Beyond chance: Predicting the unpredictable

Leonard M. Wapner[1]

[1] *Professor Emeritus of Mathematics, Division of Mathematical Sciences, El Camino College, Torrance, California 90506 (USA)*

Abstract

The outcome of a randomly tossed fair coin is generally considered unpredictable in the sense that the probability of a successful prediction is always 1/2. But there are scenarios, highly contrived, where successful predictions are expected more often than not. Two such scenarios are presented, raising questions as to their real-world relevance.

"Chance favors the prepared mind."
—Louis Pasteur

1 Introduction

A fair coin is to be tossed and we wish to predict the outcome (heads or tails) of the toss. Knowing nothing additional about the initial, physical conditions of the toss, the probability of a successful prediction is 1/2. We could, of course, do much better if the coin were biased, one way or the other, and we knew the nature of the bias. But there is a sense of being locked in to a prediction success probability of 1/2 for any binomial experiment where the two possible outcomes are equiprobable.

Under the right conditions and with a bit of mathematical creativity, the prediction success probability can be increased. The first scenario is a repurposing of the well-known Martingale betting strategy. The second is a variation of a paradox appearing in [Gal98], drawn from an example given by David Blackwell [Bla51].

2 The Martingale betting strategy

The Martingale betting strategy, dating back to 18^{th} century France, is a well-known but flawed betting system mostly applied to even money wagering. To illustrate, imagine

DOI 10.2478/rmm-2025-0008

placing a \$1 bet on heads (H) as the outcome of a fair coin toss. If the coin lands heads you win \$1 and may walk away or place another \$1 bet on heads for the next toss outcome. But whenever the coin lands tails (T) you must bet again, doubling the amount previously wagered. If the coin continues to land tails on subsequent tosses, you must continue betting, doubling your bet each time until the coin ultimately lands heads, at which time you may take your winnings and leave, or continue play by starting the process over again with a \$1 bet on heads. Say, for example, the first four toss outcomes are T, T, T and H. Then you lose \$1, lose \$2, lose \$4 and finally win \$8, giving you a net gain of \$1. Each run of outcomes of the form TTT...H provides you a net gain of \$1. It is almost certain that the coin will from time to time land heads and your bankroll will almost certainly increase, albeit slowly and nonmonotonically.

The system's flaw is well known. Excessively long losing streaks of tails are inevitable, requiring bets which will either bankrupt the player or exceed the table limit (maximum allowable bet) imposed by the house. And such long losing streaks are more likely than one might expect.

To investigate further, consider the length (l) of a run of outcomes of the form TTT...TH which ultimately increases your bankroll by \$1. See Table 1.

Run	Length l	Probability
H	1	$1/2$
TH	2	$(1/2)^2 = 1/4$
TTH	3	$(1/2)^3 = 1/8$
TTT...TH	n	$(1/2)^n = 2^{-n}$

Table 1: Winning Martingale runs

The expected value of l is

$$E(l) = 1\left(\frac{1}{2}\right) + 2\left(\frac{1}{2}\right)^2 + 3\left(\frac{1}{2}\right)^3 + \cdots = \sum_{n=1}^{\infty} n\left(\frac{1}{2}\right)^n = 2$$

(To see why the series converges to 2, use the infinite geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{for} \quad |x| < 1$$

Differentiating with respect to x,

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2} \Rightarrow \sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2}$$

Evaluating at $x = 1/2$ gives the desired result.)

Despite $E(l)$ being relatively small, a surprisingly simple result in [Sch12] suggests long, potentially bankrupting streaks of tails are more common than one might expect. When a fair coin is tossed n consecutive times, we expect roughly $n/2$ heads outcomes, each one of which marks an opportunity for a run of tails to begin. (This ignores the singular possibility of a run of tails starting at the first flip of the coin.) Roughly $1/2$ of these $n/2$ opportunities would be followed by at least one tails outcome, $(1/2)^2 = 1/4$ of these opportunities would be followed by a run of two or more tails, and so on. This suggests we should expect $(n/2)(1/2)^l$ runs of tails of length at least l. Now let L be the largest value of l expected to occur, of which we would expect one run. That is,

$$\left(\frac{n}{2}\right)\left(\frac{1}{2}\right)^L = 1$$

yielding

$$L = \log_2 \frac{n}{2}$$

So, for a fair coin tossed $n = 1024$ times, the most likely longest streak of tails would be

$$L = \log_2 \frac{1024}{2} = 9$$

which in some cases would either bankrupt the player or force a bet which would exceed the maximum bet allowed by the house.

3 The Martingale scheme repurposed

Disappointing as this may be for gamblers seeking a get rich quick scheme, the Martingale system has merit if applied differently. Assume a professional forecaster (business, weather, political events, etc.) wishes to promote themselves by documenting and publishing their high success rate on predicting events having an estimated probability of $1/2$. To establish credibility, the forecaster will predict outcomes and publish the success rate for sequences of coin tosses. The forecaster will predict heads each time the coin is tossed. If the outcome is heads on the first toss of the coin, the forecaster is correct and may go on to predict heads on the next toss. But if the coin lands tails, the forecaster will now make two separate predictions that the next outcome will be heads. If the coin continues to land tails, the forecaster will continue doubling the number of heads predictions on subsequent tosses until the coin ultimately lands heads, at which time the forecaster will have made one more correct prediction than incorrect predictions. To be precise, for a string of outcomes of the form TTT…TH, the forecaster will have made a total of

$$1 + 2 + 4 + \cdots + 2^{n-1} = 2^n - 1$$

predictions, of which 2^{n-1} are correct, giving the forecaster a success rate of

$$\frac{2^{n-1}}{2^n - 1} = \frac{1}{2 - \frac{1}{2^{n-1}}} > \frac{1}{2}$$

Possible outcomes are shown in Table 2.

Run	Successful predictions	Probability
H	1 of 1	1/2
TH	2 of 3	$(1/2)^2 = 1/4$
TTH	4 of 7	$(1/2)^3 = 1/8$
TTT...TH	2^{n-1} of $2^n - 1$	$(1/2)^n$

Table 2: The Martingale prediction scheme

$$\begin{aligned} E(\text{success rate}) &= 1\left(\frac{1}{2}\right) + \left(\frac{2}{3}\right)\left(\frac{1}{2}\right)^2 + \left(\frac{4}{7}\right)\left(\frac{1}{2}\right)^3 + \cdots \\ &= \sum_{n=1}^{\infty} \frac{2^{n-1}}{2^n - 1} \cdot \frac{1}{2^n} \\ &= \sum_{n=1}^{\infty} \frac{1}{2^{n+1} - 2} \\ &\approx 0.80 \end{aligned}$$

The Martingale scheme applied in this manner almost certainly guarantees an overall prediction success rate greater than 1/2. However, the probability of any single prediction being correct remains 1/2. In the following sections, a more confounding scheme is described where individual outcomes of experiments similar to that of tossing a fair coin are correctly predictable more often than not.

4 The Willoughby postdiction scheme

This scheme gets its name from “A Stop at Willoughby”, an episode of the American television series *The Twilight Zone*. Imagine a rail line extending infinitely far to the east and west. Train stations are spaced one mile apart, with signal lights positioned midway between each pair of adjacent stations. A train moves in random walk fashion from one station to an adjacent station departing each hour on the hour in the following manner. At 20 minutes into each hour, the train engineer flips a coin. If the coin lands heads the train will depart, at the top of the hour, to the station one mile east. If the coin lands tails the train will depart, at the top of the hour, to the station one mile west. Throughout this process and at exactly 40 minutes into each hour, exactly one of the infinitely many signal lights will flash for testing purposes only. The flash is visible infinitely far to the east and west. Passengers have no idea which light will be the next to flash nor would they necessarily care. The lights flash independently of the tossed coin and do not effect which way the train travels. See Figures 1 and 2.

Now imagine you awake at 2:30 p.m. from a drug induced sleep and find yourself on this train, stopped at some unknown station. You have no idea where, on this infinitely extending rail line, you are currently stopped. Just then the conductor passes through

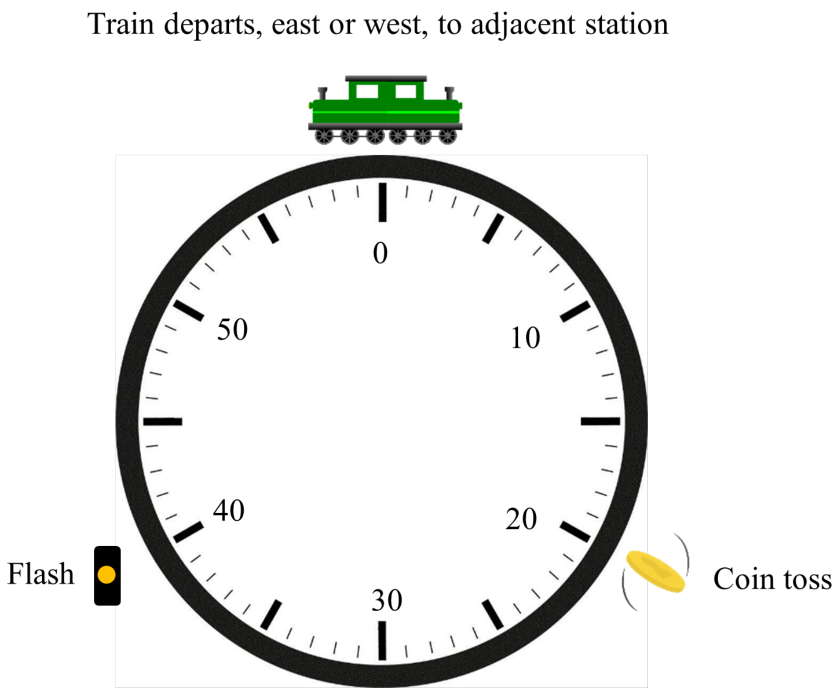

Figure 1: Willoughby events - hourly

and announces that the next stop, as determined by the 2:20 p.m. coin flip, is Willoughby. You check a route schematic posted by your seat and conclude you are presently at East Willoughby or West Willoughby. Given only these two possibilities and that there is no reason for one to be more likely than the other, you conclude the two possibilities are equiprobable. See Figure 2. Readers may recognize such reasoning as a special case of the *principle of insufficient reason*, later named the *principle of indifference*, which states that if there are n possible outcomes to an experiment and there is no reason to believe either outcome is more likely than another, then the probability $1/n$ should be assigned to each. The principle has been viewed by some as controversial because in some unusual cases it contradicts the laws of probability. But in cases such as that presented here, its application seems reasonable. For an in depth discussion of this principle, its potential contradictions and its application to prediction see [Wap19].

Prior to the train's scheduled 3:00 p.m. departure, is it possible for you to guess, technically *postdict*, the outcome of the engineer's 2:20 p.m. coin flip with a success probability greater than 1/2? Suprisingly, it can be done. You watch for the next signal light to flash, scheduled at 2:40 p.m. There is a non-zero probability p that the light to flash will be one of the two lights adjacent to Willoughby, in which case it will indicate to you, with certainty, the direction (east or west) of Willoughby. See Figure 2. The probability that the 2:40 p.m. flash comes from neither of the two lights adjacent to Willoughby is $1 - p$, in which case, due to the equiprobability of your being at East Willoughby or West Willoughby, the probability is 1/2 that the flash will point in the direction of Willoughby.

DOI 10.2478/rmm-2025-0008

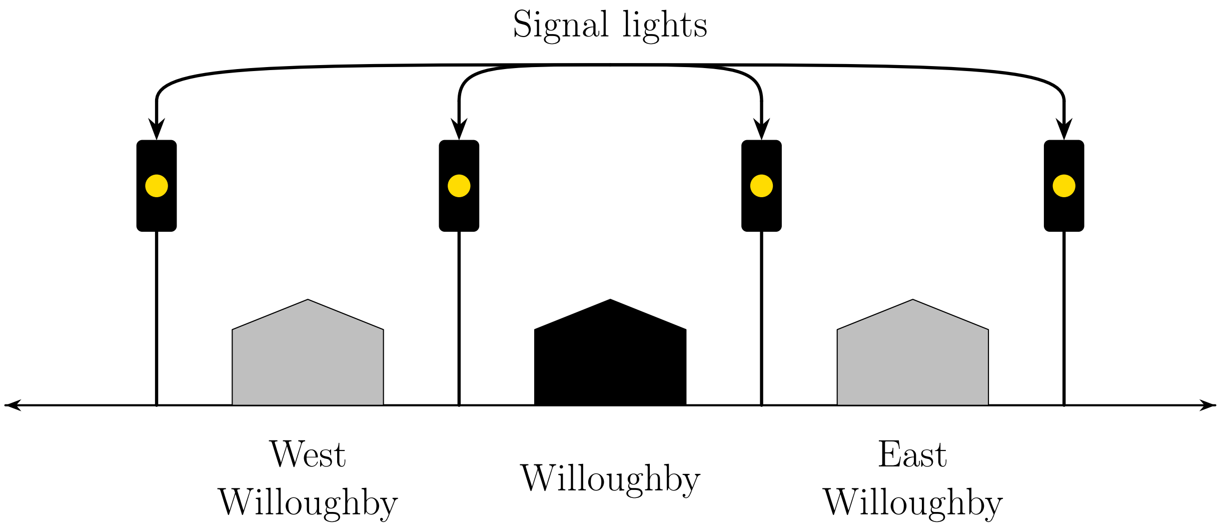

Figure 2: The Willoughby scheme

Overall,

$$P(signal\ points\ in\ direction\ of\ Willoughby) = p + \frac{1}{2}(1-p) = \frac{1}{2} + \frac{p}{2} > \frac{1}{2}$$

So, if the light flashes to your east, you would guess that the train will be departing to the east and that the engineer's coin landed heads. If the light flashes to the west, you would guess that the train will depart to the west and that the engineer's coin landed tails. You should expect your guess (east/heads or tails/west) to be successful more often than not. Your guess, as it relates to the coin toss outcome, is here described as a *postdiction*, as opposed to a prediction, because it involves a past event (the coin toss has already occurred) as opposed to a future event. The next section converts the postdiction scheme to a prediction, while maintaining the expectation of a correct guess being more likely than not.

5 The Willoughby prediction scheme

Converting the postdiction scheme to a prediction scheme requires only a minimal adjustment.

Assume the engineer dozes off just prior to 2:20 p.m. and fails to flip the coin as scheduled. The engineer awakens at 2:50 p.m. and quickly flips the coin, 10 minutes after you saw the signal flash and made your guess, now a *prediction*, as to the toss outcome and direction of departure. Your success rate is still greater than 1/2 because the toss outcomes are independent of the signal flashes. It is irrelevant which occurs first.

DOI 10.2478/rmm-2025-0008

6 Conclusion

Of the two scenarios (Martingale repurposed and Willoughby prediction), the first is far more likely to be realized in actuality. The publication of successful prediction rates associated with the Martingale scheme involves no falsification, although the manner in which success rates are so achieved may be deceiving, falsely suggesting that any given prediction is most likely to be correct.

The Willoughby prediction scheme, though mathematically valid, is far too contrived for it to be achieved in actuality. But there being no mathematical contradictions, the door remains open to variations and applications.

Acknowledgement

I wish to thank Professor James Stein for his input relating to this topic.

References

[Bla51] D. Blackwell. "On the Translation Parameter Problem for Discrete Variables". In: *Annals of Mathematical Statistics* 22(3), (1951) pp. 393–399.

[Gal98] D. Gale. *Tracking the Automatic Ant and Other Mathematical Explorations*. New York: Springer-Verlag, 1998.

[Sch12] M. Schilling. "The Surprising Predictability of Long Runs". In: *Mathematics Magazine*, 85(2) (2012) pp. 141–149.

[Wap19] L. Wapner. "The Prescient Power of Indifference". In: *Mathematical Gazette* 103(556), (2019) pp. 41–46.