Analysis of the Chow-Robbins Game with Biased Coins

Arjun Mithal

May 17, 2018

Contents

Abstract

The rules of the Chow-Robbins game are simple: flip a coin repeatedly, and stop whenever you choose. The goal: maximize the ratio of heads to total flips. This simple problem has no known closed form solution, however [1] gives the tightest known bounds on the expected payoff under optimal play. The traditional formulation of the game assumes a fair coin ($p = q = 1/2$). This allows for relatively straightforward analysis. In this paper, we extend the analysis to games with a biased coin ($p = 1 - q \neq 1/2$). The treatment and approach of the analysis is similar to [1] and [2], though the results for biased coins are original.

1 Introduction to Chow-Robbins

Yuan-Shih Chow and Herbert Robbins formulated the problem now known as Chow-Robbins in 1965 [3] as part of a larger analysis of optimal stopping rules for i.i.d. random variables. The problem is simple: flip a coin as many times as desired, and receive a payoff equal to the ratio of heads to total coin flips.

In [1], Chow and Robbins show that there exists a stopping rule for this game that triggers with probability 1. Other analyses of the problem, developed independently by Dvoretzsky, Shepp, Häggström and Wästlund, and Medina and Zeilberger all focus on deriving asymptotically optimal conditions of the game. In particular, [1] derives bounds on different game states analytically and then computationally. [2] derives the asymptotic ratio of heads required to stop after n flips have been made. Finally, [4] shows that the optimality conditions for certain game states are still unknown and not proven.

We now continue by presenting a recursive approach to the problem. Much of the notation used in this paper is taken from [1], however the extensions to biased coins are original results.

2 Recursive Framework for Chow-Robbins

Let $V(h,n)$ denote the expected payoff under optimal play from a game state with h heads in the first n flips. The following recursive formula for $V(h,n)$ relates the options of either stopping or continuing from game state (h,n) under a fair coin:

$$V(h,n) = \max\left(\frac{h}{n}, \frac{V(h+1,n+1)+V(h,n+1)}{2}\right) \tag{1}$$

The payoff of stopping the game at game state (h,n) is exactly h/n. In the fair game, with probability $1/2$ a head is observed on the $n+1$ flip. So, with probability $1/2$, we receive the expected payoff of either game state $(h+1,n+1)$ or $(h,n+1)$.

Next, we will introduce an important probabilistic concept which will be used extensively in the remainder of our analysis. For the sake of generality, we specify the Lemma independent of the bias p:

Lemma 2.1. *From any finite game state (h,n), and for a coin with any bias p, we can achieve $V(h,n) \geq p$ with probability 1.*

Proof. Clearly, for $\frac{h}{n} \geq p$, we can stop the game immediately and collect payoff of at least p. For $\frac{h}{n} < p$, we make use of the Strong Law of Large Numbers (SLLN). In particular, the SLLN asserts the following:

> Let X be a real-valued random variable with finite mean $\bar{X}$, and let $X_1, X_2, \ldots$ be a sequence of i.i.d. copies of X. Denote $\bar{X_n}$ as the empirical average $\bar{X_n} := \frac{1}{n}(X_1 + \cdots + X_n)$. Then, $\mathbb{P}(\lim_{n\to\infty} \bar{X_n} = \bar{X}) = 1$.

In other words, the sample average of any sequence of i.i.d. random variables will converge in probability to the mean of one copy of the random variable. To apply this theorem to our problem, we notice that with probability 1, the game will converge to a game state where $h/n = p$. A corollary of the SLLN is that no finite intermediate state of the sequence $X_1, X_2, \ldots$ will contradict the asymptotic convergence of the sample mean.

Given the (potentially) infinite nature of Chow-Robbins, we can continue playing from (h, n) until the proportion p is achieved. The SLLN gives the notion that we will reach this stopping state with probability 1. This completes the proof.

□

Using the Lemma, we can now assert that for any game state,

$$V(h,n) \geq \max\left(\frac{h}{n}, p\right). \tag{2}$$

Moreover, we can generalize (1) to coins with bias p:

$$V(h,n) = \max\left(\frac{h}{n}, p \cdot V(h+1, n+1) + (1-p) \cdot V(h, n+1)\right) \tag{3}$$

Now, while (3) provides a useful recursive formula for solving $V(h, n)$, the only base cases have an infinite horizon ($n = \infty$). Since we are essentially solving using "backwards induction," knowledge of the extreme values of V are required in order to solve explicitly. In the absence of this knowledge, we can only do so well as to bound V for large values of n, and solve backwards.

3 Generalizing the Lower Bound

Combining equations (3) and (4), we offer a method to compute a lower bound $V^-(0, 0)$. The bound will be tight only up to the level of computation permitted. In particular, we choose N sufficiently large, and approximate $V^-(h, N)$ for $h \leq N$ using (2). So, we assign the following values:

$$V^-(h, N) = \begin{cases} p, & \text{for } h \leq p \cdot N \\ \frac{h}{N}, & \text{for } h > p \cdot N \end{cases}$$

Computing the $V^-(h, n)$ values for $n < N$ is now straightforward from (3). We use a dynamic programming algorithm to implement the recursion computationally, and report the values for $V^-(0, 0)$, the lower bound on the expected value of the game under optimal play. In addition, we report values for certain early game states. In particular, if $V^-(h, n) > h/n$, we know unequivocally that the optimal strategy in game state (h, n) is to continue. All results are summarized in section 5, and in Figures 1, 2, and 3.

4 Generalizing the Upper Bound

The tightest upper bound we present in this paper is far less straightforward than the lower bound, and requires far more computation to derive. The motivation for the Theorem presented in this section is also attributed to Häggström and Wästlund in [1]. The extension of the bound to biased coins is an original result.

Theorem 4.1.

$$V(h,n) \leq \max\left(\frac{h}{n}, p\right) + \int_{\max(\frac{h}{n}, p)}^{1} \left(\frac{p}{q}\right)^{\frac{qn-h}{1-q}} dq \tag{4}$$

Before proving this theorem, we must introduce some more notation: Let

$$P(h,n,q) := \mathbb{P}(\text{eventually achieving a proportion greater than } q \text{ from } (h,n)). \tag{5}$$

Phrased differently, $P(h,n,q)$ is the inverse CDF (Cumulative Distribution Function) of the random variable $X_{(h,n)}$, where

$$X_{(h,n)} := \text{maximum payoff possible from } (h,n). \tag{6}$$

We can now rewrite $P(h,n,q)$ as

$$P(h,n,q) = \mathbb{P}(X_{(h,n)} > q). \tag{7}$$

This gives that $P(h,n,q) = (1-\mathbb{P}(X_{(h,n)} \leq q)) = (1-CDF(X_{(h,n)}))$. Using the well known fact from probability theory that

$$\mathbb{E}[X] = \int_X (1 - CDF(X'))dX', \tag{8}$$

we can derive an explicit formula for $V(h,n)$:

$$V(h,n) = \int_0^1 P(h,n,q)\ dq. \tag{9}$$

4.1 Proof of Theorem 4.1

Now, in deriving an upper bound for $P(h,n,q)$, consider the following procedure: given that the proportion q is eventually attained from state (h,n), say on flip m, the proportion of heads in flips $n+1,\ldots,m$ must be at least q. So, the conditional probability of heads on each flip in $n+1,\ldots,m$ must be at least q. Let us now consider k to be the minimum possible m such that proportion q is reached from (h,n). In other words, flipping k consecutive heads from (h,n) will exceed proportion q. Using this definition, we have the following inequality:

$$\mathbb{P}(k \text{ consecutive heads flipped} \mid \text{proportion } q \text{ attained}) \geq q^k \tag{10}$$

From Bayes' Rule, we also know that

$$\mathbb{P}(k \text{ consecutive heads flipped} \mid \text{proportion } q \text{ attained}) \leq \frac{\mathbb{P}(k \text{ consecutive heads flipped})}{\mathbb{P}(\text{proportion } q \text{ attained})}.$$

Rearranging the formula, and substituting $P(h,n,q)$ for the denominator on the RHS, we have

$$P(h,n,q) \leq \frac{\mathbb{P}(k \text{ consecutive heads flipped})}{\mathbb{P}(k \text{ consecutive heads flipped} \mid \text{proportion } q \text{ attained})}.$$

The unconditional probability on the RHS is simply p^k, so substituting this quantity combined with the result of (10), we arrive at the upper bound

$$P(h,n,q) \leq \left(\frac{p}{q}\right)^k. \tag{11}$$

Before evaluating our integral formula, we first solve explicitly for k. Since flipping k consecutive heads from (h, n) gives a proportion strictly greater than q, we know that

$$\frac{h+k-1}{n+k-1} \leq q < \frac{h+k}{n+k},$$

and subsequently,

$$k \geq \frac{qn-h}{1-q}.$$

Since we also know $P(h, n, q) = 1$ for $q < \max(\frac{h}{n}, p)$, we can now rewrite $V(h, n)$ as

$$V(h,n) = \int_0^1 P(h,n,q)\ dq \leq \max\Big(\frac{h}{n}, p\Big) + \int_{\max(\frac{h}{n},p)}^1 \Big(\frac{p}{q}\Big)^{\frac{qn-h}{1-q}} dq. \tag{12}$$

□

This completes the proof of Theorem 4.1, however we now proceed to show a more rigorous formulation of the upper bound of $V(0, 0)$.

4.2 A More Rigorous Upper Bound

Applying the substitution $2q = 1 + t$, the integrand and bounds of integration change as follows,

$$\int_{\max(\frac{h}{n},p)}^1 \Big(\frac{p}{q}\Big)^{\frac{qn-h}{1-q}} dq = \int_{\max(\frac{2h-n}{n},2p-1)}^1 \Big(\frac{2p}{1+t}\Big)^{\frac{(1+t)n-2h}{1-t}} \frac{dt}{2}.$$

Using the inequality

$$\frac{\log(1+t)}{1-t} \geq t,$$

we arrive at

$$\begin{aligned}
&\int_{\max(\frac{2h-n}{n},2p-1)}^1 \Big(\frac{2p}{1+t}\Big)^{\frac{(1+t)n-2h}{1-t}} \cdot \frac{dt}{2} \\
&= \frac{1}{2}\int_{\max(\frac{2h-n}{n},2p-1)}^1 \exp\Big(\frac{(1+t)n-2h}{1-t}(\log(2p) - \log(1+t))\Big) dt \\
&\leq \frac{1}{2}\int_{\max(\frac{2h-n}{n},2p-1)}^1 \exp\Big(\log(2p)\frac{(1+t)n-2h}{1-t} - (1+t)tn + 2ht\Big) dt.
\end{aligned}$$

Next we set $u = t\sqrt{n}$ and obtain

$$\begin{aligned}
V(h,n) \leq &\max\Big(\frac{h}{n}, p\Big) + \\
&\frac{1}{2\sqrt{n}} \int_{\max(\frac{2h-n}{\sqrt{n}},(2p-1)\sqrt{n})}^{\sqrt{n}} \exp\Big(\log(2p)\frac{n\sqrt{n}+un-2h\sqrt{n}}{\sqrt{n}-u} - u^2 + \frac{2h-n}{\sqrt{n}} \cdot u\Big) du.
\end{aligned} \tag{13}$$

When we consider the case of a fair coin ($p = \frac{1}{2}$), we can replace the bounds of integration and achieve the form

$$V(h,n) \leq \max\Big(\frac{h}{n}, \frac{1}{2}\Big) + \frac{1}{2\sqrt{n}} \int_0^\infty \exp\Big(-u^2 - \frac{|2h-n|}{\sqrt{n}} \cdot u\Big) du. \tag{14}$$

The integral term in (14) can further be bounded by each of the terms in the exponential, which have closed form solutions,

$$\begin{aligned} \frac{1}{2\sqrt{n}}\int_0^\infty \exp(-u^2)du &= \frac{1}{4}\sqrt{\frac{\pi}{n}} \\ \frac{1}{2\sqrt{n}}\int_0^\infty \exp\Big(-\frac{|2h-n|}{\sqrt{n}}\cdot u\Big)du &= \frac{1}{2\cdot|2h-n|} \end{aligned}$$

Combining these solutions, the upper bound for the fair coin becomes

$$V(h,n) \le \max\Big(\frac{h}{n},\frac{1}{2}\Big) + \min\Big(\frac{1}{4}\sqrt{\frac{\pi}{n}}, \frac{1}{2\cdot|2h-n|}\Big), \tag{15}$$

as derived in [1].

The upper bound of the biased coin has no known closed form, so we instead present results based on a computational approximation of the upper bound for sufficiently large N. In particular, the computation is easiest on the version of the bound presented in (12), before any change of variables is performed. The resulting $V^+(h,N)$ values for our limiting horizon N are

$$V^+(h,N) = \begin{cases} p + \int_p^1 \Big(\frac{p}{q}\Big)^{\frac{qN-h}{1-q}} dq, & \text{for } h \le p\cdot N \\ \frac{h}{N} + \int_{\frac{h}{N}}^1 \Big(\frac{p}{q}\Big)^{\frac{qN-h}{1-q}} dq, & \text{for } h > p\cdot N \end{cases}$$

Again, we use a dynamic programming algorithm to compute all $V(h,n)$ values for $h \le n,\ n < N$ by (3).

5 Computational Results for $V(0,0)$

The results presented in this paper are based on an $n = 10^4$ horizon. Due to computational constraints, we were unable to reproduce the results produced in [1], which were built on an $n = 10^7$ horizon, and had an error margin of $< 10^{-5}$ for the $V(0,0)$ bound for a fair coin. Specifically, their results bounded $V(0,0)$ as

$$.79295301268091 < V(0,0) < .79295559864361,$$

whereas our results show

$$.792881121471 < V(0,0) < .793131491473,$$

giving bounds about 100 times as wide as [1]. The complete bounds for different values of p are plotted in Figures 1 and 2. The error, or bound tightness, as a function of the bias p is plotted in Figure 3. Interestingly, Figures 1 and 2 seem to indicate a smooth functional relationship between $V(0,0)$ and p.

6 Application to the "Stop When Ahead" Problem

One interesting corollary to the results presented in the previous section is the observation that, for some $p^\star \in (.2,.25)$, we first obtain $V(0,0) > .5$. This means that, in

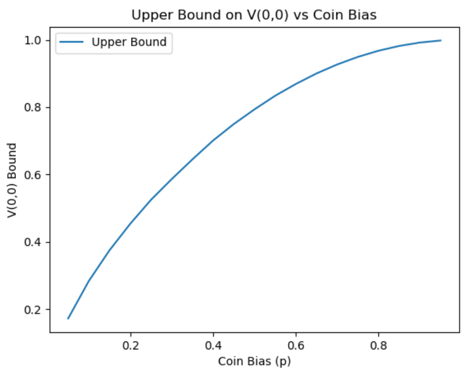

Figure 1: Upper bound on $V(0,0)$

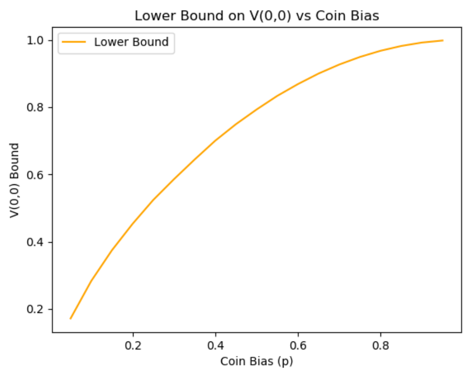

Figure 2: Lower bound on $V(0,0)$

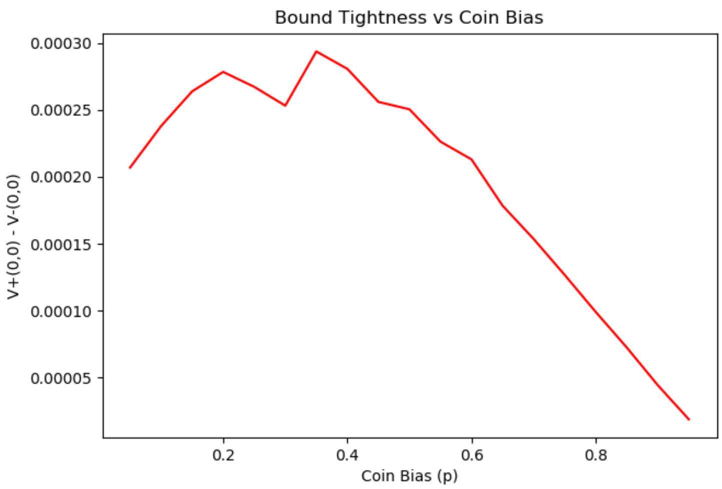

Figure 3: Bound Tightness

expectation, the optimal stopping rule for coin with bias $p^\star$ necessarily yields more heads than tails. Let us now apply this observation to a variation of a classic gambling problem. Consider a game with the following rules:

- You are given a (possibly biased) coin.
- Flip the coin repeatedly, and collect payoff equal to the ratio of heads to total flips minus the ratio of tails to total flips when you decide to stop.

When should you be willing to play this game? The results indicate that under optimal play, for all $p > p^\star$, the game has a positive expected value. This is even more interesting considering that for $p = .25$, you should play the game. Under this value of p, we already have a $.75 \times .75 = .5625$ probability of reaching state (0,2), so intuition might lead us to believe the game is not profitable. Our results show that we do, in fact, have sequence of optimal moves that will yield a profit in the game, on average.

7 Conclusion

In this paper, we extend the work of Häggström and Johan Wästlund in [1] to study the case of a biased coin in the Chow-Robbins game. We use a similar methodology to work through upper and lower bounds on the expected value of playing the game with a coin with particular bias p. Using a computational approach with a horizon $N = 10^4$, and solving backwards using the recursive formula in (3), we derive tight bounds on $V(0,0)$. The results are summarized in Figures 1 and 2. In each solution, if the value of a given $V^-(h,n) > h/n$, then the optimal move at game state (h,n) is necessarily to continue. Likewise, if $V^+ = h/n$, then the optimal move is to stop at (h,n). With more computational resources, we should be able to approximate arbitrarily tight bounds for $V(0,0)$. Finally, Figures 1 and 2 seem to indicate the existence of a smooth functional relationship between $V(0,0)$ and p. The search for such a function, as well as a closed form for $V(0,0)$ remain open problems.

References

[1] Olle Häggström and Johan Wästlund, *Rigorous computer analysis of the Chow Robbins game*, The American Mathematical Monthly 120(10): 893-900 (2013).

[2] Larry A. Shepp, *Explicit solutions to some problems of optimal stopping*, The Annals of Mathematical Statistics, 40:993-1010, 1969.

[3] Yuan-Shih Chow and Herbert Robbins, *On optimal stopping rules for s_n/n*, Ill. J. Math., 9: 444-454, 1965.

[4] Luis A. Medina and Doron Zeilberger, *An Experimental Mathematics Perspective on the Old, and still Open, Question of When To Stop?*, in "Gems in Experimental Mathematics", Contemporary Mathematics series v. 517 (AMS), eds T. Amdeberhan, L. Medina, and V. Moll, 265-274, also arXiv:0907.0032v2 [math.PR].