Behaviour of Sequential Predictors of Binary Sequences

T. M. Cover

1 Introduction

This paper explores the effectiveness of sequential predictors on finite binary sequences. Prior work by Robbins, Hannan, and Blackwell has demonstrated algorithms capable of asymptotically achieving high scores by leveraging the empirical distribution of the sequences. Fogel has notably extended this by adapting the predictive model based on the sequence observed, achieving near-perfect accuracy on sequences displaying non-random patterns, such as prime numbers. This study investigates the capacity of sequential predictors to consistently generate high scores across a broad spectrum of sequences, delving into the inherent limits and potential of such algorithms in systematic and non-random environments.

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.

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

$$\bar{s} = \mathbb{E}_p(s) = \frac{1}{n} \sum_{i=1}^{n} \left(p_i \Theta_i + (1 - p_i)(1 - \Theta_i) \right) \qquad (2.2)$$

(I) 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

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

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

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

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

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

and

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

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

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

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

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

Proof:
For $k = n$, we need to show that

$$b_n = \left(\frac{1}{2}\right) \sum_{\Theta_n \in \{0,1\}} w(\Theta_1, \ldots, \Theta_{n-1}, \Theta_n)(2\Theta_n - 1)$$

we have from 3.4 that

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

For general b_k it holds that

$$\sum_{(\Theta_k, \ldots, \Theta_n)} w(\Theta_1, \ldots, \Theta_{k-1}, \Theta_k, \ldots, \Theta_n)(2\Theta_k - 1) =$$
$$\sum_{(\Theta_{k+1}, \ldots, \Theta_n)} w(\Theta_1, \ldots, \Theta_{k-1}, 1, \Theta_{k+1}, \ldots, \Theta_n) - w(\Theta_1, \ldots, \Theta_{k-1}, 0, \Theta_{k+1}, \ldots, \Theta_n) =$$
$$2 \cdot b_k \cdot 2^{n-(k+1)+1} = 2 \cdot b_k \cdot 2^{n-k} = b_k \cdot 2^{n-k+1}$$
$$\implies$$
$$b_k = \left(\frac{1}{2}\right)^{n-k+1} \sum_{(\Theta_k, \ldots, \Theta_n) \in \{0,1\}^{n-k+1}} w(\Theta)(2\Theta_k - 1).$$

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).
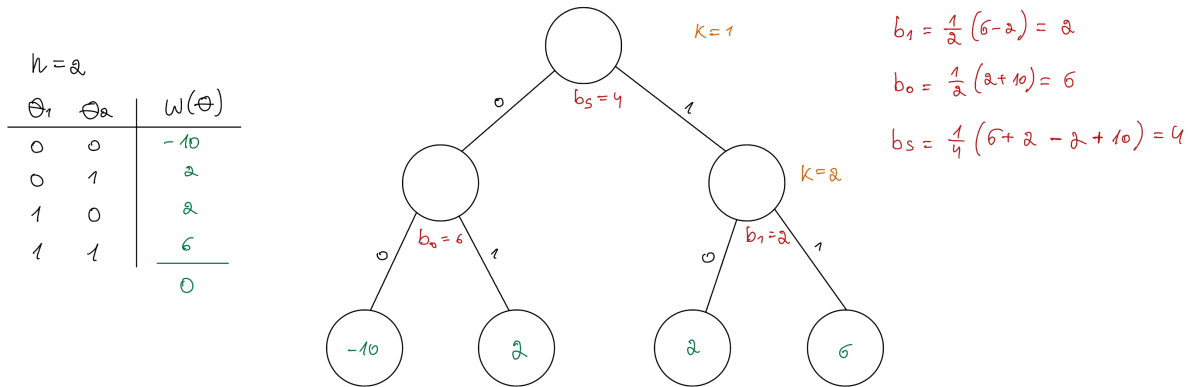

Figure 1: Sequential Betting Scheme

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

3

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 $\sum_{i=1}^n \Theta_i$

We may be interested in winnings w which are functions only of $\sum_{i=1}^n \Theta_i$, the number of 1's in Θ. In this case, define, for every $\Theta \in \{0,1\}^n$,

$$\hat{w}\left(\sum_{i=1}^n \Theta_i\right) = w(\Theta). \tag{3.9}$$

Then the conditions of 3.7 and 3.8 become respectively

$$\sum_{k=0}^n \binom{n}{k} \hat{w}(k) = 0 \tag{3.10}$$

and

$$|b_k(i)| = \left| \left(\frac{1}{2}\right)^{n-k+1} \sum_{j=0}^{n-k} (\hat{w}(i+j+1) - \hat{w}(i+j)) \binom{n-k}{j} \right| < b \tag{3.11}$$

For $i = 0, 1, \ldots, k-1$ and $k = 1, 2, \ldots, n$ where $b_k(i)$ is the bet at stage k when the sum of the first $k-1$ outcomes is i.

Proof:

We start by expanding the expression from Equation (3.8) under the assumption that $\hat{w}\left(\sum_{i=1}^n \Theta_i\right) = w(\Theta)$ from Equation (3.9):

LHS of $(3.8) = |b_k(i)|$

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

$$= \left| \left(\frac{1}{2} \right)^{n-k+1} \sum_{j=0}^{n-k+1} \left(\sum_{(\Theta_k,\ldots,\Theta_n) \in \{0,1\}^{n-k+1}:\sum \Theta_i = j} w(\Theta)(2\Theta_k - 1) \right) \right|$$

$$= \left| \left(\frac{1}{2} \right)^{n-k+1} \sum_{j=0}^{n-k} \left(\sum_{(\Theta_{k+1},\ldots,\Theta_n)} w(\Theta_1,\ldots,\Theta_k,1,\Theta_{k+1},\ldots,\Theta_n) - w(\Theta_1,\ldots,\Theta_k,0,\Theta_{k+1},\ldots,\Theta_n) \right) \right|$$

$$= \left| \left(\frac{1}{2} \right)^{n-k+1} \sum_{j=0}^{n-k} (\hat{w}(i+j+1) - \hat{w}(i+j)) \binom{n-k}{j} \right|,$$

where the binomial coefficient $\binom{n-k}{j}$ counts the number of ways in which $\sum \Theta_i = j$ can occur. The term inside the summation represents the change in winnings $\hat{w}$ when the count of 1's increases by one, hence reflecting the derivative (or difference) $\hat{w}(i+j+1) - \hat{w}(i+j)$. The multiplication by $\binom{n-k}{j}$ adjusts for the number of sequences that result in a particular sum j of 1's among the last $n - k$ flips.

Thus, we have shown that:

$$\text{LHS of } (3.8) = \text{RHS of } (3.11),$$

Now we will find a sufficient bound on $|b_k(i)|$ such that the conditions of 3.11 are satisfied. Letting $k = n$ in (3.11) we have the condition:

$$|b_n(i)| = \frac{1}{2} |\hat{w}(i+1) - \hat{w}(i)| < b \tag{3.12}$$

for $i = 0, 1, \ldots, n - 1$.

All other conditions of (3.11) are consequences of (3.12), since, assuming (3.12) true,

$$|b_k(i)| = \left| \left(\frac{1}{2} \right)^{n-k+1} \sum_{j=0}^{n-k} (w(i+j+1) - w(i+j)) \binom{n-k}{j} \right| \leq$$
$$\left(\frac{1}{2} \right)^{n-k+1} \sum_{j=0}^{n-k} \binom{n-k}{j} 2b = b. \tag{3.13}$$

(V) A terminal score $\hat{w}$ depending only on $\sum_{i=1}^{n} \Theta_i$ is achievable by a sequential betting scheme with bounded bet size b if and only if

$$\sum_{k=0}^{n} \binom{n}{k} \hat{w}(k) = 0 \tag{3.14}$$

and

$$|\hat{w}(k+1) - \hat{w}(k)| < 2b, \quad k = 1, 2, \ldots, n - 1. \tag{3.15}$$

3.3 Examples

3.3.1 Example 1 - foreknowledge of a sequence which will not occur

Consider a gambler betting on a binary sequence of length n, consisting of 1's and 0's. The gambler can choose his bet amount at each stage, based on the sequence observed so far. Suppose the gambler knows in advance that a specific sequence, say Θ^*, will not occur. The question is whether the gambler can guarantee a profit. The answer is yes; he can potentially win an infinite amount.

For any goal function $w(\Theta)$, there exists a betting strategy that guarantees a win of $w(\Theta)$ for any sequence $\Theta \neq \Theta^*$. This is achieved by setting:

$$w(\Theta^*) = - \sum_{\Theta \neq \Theta^*} w(\Theta)$$

and applying the betting strategy outlined in section 3.6. This example illustrates that, under certain conditions, a gambler can manipulate his bets based on a probability distribution over possible outcomes to achieve a desired terminal wealth distribution.

3.3.2 Example 2 - independent flips of a fair coin

Let $\Theta_1, \ldots, \Theta_n$ be independent flips of a fair coin. For a desired distribution function F, there exists a sequential betting scheme achieving a terminal distribution F_n such that

$$\sup_x |F(x) - F_n(x)| < \frac{1}{2^n}$$

To achieve F_n in the case of continuous F, choose w_i such that $F(w_i) = \frac{i}{2^n}$ for $i = 1, 2, \ldots, 2^n - 1$ and $w_0 = - \sum_{i=1}^{2^n-1} w_i$. Associate the winnings w with the outcomes Θ's in an arbitrary fashion and use the betting scheme of 3.6.

3.4 Summary

Our investigations reveal that while almost any probability distribution for terminal capital can theoretically be achieved through sequential betting strategies, in practice, most terminal distributions are not appealing to gamblers. This disinterest largely stems from the nature of popular gambling systems like Martingale and Progression Systems, which typically offer small gains offset by a significant risk of substantial losses.

Most betting systems fail to optimize gambler's utilities because they do not sufficiently reward the risk of extreme outcomes, leading to inherently suboptimal strategies. The utility functions, if properly utilized to influence betting decisions, should account for the potential gains at the extremes of the distribution. However, the increase at a few terminal points, as suggested by theoretical models, often does not compensate for the overall risk, making these strategies less favorable.

4 Random Predictors

4.1 Average Scores of Random Predictors as Winnings in Sequential Betting

In 3.1 we have seen that the net winnings $w(\Theta)$ of a series of n bets $b = (b_1, b_2, \ldots, b_n)$ on the outcomes of a sequence $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n) \in \{0, 1\}^n$ can be expressed as

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

We now consider the case where the bets b_k are random variables, and the gambler's strategy is to choose the b_k's to maximize his expected winnings. The random predictor $p = (p_1, p_2, \ldots, p_n)$ yields an average score

$$\bar{s}(\Theta) = \frac{1}{n} \sum_{k=1}^{n} (p_k \Theta_k + (1 - p_k)(1 - \Theta_k)) \tag{4.1}$$

The characterization of the set of all achievable scores for all $p \in P$ is a consequence of the previous section under the correspondence of bets with probabilities given by

$$b_k = \frac{1}{n}(p_k - \frac{1}{2}) \quad \left(\Rightarrow \quad p_k = \frac{1}{2} + nb_k \right) \tag{4.2}$$

The probability constraints $0 \le p_k \le 1$ for all $k \in [n]$ imply bounded sets

$$|b_k| \le \frac{1}{2n} \tag{4.3}$$

Thus, if $b = (b_1, b_2, \ldots, b_n)$ yields $w(\Theta)$, the corresponding predictor $p = (p_1, p_2, \ldots, p_n)$ yields score

$$\begin{aligned} \bar{s}(\Theta) &= \frac{1}{n} \sum_{k=1}^{n} (p_k \Theta_k + (1 - p_k)(1 - \Theta_k)) \\ &= \frac{1}{n} \sum_{k=1}^{n} \left((\frac{1}{2} + nb_k)\Theta_k + (\frac{1}{2} - nb_k)(1 - \Theta_k) \right) \\ &= \frac{1}{2} + \frac{1}{n} \sum_{k=1}^{n} nb_k(2\Theta_k - 1) = \frac{1}{2} + w(\Theta) \end{aligned} \tag{4.4}$$

(VI) We deduce from previous sections (3.5) that an average score function $\bar{s}(\Theta)$ for $\Theta \in \{0,1\}^n$ is achievable by a sequential predictor $p \in P$ if and only if:

$$\sum_{\Theta} w(\Theta) = 0 \quad \Longleftrightarrow \quad \sum_{\Theta \in \{0,1\}^n} \bar{s}(\Theta) = \frac{1}{2} 2^n \quad \Longleftrightarrow \quad \left(\frac{1}{2}\right)^n \sum_{\Theta} \bar{s}(\Theta) = \frac{1}{2} \tag{4.5}$$

and

$$\left| \left(\frac{1}{2}\right)^{n-k+1} \sum_{(\Theta_k, \ldots, \Theta_n) \in \{0,1\}^{n-k+1}} \bar{s}(\Theta)(2\Theta_k - 1) \right| \le b \le \frac{1}{2n} \tag{4.6}$$

for all $\Theta \in \{0,1\}^n$ and each $k = 1, 2, \ldots, n$.

($w(\Theta) = \bar{s}(\Theta) - \frac{1}{2}$ and $\sum_{\Theta_k, \ldots, \Theta_n}(-\frac{1}{2})(2\Theta_k - 1) = 0.$)

(VII) A score

$$\bar{s}(\Theta) = \hat{s}\left(\frac{1}{n}\sum_{k=1}^{n}\Theta_k\right) \tag{4.7}$$

depending only on the weight of the sequence Θ, is achievable by a sequential predictor $\hat{p} \in P$ if and only if:

$$\left(\frac{1}{2}\right)^n \sum_{k=0}^{n}\binom{n}{k}\hat{s}\left(\frac{k}{n}\right) = \frac{1}{2} \tag{4.8}$$

and

$$\left|\hat{s}\left(\frac{k+1}{n}\right) - \hat{s}\left(\frac{k}{n}\right)\right| \le 2b \le \frac{1}{n} \tag{4.9}$$

for $k = 0, 1, 2, \ldots, n-1$.

In accordance with the previous notation, let the predicator at stage k given history i be denoted by

$$\hat{p}_k(i) = Pr\{\delta_k = 1 | \sum_{j=1}^{k-1}\Theta_j = i\} \tag{4.10}$$

for $i = 1, 2, \ldots, k-1$ and $k = 1, 2, \ldots, n$.
Then, utilizing (3.6), (4.2) and (4.6) (as we did in (3.11))

$$\hat{p}_k(i) = \frac{1}{2} + nb_k(i)$$
$$= \frac{1}{2} + n\left(\frac{1}{2}\right)^{n-k+1}\sum_{j=0}^{n-k}\left(\hat{s}\left(\frac{i+j+1}{n}\right) - \hat{s}\left(\frac{i+j}{n}\right)\right)\binom{n-k}{j} \tag{4.11}$$

4.2 Examples and uniform distribution

4.2.1 Example 1:

$\bar{s}(\Theta) = \frac{1}{2}$ for all $\Theta \in \{0,1\}^n$ is achievable by the predictor $p = (\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$.
proof:
Given the score function

$$\bar{s}(\Theta) = \frac{1}{2},$$

for all $\Theta \in \{0,1\}^n$, we want to demonstrate that it satisfies the conditions of achievability defined by Equations 4.5 and 4.6.
Condition 1: Firstly, according to Equation 4.5, the total sum of scores over all possible Θ configurations must be half of 2^n:

$$\sum_{\Theta}\bar{s}(\Theta) = \frac{1}{2} \times 2^n = 2^{n-1}.$$

This simplifies to:

$$\left(\frac{1}{2}\right)^n \sum_{\Theta}\bar{s}(\Theta) = \left(\frac{1}{2}\right)^n \times \frac{1}{2} \times 2^n = \frac{1}{2}.$$

Hence, condition 4.5 is satisfied as the average of the scores equals $\frac{1}{2}$.

Condition 2: Next, to verify Equation 4.6, we compute:

$$\left| \left(\frac{1}{2}\right)^{n-k+1} \sum_{(\Theta_k,\ldots,\Theta_n)\in\{0,1\}^{n-k+1}} \bar{s}(\Theta)(2\Theta_k - 1) \right|.$$

Since $\bar{s}(\Theta) = \frac{1}{2}$ for all Θ, and noting the symmetry between $\Theta_k = 1$ and $\Theta_k = 0$, the weighted sum for each k simplifies to zero:

$$\sum_{(\Theta_k,\ldots,\Theta_n)} (2\Theta_k - 1) = 0.$$

Thus, each term in the sum is zero and the absolute condition in 4.6 is trivially satisfied as $0 \le b \le \frac{1}{2n}$.

Then we can see that the predictor $p = (\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$ achieves the average score $\bar{s}(\Theta) = \frac{1}{2}$ for all $\Theta \in \{0,1\}^n$:

$$\bar{s}(\Theta) = \frac{1}{n} \sum_{k=1}^{n} (p_k\Theta_k + (1 - p_k)(1 - \Theta_k))$$

$$= \frac{1}{n} \sum_{k=1}^{n} \left(\frac{1}{2}\Theta_k + \frac{1}{2}(1 - \Theta_k)\right)$$

$$= \frac{1}{n} \sum_{k=1}^{n} \frac{1}{2} = \frac{1}{2}$$

Intuitively, the predictor $p = (\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$ is the most conservative strategy, and for every flip we will predict correctly with probability $\frac{1}{2}$.

4.2.2 Example 2:

$$\hat{s}\left(\frac{k}{n}\right) = \frac{k}{n}$$

proof:

Given the score function

$$\hat{s}\left(\frac{k}{n}\right) = \frac{k}{n},$$

we need to verify the conditions stated in the given constraints for achievability by a sequential predictor.

Condition 1:

$$\left(\frac{1}{2}\right)^n \sum_{k=0}^{n} \binom{n}{k} \hat{s}\left(\frac{k}{n}\right) = \frac{1}{2}$$

This condition simplifies as follows:

$$\left(\frac{1}{2}\right)^n \sum_{k=0}^{n} \binom{n}{k} \frac{k}{n} = \left(\frac{1}{2}\right)^n \frac{1}{n} \sum_{k=0}^{n} k \binom{n}{k} = \left(\frac{1}{2}\right)^n \frac{n2^{n-1}}{n} = \frac{1}{2}$$

which holds true by the binomial theorem where the expected value of k in a binomial distribution $B(n, \frac{1}{2})$ is $\frac{n}{2}$.

Condition 2:

$$\left| \hat{s}\left(\frac{k+1}{n}\right) - \hat{s}\left(\frac{k}{n}\right) \right| = \left| \frac{k+1}{n} - \frac{k}{n} \right| = \frac{1}{n} \le 2b \le \frac{1}{n}$$

This also holds as the absolute difference between successive terms of $\hat{s}$ is $\frac{1}{n}$, which is within the bounds specified by the constraint.

Construction of Sequential Predictor $\hat{p}$**:** Given $\hat{s}$, the predictor $\hat{p}_k$ for each stage k based on the history of sums can be designed according to Equation 4.11:

$$\hat{p}_k(i) = \frac{1}{2} + n \left(\frac{1}{2}\right)^{n-k+1} \sum_{j=0}^{n-k} \left(\hat{s}\left(\frac{i+j+1}{n}\right) - \hat{s}\left(\frac{i+j}{n}\right)\right) \binom{n-k}{j}$$

Let's define $\hat{s}\left(\frac{k}{n}\right) = \frac{k}{n}$. Then for each k, we can calculate the change in score needed for the next predictor:

$$\Delta s_k = \hat{s}\left(\frac{k+1}{n}\right) - \hat{s}\left(\frac{k}{n}\right) = \frac{k+1}{n} - \frac{k}{n} = \frac{1}{n}.$$

This Δs_k affects the predictor $\hat{p}_k(i)$, which depends on the history i as:

$$\hat{p}_k(i) = \frac{1}{2} + n \left(\frac{1}{2}\right)^{n-k+1} \sum_{j=0}^{n-k} \left(\hat{s}\left(\frac{i+j+1}{n}\right) - \hat{s}\left(\frac{i+j}{n}\right)\right) \binom{n-k}{j}.$$

Given $\Delta s_k = \frac{1}{n}$, the formula simplifies because $\hat{s}\left(\frac{i+j+1}{n}\right) - \hat{s}\left(\frac{i+j}{n}\right)$ equals $\frac{1}{n}$ for all j. Thus,

$$\hat{p}_k(i) = \frac{1}{2} + n \left(\frac{1}{2}\right)^{n-k+1} \sum_{j=0}^{n-k} \left(\frac{1}{n}\right) \binom{n-k}{j}.$$

4.3 Uniformly Asymptotically Approachable Scores

4.3.1 Definition

A score $\hat{s}(\frac{k}{n})$ which satisfies (4.9) but not (4.8) can be achieved with error

$$\varepsilon_n = \frac{1}{2} - \left(\frac{1}{2}\right)^n \sum_{k=0}^{n} \binom{n}{k} \hat{s}\left(\frac{k}{n}\right) \tag{4.12}$$

uniformly in Θ by the p which achieves $\hat{s} + \varepsilon_n$.
If $\varepsilon_n \to 0$, we shall say that $\hat{s}$ is ***uniformly asymptotically approachable*** by $\hat{p}$.

4.3.2 Bayes Envelope

Of particular interest is the Bayes envelope

$$\max\left\{\frac{1}{n}\sum_{i=1}^{n}\Theta_i, 1 - \frac{1}{n}\sum_{i=1}^{n}\Theta_i\right\},$$

considered by Hannan and Robbins, corresponding to the score function

$$\hat{s}(\eta) = \max\{\eta, 1 - \eta\}, \quad 0 < \eta < 1. \tag{4.13}$$

$\hat{s}$ satisfies (4.9) and hence can be achieved uniformly in Θ with error

$$\epsilon_n = \frac{1}{2} - \left(\frac{1}{2}\right)^n \sum_{k=0}^{n} \binom{n}{k} \max\left\{\frac{k}{n}, 1 - \frac{k}{n}\right\} \approx \frac{-1}{\sqrt{2\pi n}}. \tag{4.14}$$

Proof of the Approximation for ϵ_n

Given:

$$\epsilon_n = \frac{1}{2} - \left(\frac{1}{2}\right)^n \sum_{k=0}^n \binom{n}{k} \max\left\{\frac{k}{n}, 1 - \frac{k}{n}\right\}$$

we seek to prove that:

$$\epsilon_n \approx \frac{-1}{\sqrt{2\pi n}}.$$

An elementary form of the CLT states the following.

Let $\Theta_1, \Theta_2, \ldots, \Theta_n$ denote a random sample of n independent observations from a population with overall expected value (average) μ and finite variance σ^2, and let $\overline{\Theta}_n$ denote the sample mean of that sample (which is itself a random variable). Then the limit as $n \to \infty$ of the distribution of

$$\frac{\overline{X}_n - \mu}{\sigma_{\overline{X}_n}},$$

where $\sigma_{\overline{X}_n} = \frac{\sigma}{\sqrt{n}}$, is the standard normal distribution, i.e., $\overline{\Theta}_n \sim N(\mu, \sigma^2/n)$.

The PDF of Folded Normal Distribution is given by:

$$f_Y(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x+\mu)^2}{2\sigma^2}}$$

and the mean is given by:

$$\mu_Y = \sigma\sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} + \mu\left(1 - 2\Phi\left(-\frac{\mu}{\sigma}\right)\right)$$

where Φ is the normal cumulative distribution function.

We get:

$$\epsilon_n = \frac{1}{2} - \left(\frac{1}{2}\right)^n \sum_{k=0}^n \binom{n}{k} \max\left\{\frac{k}{n}, 1 - \frac{k}{n}\right\}$$

$$= \frac{1}{2} - \left(\frac{1}{2}\right)^n \sum_{k=0}^n \binom{n}{k} \left(\frac{1}{2} + \left|\frac{k}{n} - \frac{1}{2}\right|\right)$$

$$= \frac{1}{2} - \left(\frac{1}{2}\right)^n \sum_{k=0}^n \left(\frac{1}{2}\right)\binom{n}{k} - \left(\frac{1}{2}\right)^n \sum_{k=0}^n \binom{n}{k}\left|\frac{k}{n} - \frac{1}{2}\right|$$

$$= -\left(\frac{1}{2}\right)^n \sum_{k=0}^n \binom{n}{k}\left|\frac{k}{n} - \frac{1}{2}\right|$$

$$= -\sum_{k=0}^n \frac{\binom{n}{k}}{2^n}\left|\frac{k}{n} - \frac{1}{2}\right|$$

$$\sim -\left(\text{The expectation of the folded normal distribution with } \mu = 0, \sigma^2 = \frac{1}{4n}\right)$$

$$= -\sqrt{\frac{1}{4n}}\sqrt{\frac{2}{\pi}} e^{-\frac{0}{2\cdot\frac{1}{4n}}} + 0\left(1 - 2\Phi\left(-\frac{0}{\sqrt{\frac{1}{4n}}}\right)\right)$$
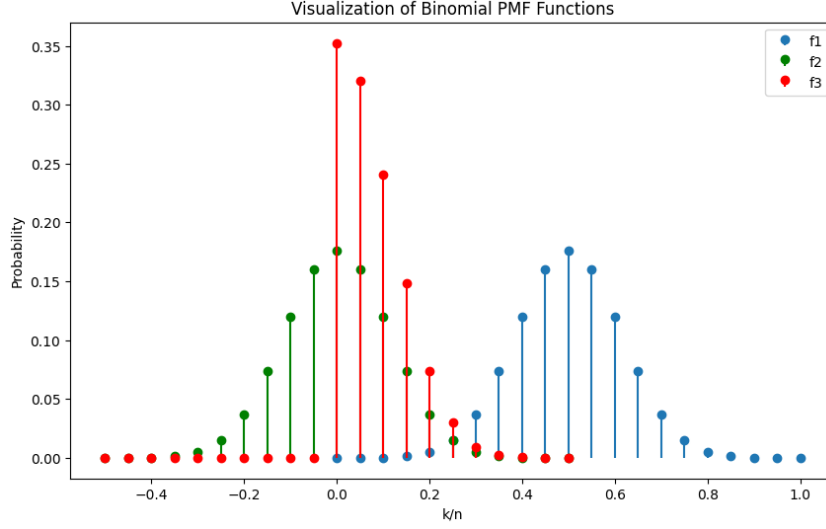
$$= -\frac{1}{\sqrt{2\pi n}}$$
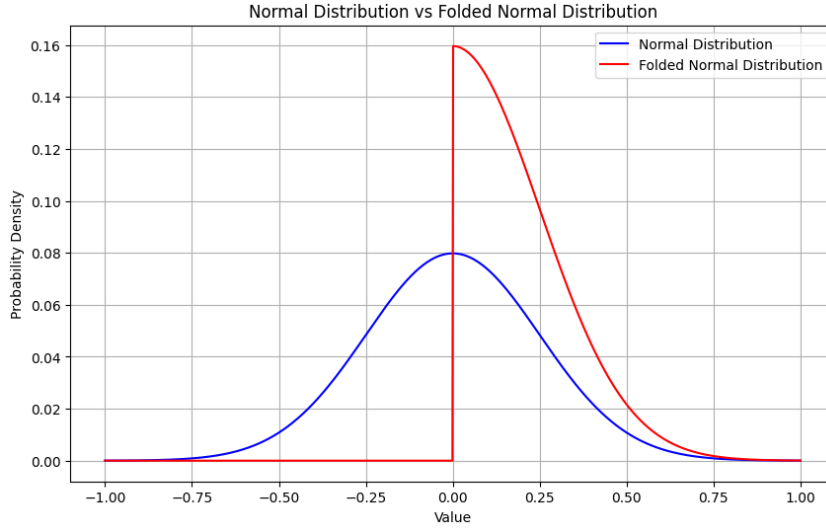
Figure 2: Visualization of Binomial PMF Functions

Figure 3: Normal Distribution vs Folded Normal Distribution

4.3.3 The Set of Asymptotically Approachable Scores

It is natural to ask what set of score functions $\hat{s}(\eta), 0 \leq \eta \leq 1$, satisfy (4.8) and (4.9) in the limit as $n \to \infty$. Clearly, the set of all continous differentiable such functions satisfies the constraints:

$$\hat{s}(\frac{1}{2}) = \frac{1}{2} \tag{4.15}$$

$$|\hat{s}(\eta)| < 1, \quad 0 \leq \eta \leq 1 \tag{4.16}$$

We define the set of uniformly asymptotically approachable scores in the sense of Blackwell as follows:

$$\hat{\mathcal{S}} = \{\hat{s}(\eta), 0 \leq \eta \leq 1 : \exists p = (p_1, p_2, \ldots) \in \{0, 1\}^\infty \text{ s.t. } \forall \Theta \in \{0, 1\}^\infty$$
$$\mathbb{P}\left\{ \lim_{n \to \infty} \left| \frac{1}{n} \sum_{i=1}^n s_i(\Theta_i) - \hat{s}\left(\frac{1}{n} \sum_{i=1}^n \Theta_i \right) \right| = 0 \right\} = 1 \} \tag{4.17}$$

(Recall that $s_i(\Theta_i) = \delta_i \Theta_i + (1 - \delta_i)(1 - \Theta_i)$ is a Random Variable govern by p_i.)

Blackwell reports that necessary and sufficient conditions for approachability and exclaudability are known only for convex functions.

In particular, $\eta \alpha + (1 - \eta)(1 - \alpha)$, for $0 \le \alpha \le 1$, and $\max \eta, 1 - \eta$ are approachable.

Cover offer the conjecture that $\hat{s}$ is approachable if and only if $\hat{s}\left(\frac{1}{2}\right) = \frac{1}{2}$ and

$$\left| \hat{s}\left(\frac{k+1}{n}\right) - \hat{s}\left(\frac{k}{n}\right) \right| < \frac{1}{n},$$

for all $k = 0, 1, 2, \ldots, n - 1$ and $n = 1, 2, \ldots$.

5 Conclusions