Theta_2, \ldots, \Theta_{n-1})$$

taking values in $[0,1]$.

- *If the p_is are restricted to $\{0,1\}$, the predictor is called a **deterministic predictor**.*
- *If the p_is are independent of the Θs, the predictor is called a **memoryless predictor**.*
- *If the p_is are also independent of i, the predictor is called a **constant/time invariant predictor**.*

Let $\delta = (\delta_1, \delta_2, \ldots, \delta_n) \in \{0,1\}^n$ be the sequence of R.V.s resulting from the predictor $p = (p_1, p_2, \ldots, p_n)$ and the sequence $\Theta \in \{0,1\}^n$.
Then the empirical average score (the fraction of correct predictions) is given by

$$s = \frac{1}{n}\sum_{i=1}^{n}[\delta_i\Theta_i + (1-\delta_i)(1-\Theta_i)] \tag{2.1}$$

and the expected empirical average score is given by

$$\bar{s} = \mathbb{E}_p(s) = \frac{1}{n}\sum_{i=1}^{n}[p_i\Theta_i + (1-p_i)(1-\Theta_i)] \tag{2.2}$$

Theorem 2.1
Any sequential deterministic predicator attains a score of $\frac{k}{n}$ on precisely $\binom{n}{k}$ sequences in $\{0,1\}^n$ where $k \in [n]$. For any deterministic predictor, there exists a sequence upon which a score of 0 is attained.

3 Sequential Betting Systems

3.1 Achievable Winnings in Sequential Betting

A series of n bets $b = (b_1, b_2, \ldots, b_n)$ is made by a gambler on the outcomes of a sequence $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n) \in \{0,1\}^n$. The gambler's net gain at bet k is b_k if $\Theta_k = 1$ and $-b_k$ if $\Theta_k = 0$. Hence, his net winnings $w(\Theta)$ using strategy b against sequence Θ is

$$w(\Theta) = \sum_{k=1}^{n} (b_k\Theta_k - b_k(1-\Theta_k)) = \sum_{k=1}^{n} b_k(2\Theta_k - 1), \tag{3.1}$$

where, in general, b_k will be a real valued function of Θ.

Notice that a gambler may win any preassigned amount $w(\Theta)$ if Θ is known a priori. For example, any w could be achieved with the betting system

$$\begin{aligned} &b_1 = w(\Theta)\Theta_1 - w(\Theta)(1-\Theta_1), \\ &b_2 = b_3 = \ldots = b_n = 0. \end{aligned} \tag{3.2}$$

However, if he knows only $\Theta_1, \Theta_2, \ldots, \Theta_{k-1}$ when he must place his bet b_k, his set of achievable winnings w on $\{0,1\}^n$ is limited. For, if $\{b_1, b_2, \ldots, b_n\}$ achieves w, then manipulation of the above sum, noting the functional independence of b_k and Θ_k, yields

$$w(\Theta_1, \ldots, \Theta_{n-1}, 1) + w(\Theta_1, \ldots, \Theta_{n-1}, 0) = 2\sum_{k=1}^{n-1} b_k(2\Theta_k - 1), \tag{3.3}$$

and

$$w(\Theta_1, \ldots, \Theta_{n-1}, 1) - w(\Theta_1, \ldots, \Theta_{n-1}, 0) = 2b_n. \tag{3.4}$$

So, b_n is determined and 3.1 is replaced by 3.3 for the determination of b_{n-1}.
Proceeding, we find

$$\sum_{\Theta} w(\Theta) = 0 \tag{3.5}$$

Proof:

$$\begin{aligned} \sum_{\Theta} w(\Theta) &= \sum_{\Theta}\sum_{k=1}^{n} b_k(2\Theta_k - 1) \\ &= \sum_{k=1}^{n} b_k \sum_{\Theta}(2\Theta_k - 1) \\ &= \sum_{k=1}^{n} b_k \cdot 0 = 0. \end{aligned}$$

and

$$b_k = (\frac{1}{2})^{n-k+1} \sum_{(\Theta_k, \Theta_{k+1}, \ldots, \Theta_n) \in \{0,1\}^{n-k+1}} w(\Theta)(2\Theta_k - 1) \quad (k = 1, 2, \ldots, n). \tag{3.6}$$

Proof:
From Equation (3.3) and (3.4), the last bet b_n can be uniquely determined if all prior outcomes $\Theta_1, \ldots, \Theta_{n-1}$ are known.

To generalize for any b_k, consider the summation over all possible sequences of Θ from k to n, and the functional form $(2\Theta_k-)$ flips the impact of the bet b_k based on the outcome Θ_k. Thus:

$$b_k = \left(\frac{1}{2}\right)^{n-k+1} \sum_{(\Theta_k,\Theta_{k+1},\ldots,\Theta_n)\in\{0,1\}^{n-k+1}} w(\Theta)(2\Theta_k - 1)$$

Hence, for $w(\Theta)$ to be achievable by a sequential betting scheme, it is necessary and sufficient (3.5) be satisfied. The betting scheme achieving w is unique and is given by (3.6).

Summary:
Consider a betting strategy for a game against sequences: $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n) \in \{0,1\}^n$, which allows the bet b_k at stage k to be some element in a subset B_k of the collection B of all functions from $\{0,1\}^n$ to $\mathbb{R}$. Let $w : \{0,1\}^n \to \mathbb{R}$ be a desired set of net winnings defined for each sequence Θ in $\{0,1\}^n$. As before, 3.1 expresses the net winnings $w(\Theta)$ as a function of $\{b_1, b_2, \ldots, b_n\}$. Then:

(II) Trivially, if $B_k = B$, any w is achievable.

(III) If B_k is the set of all functions in B depending only on $\Theta_1, \Theta_2, \ldots, \Theta_{k-1}$, then w is achievable if and only if (3.5) is satisfied.

(IV) If, for $k = 1, 2, \ldots, n$, $B_k \subseteq B$ is the set of functions bounded in absolute value by b, depending only on $\Theta_1, \Theta_2, \ldots, \Theta_{k-1}$, then w is achievable if and only if

$$\sum_{\Theta} w(\Theta) = 0 \tag{3.7}$$

and if, for $k = 1, 2, \ldots, n$,

$$\left| \left(\frac{1}{2}\right)^{n-k+1} \sum_{(\Theta_k,\Theta_{k+1},\ldots,\Theta_n)\in\{0,1\}^{n-k+1}} w(\Theta)(2\Theta_k - 1) \right| < b \tag{3.8}$$

for every $(\Theta_1, \Theta_2, \ldots, \Theta_{k-1}) \in \{0,1\}^{k-1}$. This is the sequential betting scheme with bounded bet size.

3.2 Winnings which are functions of $\sum_{i=1}^{n} \Theta_i$

4 Random Predictors

5 Conclusions