ALICE AND BOB ON $\mathbb{X}$: REVERSAL, COUPLING, RENEWAL

GEOFFREY R. GRIMMETT

Abstract. A neat question involving coin flips surfaced on $\mathbb{X}$, and generated an intensive outbreak of 'social mathematics'. In a sequence of flips of a fair coin, Alice wins a point at each appearance of two consecutive heads, and Bob wins a point whenever a head is followed immediately by a tail. Who is more likely to win the game? The subsequent discussion illustrated conflicting intuitions, and concluded with the correct answer (it is a close thing). It is explained here why the context of the question is interesting and how it may be answered in a quantitative manner using the probabilistic techniques of reversal, coupling, and renewal.

1. The problem

Question 1.1. *A fair coin is tossed n times. Alice scores one point at each appearance of two consecutive heads, and Bob scores one point each time a head is followed immediately by a tail. The winner is the player who accumulates more points. Who is the more likely to win?*

Here are some intuitive arguments in order of decreasing naiveté.

(a) In any pair of consecutive coin tosses, Alice wins a point with probability $\frac{1}{4}$, and Bob wins a point with the same probability. Therefore, each has a mean total score of $\frac{1}{4}(n-1)$. Since these means are equal, each player has the same probability of winning.
(b) Alice's points tend to arrive in clusters, whereas Bob's are isolated. That favours Alice, so she is more likely to win.
(c) When Alice wins a point, she has an increased chance of winning again. Therefore, she tends to win the game by a wider margin than Bob. However, her mean total is the same as Bob's. It follows that Bob has a greater chance of winning.

This question (with $n = 100$) was posed by Daniel Litt on his $\mathbb{X}$ feed [8] on 16 March 2024. At the current time of writing, his post has attracted 1.2M views. First responders tended to favour Alice above Bob on the grounds of argument (b), and a vote was reported as placing Alice (26.3%) over Bob (10.2%), with 42.8% of the 51,588 voters supporting equality. Later simulations appeared to show that Bob has a slight advantage.

ChatGPT has changed its position on the question over the intervening months. Initially it favoured Alice on the grounds given in (b) above. At the time of writing it has veered

Date: August 30, 2024.
2010 *Mathematics Subject Classification.* 60C05.
Key words and phrases. Coin tosses, random game, local central limit theorem.

towards (a), with the conclusion that "The game is fair to both players in terms of expected outcomes." Fair enough, but not very helpful. There remains hope for mathematicians.

More complete answers began to surface on Litt's $\mathbb{X}$ feed fairly soon after the original post, and suggestions were made for rigorous proof (see Remark 1.4 below). It is not easy to convey the details and to check the correctness of a mathematical argument within the format confines of $\mathbb{X}$, and hence this note. We present three conclusions.

Theorem 1.2. *Consider n flips of a fair coin.*

(a) *Bob is (strictly) more likely than Alice to win when $n \ge 3$.*
(b) $\mathbb{P}(\text{Bob wins}) - \mathbb{P}(\text{Alice wins}) \sim c/\sqrt{n}$ *as $n \to \infty$, where*

$$c = \frac{1}{2\sqrt{\pi}} \approx 0.282.$$

(c) *The probability of a tie is asymptotically $2c/\sqrt{n}$, with c as in part (b).*

Explicit representations for the probabilities in (b) and (c) are given in equations (3.6)–(3.7). Parts (b) and (c) imply that

$$\frac{1}{2} - \mathbb{P}(\text{B wins}) \sim \frac{c}{2\sqrt{n}}, \qquad \frac{1}{2} - \mathbb{P}(\text{A wins}) \sim \frac{3c}{2\sqrt{n}}. \tag{1.1}$$

It is immediate that Alice and Bob are equally likely to win when $n = 1, 2$; by considering the eight possible outcomes when $n = 3$, we have $\mathbb{P}(\text{Bob wins}) - \mathbb{P}(\text{Alice wins}) = \frac{1}{8}$.

The methods of proof may be summarised as reversal, coupling, and renewal, and the proof of Theorem 1.2 illustrates these standard techniques. Reversing a random sequence is a long-established activity which has contributed enormously to probability and especially to the theory of random walks (see, for example, [7, Sect. 3.10]). The coupling approach enables us to study the 'pathwise' relationship between winning sequences for Alice and for Bob, rather than by simply calculating probabilities (see, for example, [7, Sect. 4.12]). Renewal theory is the study of random processes that renew themselves at random times (see for example, [7, Chap. 10]).

Remark 1.3. *The above problem may be phrased as '*HT *vs* HH*'. The methods described here yield the same conclusions for* TH *vs* HH*, for the following reason. Take a sequence ω of heads/tails and reverse it in time to obtain $\rho(\omega)$. Then ω and $\rho(\omega)$ have the same probability distribution, and the same count of consecutive head pairs* HH*. However, every* HT *in ω becomes* TH *in $\rho(\omega)$.*

As mathematicians we insist on unambiguous definitions of the objects of our study. One may capture pretty well all of probability theory by defining it as the theory of a countably infinite sequence of tosses of a fair coin. Perhaps not everything worth knowing is yet known about this primeval experiment.

Remark 1.4. *The target of this note is to obtain Theorem 1.2 using probabilistic methods. Since writing it, we have learned of related contributions available online. The idea of reversal has arisen in certain contributions on $\mathbb{X}$ to Litt's post. Mention is made also of [5] (using symbolic computation) and [11] (using analysis); each of these works includes a version of Theorem 1.2.*

2. Notation and basic observations

Here is some notation. Abbreviate Alice to A, and Bob to B; write H for a head and T for a tail. Let $\mathbb{N}$ denote the natural numbers, and let $\Omega_\infty = \{\mathrm{H}, \mathrm{T}\}^{\mathbb{N}}$ be the set of all sequences $\omega = (\omega_1, \omega_2, \dots)$ each element of which is either H or T. We shall normally express such sequences as $\omega = \omega_1\omega_2\cdots$. Similarly, $\Omega_n := \{\mathrm{H}, \mathrm{T}\}^n$, the set of all sequences of n coin tosses.

Let $\omega \in \Omega_\infty$. Player A scores -1 at each appearance of HH; player B scores 1 at each appearance of HT. The score $S_k(\omega)$ is the aggregate score after k steps of ω, that is, $S_k(\omega)$ is the number of appearances of HT minus the number of appearances of HH in the subsequence $\omega_1\omega_2\cdots\omega_k$.

Here is an outline of the method to be followed here. First, one defines epochs of renewal, at which the scoring process restarts. Neither player scores until the first appearance of H.

(i) If the following flip is T then Bob enters a winning period (of some length l_{B}), which lasts until the next time that the aggregate score is 0; this must happen at an appearance of HH.
(ii) If the following flip is H then Alice enters a winning period (of some length l_{A}), which lasts until the next time that the aggregate score is 0; this must happen at an appearance of HT.

The process then restarts (subject to certain details to be made specific). After n coin flips, Alice is winning if she is then in a winning period, and similarly for Bob. Ties occur between winning periods. We will see that the representative periods l_{A}, l_{B} are such that l_{A} is (stochastically) smaller than l_{B}, and Theorem 1.2(a) will follow.

The stochastic domination is proved by displaying a concrete coupling of l_{A} and l_{B}. The above explanation is given in more detail in Section 4.

The asymptotic part (b) of Theorem 1.2 is of course connected to the local central limit theorem (see, for example, [7, p. 219]). It suffices to use the earliest such theorem ever proved, namely the 1733 theorem of de Moivre [3, 10] (see also [4, p. 243–254] and [12, Thm 1.1]), although one may equally use the (de Moivre–)Stirling formula alone (see the historical note [9]).

We have assumed implicitly in the above that the lengths l_{A}, l_{B} are finite (almost surely), and this requires proof. It follows from the next lemma.

Let $\mathbb{P}$ be the probability measure on Ω_∞ under which the coin tosses are independent random variables, each being H (respectively, T) with probability $\frac{1}{2}$. An event E is said to occur *almost surely* if $\mathbb{P}(E) = 1$.

Lemma 2.1. *The number of times r at which the aggregate score satisfies $S_r(\omega) = 0$ is infinite almost surely.*

Remark 2.2. *Many of the arguments of this note are valid in the more general setting where heads occurs with some probability $p \in (0, 1)$. However, the conclusion of Lemma 2.1 is false when $p \neq \frac{1}{2}$, and indeed $\mathbb{P}(\mathrm{A\ wins})$ approaches 1 if $p > \frac{1}{2}$, and approaches 0 if $p < \frac{1}{2}$, in the limit as the number of coin flips grows to ∞.*

Proof of Lemma 2.1. Write $1[A]$ for the indicator function of an event A. The score $S(r)$ after r appearances of T forms a random walk with typical jump-size $\Delta = (2-R)1[R \geq 1]$ where R is the length of a typical run of heads; we allow $R = 0$, so that $\mathbb{P}(R = r) = (\frac{1}{2})^{r+1}$ for $r \geq 0$. Now,

$$\mathbb{E}(\Delta) = \sum_{r=1}^{\infty} (2-r)\left(\tfrac{1}{2}\right)^{r+1} = 0.$$

By a classical theorem of Chung and Fuchs [1] (see, for example, [7, Thm 5.10.16]) this random walk is recurrent, and therefore the set $\{r : S(r) = 0\}$ is almost surely infinite. □

3. Proof of Theorem 1.2(a)

A finite sequence $\omega = \omega_1 \cdots \omega_k$ is called

a B-*excursion* if it starts HT and ends HH, and satisfies:

$$S_l(\omega) > 0 \text{ for } 1 < l < k, \text{ and } S_k(\omega) = 0,$$

an A-*excursion* if it starts HH and ends HT, and satisfies:

$$S_l(\omega) < 0 \text{ for } 1 < l < k, \text{ and } S_k(\omega) = 0,$$

an Â-*excursion* it comprises an A-excursion followed by a (possibly empty) run of tails and then a single head.

We write $\mathcal{B}$ for the set of B-excursions, and $\mathcal{B}_k$ for those of length k, with similar notation for A-excursions and Â-excursions (using the notation $\mathcal{A}$ and $\hat{\mathcal{A}}$, respectively).

Â-excursions are introduced for the following reason. Players A and B can score only following an appearance of H. We shall use an embedded renewal argument, and to that end it is convenient that excursions finish with a head. Since an A-excursion ends with a tail, we simply extend it to include any following run of tails, followed by the subsequent head; this extension contains neither HH nor HT and hence the score of the excursion is unchanged.

Remark 3.1. *A* headrun *(respectively,* tailrun*) is a maximal consecutive sequence of heads (respectively, tails). Any finite sequence ω comprises alternating headruns and tailruns. Alice's score equals the number of heads minus the number of headruns. Bob's score equals the number of tailruns having a preceding head. Let h be the total number of heads, and r the total number of runs (each comprising either heads or tails). The aggregate score $S(\omega)$ equals $r - h$ if $\omega_1 = \mathrm{H}$ (respectively, $r - 1 - h$ if $\omega_1 = \mathrm{T}$).*

We define sequence-reversal next. Let $\Phi = \bigcup_{k=1}^{\infty} \Omega_k$ be the set of all non-empty finite sequences, and let Φ^{H} be the subset of Φ containing all finite sequences that begin H. For $\omega \in \Phi$, we write $S(\omega)$ for the aggregate score of ω. The score function is additive in the sense that

$$S(\omega_1 \cdots \omega_{m+n}) = S(\omega_1 \cdots \omega_m) + S(\omega_m \cdots \omega_{m+n}), \qquad m, n \geq 1. \tag{3.1}$$

For a finite sequence $\omega = \omega_1\omega_2 \dots \omega_k \in \Phi$, define its *reversal* $\rho(\omega)$ by $\rho(\omega) = \omega_k\omega_{k-1} \dots \omega_1$.

Lemma 3.2.

(a) *If $\omega = \omega_1 \cdots \omega_k \in \Phi^{\mathrm{H}}$ ends in* H *(so that $\omega_1 = \omega_k = \mathrm{H}$), then $S(\omega) = S(\rho(\omega))$.*

(b) *Let $k \ge 2$. The mapping ρ is a measure-preserving bijection between the sets $\mathcal{B}_k$ and $\hat{\mathcal{A}}_k$.*

Proof. (a) Let $\omega \in \Phi^{\mathrm{H}}$ end in H, so that $\rho(\omega) \in \Phi^{\mathrm{H}}$. By Remark 3.1, $S(\omega)$ equals the number of runs minus the number of heads. These counts are invariant under ρ.

(b) A B-excursion $b = b_1 b_2 \cdots b_k \in \mathcal{B}_k$ has the form $b = \mathrm{HT}^r\mathrm{H}\cdots\mathrm{HH}$ for some $r \ge 1$, whence $\rho(b) = b_k b_{k-1} \cdots b_1$ has the form $\mathrm{HH}\cdots\mathrm{HT}^r\mathrm{H}$. By part (a), $S(\rho(b)) = S(b) = 0$, which implies

$$S(b_k b_{k-1} \cdots b_{r+2}) = 0. \tag{3.2}$$

We claim that

$$S(b_k b_{k-1} \cdots b_j) < 0 \qquad \text{for } j \in [r+3, k-1]. \tag{3.3}$$

Suppose, on the contrary, that $j \in [r+3, k-1]$ is such that $S(b_k b_{k-1} \cdots b_j) = 0$. Then $s = \max\{t \le j : \omega_t = \mathrm{H}\}$ is such that $s \ge r+2$ and $S(b_k b_{k-1} \cdots b_s) = 0$. By part (a) and the fact that $b \in \mathcal{B}_k$, we have $S(b_s b_{s+1} \cdots b_k) = S(b_k b_{k-1} \cdots b_s) = 0$. By the additivity of S, (3.1),

$$S(b_1 b_2 \cdots b_s) = S(b_1 b_2 \cdots b_k) - S(b_s b_{s+1} \cdots b_k) = 0,$$

in contradiction of the assumption that b is a B-excursion, and (3.3) follows.

By (3.2)–(3.3), $b_k b_{k-1} \cdots b_{r+2}$ is an A-excursion, and hence $\rho(b)$ is an Â-excursion. Therefore, ρ is an injection. By the same argument applied to ρ^{-1}, we have that ρ is a surjection. The measure-preserving property follows from the fact that ω and $\rho(\omega)$ have the same numbers of heads and tails. □

We make three observations, in amplification of the remarks of the last section. Let $\omega \in \Omega_\infty$, and observe the initial evolution of scores.

(a) There is no score until the first H appears. Let M be the position of the first H, and note that $\mathbb{P}(M = m) = (\frac{1}{2})^m$ for $m \ge 1$. That is, there is an initial sequence of tails of some random length $M - 1$ (≥ 0), followed by a head (so that $\omega = \mathrm{T}^{M-1}\mathrm{H}\cdots$). Following this head, there is equal probability of H or T.

(b) If the next toss is T (which is to say that $\omega_{M+1} = \mathrm{T}$), then Bob scores one point. He remains in the lead until the next epoch, $M + N$ say, at which the aggregate score equals 0. It must be the case that $\omega_{M+N-1}\omega_{M+N} = \mathrm{HH}$, and thus the sequence $\omega_M \cdots \omega_{M+N}$ is a B-excursion. The process restarts from H at time $M + N$.

(c) Suppose the next toss is H (i.e., $\omega_{M+1} = \mathrm{H}$). The situation is slightly more complicated in this case. At time $M + 1$, Alice leads by 1, and she continues in the lead until the next epoch, $M + Q$ say, when the aggregate score equals 0, and moreover $\omega_{M+Q-1}\omega_{M+Q} = \mathrm{HT}$. Thus the sequence $\omega_1 \cdots \omega_{M+Q}$ is an A-excursion. Whereas in (b) the process restarts from the state $\omega_{M+N} = \mathrm{H}$, this time we have $\omega_{M+Q} = \mathrm{T}$, and we wait a random period of time for the next H. As in (a), there is a run of tails before the next head, which is to say that the process restarts from the head at epoch $M + Q + M'$ where M' is an independent copy of M. The sequence $\omega_M \cdots \omega_{M+N+M'}$ is an Â-excursion.

We shall adjoin certain sequences by placing them in tandem, and we write either $\psi_1\psi_2$ or $\psi_1 \cdot \psi_2$ for ψ_1 followed by ψ_2. When scoring ψ_2, viewed as a subsequence of $\psi_1\psi_2$, one takes account of the final character of ψ_1, which will typically be H in the cases studied here. In order to do the necessary book-keeping we introduce the following notation: if ψ is a sequence beginning H, we write $\mathrm{H}^{-1}\psi$ for the sequence obtained from ψ by removing its initial H.

Let $(\tau_i : i \ge 1)$ be a sequence of independent random elements of $\mathcal{A}$, and let Q_i+1 be the length of τ_i (so that $\mathrm{H}^{-1}\tau_i$ has length Q_i). The τ_i are almost surely finite, by Lemma 2.1. Let $(M_i : i \ge 0)$ be independent copies of M, also independent of the Q_i. Let $N_i = Q_i + M_i$, and let $L_i = \mathrm{T}^{M_i-1}\mathrm{H}$.

We have that $\tau_i L_i$ is a random element of $\hat{\mathcal{A}}$ of length $Q_i + M_i + 1$. By Lemma 3.2, $\rho(\tau_i L_i)$ is a random element of $\mathcal{B}$ of length $Q_i + M_i + 1$. We define

$$\text{the Â-excursion } \alpha_i := \tau_i L_i,$$
$$\text{the B-excursion } \beta_i := \rho(\tau_i L_i).$$

Thus β_i is a simple reversal of α_i, and this fact will provide a coupling of Â-excursions and B-excursions which is length-conserving in that α_i and β_i have the same length.

Both α_i and β_i start and end with H; when placing them in tandem we shall strip the initial H. Another way of expressing (a)–(c) is as follows.

(i) A random sequence of coin tosses begins L_0.
(ii) This is followed by

$$\gamma_1 := \begin{cases} \mathrm{H}^{-1}\alpha_1 & \text{with probability } \frac{1}{2}, \\ \mathrm{H}^{-1}\beta_1 & \text{with probability } \frac{1}{2}. \end{cases}$$

(iii) Let $k \ge 2$, and suppose α_i, β_i have been constructed for $i = 1, 2, \dots, k$. We then let

$$\gamma_{k+1} := \begin{cases} \mathrm{H}^{-1}\alpha_{k+1} & \text{with probability } \frac{1}{2}, \\ \mathrm{H}^{-1}\beta_{k+1} & \text{with probability } \frac{1}{2}. \end{cases} \tag{3.4}$$

Note that the length of γ_i is $Q_i + M_i$.

Lemma 3.3. *The sequence $X = L_0\gamma_1\gamma_2\cdots$ is an independent sequence of fair coin tosses.*

This does not require proof beyond the above remarks. The lemma provides a representation of a random sequence in terms of an initial tailrun, followed by independent copies of γ_1 in tandem. Each γ_i occupies a time-slot, and within this slot there appears a sequence of coin-tosses; according to the flip of another fair coin, we either retain this sequence or we reverse it (see (3.4)). (Some minor details concerning initial and final heads are omitted from this overview). Lemma 3.3 provides the setting for a coupling of Â-excursions and B-excursions.

Proof of Theorem 1.2(a). Let X be as in Lemma 3.3. For a finite subsequence $\delta = X_r X_r \cdots X_s$, write $\overline{\delta} = \{r, r+1, \dots, s\}$ for the set of times spanned by δ, and $\delta^\circ = \overline{\delta} \setminus \{s\}$.

Let $n \geq 3$, and find the unique I such that $n \in \gamma_I$ (write $I = 0$ if $n \in L_0$). At epoch n,

$$
\begin{aligned}
S_n &= 0 \quad \text{if} \quad && I = 0,\\
S_n &> 0 \quad \text{if} \quad && I \geq 1,\, \mathrm{H}\gamma_I = \beta_I \in \mathcal{B} \text{ and } n \in \beta_I^\circ,\\
S_n &< 0 \quad \text{if} \quad && I \geq 1,\, \mathrm{H}\gamma_I = \tau_I L_I \in \hat{\mathcal{A}} \text{ and } n \in \tau_I^\circ,\\
S_n &= 0 \quad \text{if} \quad && \text{none of the above conditions hold.}
\end{aligned} \tag{3.5}
$$

Recalling (3.4), it follows that

$$
\begin{aligned}
\mathbb{P}(S_n > 0) - \mathbb{P}(S_n < 0) &= \tfrac{1}{2}\big[\mathbb{P}(I \geq 1,\, n \in \beta_I^\circ) - \mathbb{P}(I \geq 1,\, n \in \tau_I^\circ)\big]\\
&= \tfrac{1}{2}\mathbb{P}(I \geq 1,\, n \in \beta_I^\circ \setminus \tau_I^\circ) \geq 0.
\end{aligned} \tag{3.6}
$$

The strict positivity of the last probability (with $n \geq 3$) follows by consideration of sequences beginning $\mathrm{H}^2\mathrm{T}^m\mathrm{H}\cdots$ with $m \geq n$, for which $M_0 = 1$, $M_1 = m$, and $\gamma_1 = \mathrm{H}\mathrm{T}^m\mathrm{H}$. Thus, $\tau_1 = \mathrm{HHT}$, and the interval $\beta_1^\circ \setminus \tau_1^\circ = \{3, m+2\}$ contains the value n. $\square$

The probability of a tie is derived similarly to (3.6) as

$$
\mathbb{P}(S_n = 0) = \mathbb{P}(I = 0) + \mathbb{P}(I \geq 1,\, n \in \beta_I \setminus \beta_I^\circ) + \tfrac{1}{2}\mathbb{P}(I \geq 1,\, n \in \beta_I^\circ \setminus \tau_I^\circ). \tag{3.7}
$$

From (3.6)–(3.7) (as in (1.1)) one obtains representations for $\mathbb{P}(S_n > 0)$ and $\mathbb{P}(S_n < 0)$.

4. Proof of Theorem 1.2(b, c)

Let R be a renewal process with inter-renewal times distributed as $L_0 \cdot \mathrm{H}^{-1}\tau_1$, and let $\pi_m := \mathbb{P}(m \in R)$. By the renewal property, conditional on the event $\{m \in R\}$, the sequence $\psi_{m,n} := X_{m+1}X_{m+2}\cdots X_n$ is a sequence of independent coin flips. Since the events $\{m \in R,\, \psi_{m,n} = \mathrm{T}^{n-m}\}$, $m = 1, 2, \ldots, n$, are disjoint, we have by (3.6) that

$$
\begin{aligned}
\mathbb{P}(S_n > 0) - \mathbb{P}(S_n < 0) &= \frac{1}{2}\sum_{m=3}^{n} \mathbb{P}(m \in R,\, \psi_{m,n} = \mathrm{T}^{n-m})\\
&= \sum_{k=0}^{n-3} (\tfrac{1}{2})^{k+1}\pi_{n-k}.
\end{aligned} \tag{4.1}
$$

Theorem 1.2(b) follows once the following has been proved.

Proposition 4.1. *We have that $\pi_m \sim c/\sqrt{m}$ as $m \to \infty$, where $c = 1/(2\sqrt{\pi}) \approx 0.282$.*

The proof uses the de Moivre local central limit theorem (as in [12]), though this amounts only to sustained use of the (de Moivre–)Stirling formula, [9]. An outline of the proof of Theorem 1.2(c) is included after that of the proposition.

Proof. We call $r \geq 1$ a τ-*endtime* if, for some $i \geq 1$, r is the final end-time of τ_i (otherwise written, $r \in \tau_i \setminus \tau_i^\circ$). Let R_X be the set of τ-endtimes. It is immediate that $R \subseteq R_X$. Moreover, by (3.4), any $r \in R_X$ satisfies $r \in R$ with probability $\frac{1}{2}$ (these events being independent for different r). Therefore,

$$
\mathbb{P}(m \in R) = 2\mathbb{P}(m \in R_X), \qquad m \geq 1. \tag{4.2}
$$

By Lemma 3.3, the event $\{m \in R_X\}$ is the subset of all $\omega \in \Omega_m$ satisfying $\omega_{m-1}\omega_m = \mathrm{HT}$ and $S(\omega) = 0$. Let h be the number of heads, and r the number of runs (of either heads or tails). If r is even, (respectively, odd), then such sequences begin H (respectively, T), and they invariably finish with a tailrun of exactly one tail. By Remark 3.1, since the aggregate score is 0, we have $h = r$ if r is even, and $h = r - 1$ if r is odd.

We use the fact that the number of p-partitions of the integer q is $\binom{q-1}{p-1}$. By partitioning the heads and tails into r runs subject to the above, and interleaving headruns and tailruns, we obtain that

$$|\{m \in R_X\}| = \Sigma_1 + \Sigma_2$$

where

$$\Sigma_1 = \sum_{r \text{ odd}} \binom{m-h-2}{\frac{1}{2}(r+1)-2}\binom{h-1}{\frac{1}{2}(r-1)-1} = \sum_{r \text{ odd}} \binom{m-r-1}{\frac{1}{2}(r-3)}\binom{r-2}{\frac{1}{2}(r-3)},$$

$$\Sigma_2 = \sum_{r \text{ even}} \binom{m-h-2}{\frac{1}{2}r-2}\binom{h-1}{\frac{1}{2}r-1} = \sum_{r \text{ even}} \binom{m-r-2}{\frac{1}{2}r-2}\binom{r-1}{\frac{1}{2}r-1}.$$

Set $r = 2s+1$ (respectively, $r = 2s$) in Σ_1 (respectively, Σ_2) and add to obtain

$$\begin{aligned} |\{m \in R_X\}| &= \sum_{s=2}^{\infty} \left[\binom{m-2s-2}{s-1} + \binom{m-2s-2}{s-2}\right]\binom{2s-1}{s-1} \\ &= \sum_{s=2}^{\infty} \binom{m-2s-1}{s-1}\binom{2s-1}{s-1}. \end{aligned} \tag{4.3}$$

Therefore,

$$\mathbb{P}(m \in R_X) = 2^{-m}|\{m \in R_X\}| = \frac{1}{4}\sum_{s=2}^{\infty} \mathbb{P}(T_{m-2s-1} = s-1)\mathbb{P}(T_{2s-1} = s-1), \tag{4.4}$$

where T_k has the binom$(k, \frac{1}{2})$ distribution. By Stirling's formula (or [12, Thm 1.1]),

$$\mathbb{P}\left(T_{2s-1} = s-1\right) \sim \frac{1}{\sqrt{\pi s}} \quad \text{as } s \to \infty. \tag{4.5}$$

We may occasionally use real numbers in the following where integers are expected. The term $\mathbb{P}(T_{m-2s-1} = s-1)$ in (4.4) is a maximum when $m - 2s - 1 = 2(s-1)$, which is to say that $s = \frac{1}{4}(m+1)$. Let $\gamma \in (\frac{1}{2}, \frac{2}{3})$. We may restrict the summation in (4.4) to values of s satisfying $|s - \frac{1}{4}m| < m^\gamma$. To see this, note that

$$\sum_{s \ge \frac{1}{4}m + m^\gamma} \mathbb{P}(T_{m-2s-1} = s-1) \le \sum_{s \ge \frac{1}{4}m + m^\gamma} \mathbb{P}(T_{m/2} \ge s-1),$$

which tends to 0 as $m \to \infty$ by the moderate-deviation theorem of Cramér [2] (or, for a more modern treatment, see Feller [6, p. 549]). We have used the fact that T_k is stochastically increasing in k. A similar argument applies for $s \le \frac{1}{4}m - m^\gamma$.

Suppose $|s - \frac{1}{4}m| < m^\gamma$. By [12, Thm 1.2], there is an absolute constant C such that

$$\left|\mathbb{P}(T_{m-2s-1} = s-1) - \sqrt{\frac{2}{\pi(m-2s-1)}}\exp\left(-\frac{(4s-m-1)^2}{2(m-2s-1)}\right)\right| \le \frac{C}{m^{2/3}}.$$

It follows that

$$\sum_{|s-\frac{1}{4}m|<m^\gamma} \mathbb{P}(T_{m-2s-1} = s-1)$$

deviates from

$$\Psi_m := \sum_{|s-\frac{1}{4}m|<m^\gamma} \sqrt{\frac{2}{\pi(m-2s-1)}} \exp\left(-\frac{(4s-m-1)^2}{2(m-2s-1)}\right)$$

by at most $C(2m^\gamma+1)/m^{2/3}$, which tends to 0. Express Ψ_m as an integral, make the change of variable

$$\beta = \frac{4s-m-1}{\sqrt{m-2s-1}}$$

and let $m \to \infty$ to obtain, by the dominated convergence theorem, that

$$\Psi_m \to \frac{1}{2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\beta^2}\, d\beta = \frac{1}{2}.$$

Therefore, the terms $\mathbb{P}(T_{m-2s-1} = s-1)$ in (4.4) are (asymptotically) concentrated in the interval $\frac{1}{4}m \pm m^\gamma$, with total weight $\frac{1}{2}$. Hence, by (4.2) and (4.4)–(4.5),

$$\pi_m \sim \frac{2}{8\sqrt{\pi m/4}} = \frac{1}{2\sqrt{\pi m}},$$

as claimed. □

Outline proof of Theorem 1.2(c). One may deduce (c) from (3.7) by adapting the argument leading to (4.1). Alternatively, one may perform a direct calculation as above, and there follows a sketch of this. Let Z_m be the set of vectors $\omega \in \Omega_m$ such that $S(\omega) = 0$. Elements $\omega \in Z_m$ may be expressed as interleaved headruns and tailruns, with the counts of heads and tailruns being balanced by the condition $S(\omega) = 0$ (see Remark 3.1). Such ω may start with either H or T, and each case leads to two terms as in the first line of (4.3). Thus, $|Z_m|$ is the sum of four terms, each being the product of two binomial coefficients. The analysis continues as in the above proof. □

Acknowledgements

GRG thanks Daniel Litt for a sweet question, and Charles McGinnes for drawing his attention to it. He thanks Richard Weber and David Stirzaker for useful correspondence.

References

[1] K. L. Chung and W. H. J. Fuchs, *On the distribution of values of sums of random variables*, Mem. Amer. Math. Soc. **6** (1951), 1–12.

[2] H. Cramér, *Sur un nouveau théorème-limite de la théorie des probabilités*, Actualités Scientifiques et Industrielles **736** (1938), 2–23, transl. H. Touchette, https://arxiv.org/abs/1802.05988.

[3] A. de Moivre, *Approximatio ad Summam Terminorum Binomii $\overline{a+b}^n$ in Seriem expansi*, (1733).

[4] ———, *The Doctrine of Chances*, 3rd ed., 1756, https://archive.org/details/doctrineofchance00moiv/page/n5/mode/2up,.

[5] S. B. Ekhad and D. Zeilberger, *How to answer questions of the type: if you toss a coin n times, how likely is HH to show up more than HT?*, (2024), https://arxiv.org/abs/2405.13561.

[6] W. Feller, *An Introduction to Probability Theory and its Applications. Vol. II*, 2nd ed., John Wiley & Sons., New York, 1971.

[7] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 4th ed., Oxford University Press, 2020.

[8] D. Litt, (2024), https://x.com/littmath/status/1769044719034647001.

[9] K. Pearson, *Historical note on the origin of the normal curve of errors*, Biometrika **16** (1924), 402–404.

[10] K. Pearson, A. de Moivre, and R. C. Archibald, *A Rare Pamphlet of Moivre and some of his Discoveries*, Isis **8** (1926), 671–683, https://www.journals.uchicago.edu/doi/epdfplus/10.1086/358439.

[11] S. Segert, *A proof that HT is more likely to outnumber HH than vice versa in a sequence of n coin flips*, (2024), https://arxiv.org/abs/2405.16660.

[12] Z. Szewczak and M. Weber, *Classical and almost sure local limit theorems*, Dissertationes Math. **589** (2023), 97 pp.

Centre for Mathematical Sciences, Cambridge University, Wilberforce Road, Cambridge CB3 0WB, UK

Email address: grg@statslab.cam.ac.uk