

Optimizing Hand Cricket

Atharv Nema

September 2023

Introduction

The goal of this project is to find out how to optimally play hand cricket. The rules of hand cricket are as follows:

1. This is a 2-player game
2. First, a toss occurs(for the toss, the players choose either odd or even. Then, both play numbers from 1 to 6. If the sum of both the numbers is odd, the player who chose odd wins the toss. Similarly for even). Whoever wins the toss gets to decide if he or she wants to bat or bowl first.
3. The first inning starts. The batsman's runs are initialized to 0.
4. Each player chooses a number from 1 to 6. If both the numbers are different, add the number the batsman played to his runs. If the numbers are the same, the batsman is out and now it is the bowler's chance to bat(notice that if the batsman gets out, the number he played is not added to his runs).
5. After the batsman gets out, the second inning starts. Repeat the same process, but now the batsman will bowl and the bowler will bat.
6. Whoever has more runs at the end of both the innings wins the game. If both have the same runs, the game ends in a draw.

I mentioned that I wanted to play this game optimally. Now, what do I mean by optimally? Optimally, intuitively to me means trying to play the best I can. However, this is still very vague. We will explore this concept of 'optimal' in the following sections.

NOTE: From sections 1 to 6, I will try to maximize(if I am the batsman) or minimize(if I am the bowler) the expected value of runs. This is close to optimal(but not exactly)

1 Optimal play of bowler for '*dice*' play of batsman

What I mean by *dice* play is that the batsman chooses every number with equal probability. This is equivalent to him(the batsman) throwing a dice and playing the number that shows on the top, hence the name dice play. I want to find what is the optimal strategy for the bowler in this case. Obviously, we are assuming that the batsman will not change his strategy in the course of his play. What do I mean by '*optimal*' here? I mean that the bowler should play such that the number of runs scored by the batsman is minimized. I claim that the optimal strategy is to only play 6. My intuition is that as the batsman is playing each number with an equal chance, no matter what the bowler plays, he has the same probability of getting the batsman out(which is $1/6$). If the bowler plays 6, when the batsman gets out, the number he played when he got out(which is 6) will not be counted in his total runs. In other words, the batsman will never be able to get 6 runs in one move. As 6 is the largest number he can play and the bowler is preventing him from playing it, the bowler is minimizing the batsman's runs.

Now, I will calculate the expected value of the batsman if the bowler plays 'optimally' in this case. Let the expected value be E . Say the batsman played a number that is not 6 (this has a probability of $5/6$). What is the average number of runs he would get by this play(given that he did not play 6)? It would be:

$$\frac{(1 + 2 + 3 + 4 + 5)}{5} = 2.5$$

Now after the batsman has got an average of 2.5 runs, what is the expected value of the runs he will get in his future play(not counting the runs he got in the first turn)? **It is E !** This is because the batsman's situation is the same as it was in the beginning. Hence we can calculate the expected value by the following formula

$$\frac{5}{6} * (2.5 + E) + \frac{1}{6} * 0 = E$$

Why is there a 0 multiplied to $1/6$? This is because if the batsman plays 6 on his first turn, he gets 0 runs. Hence $E = 12.5$

2 Optimal play of batsman for '*dice*' play of bowler

I claim that the optimal play of the batsman is to only play 6(notice here that the goal of the batsman is to maximize his runs). My logic is that as the expected value of the number of turns the batsman has is constant no matter what number he plays(as the bowler is in dice play), it is optimal for the batsman to maximize the runs he gets in each turn(hence, he should play 6).

Now, I will calculate the expected value of the batsman. Let the expected value be E . Say that the batsman does not get out in his first move (the bowler did not play 6). This has a probability of $5/6$. What is the expected value of the runs he will get in all future moves (not counting the runs in the first turn)? **It is E .** Hence we can calculate the expected value by the following formula.

$$\frac{5}{6} * (6 + E) + \frac{1}{6} * 0 = E$$

Hence, $E = 30$.

3 Optimal play of bowler for a particular probability distribution of the batsman

Let $P(i)$ represent the probability that the batsman plays i . I want to find out what is the optimal strategy of the bowler. As always, the batsman cannot change his strategy and it is the bowler's goal to minimize the batsman's runs. I claim that it will be optimal for the bowler to only play W . Here W is any number from 1 to 6. I know that this claim seems pretty sus, but it has a nice proof. See the end of the document for the proof of this claim (or try thinking why this is true). Basically, the bowler needs to calculate the expected value of the runs of the batsman if he (the bowler) plays only 1, only 2, only 3... and then choose the number that leads to the minimum runs as W .

I want to calculate the expected value of runs of the batsman if the bowler only plays some number x . The x that leads to the minimum expected value of the runs of the batsman will be chosen as W . We will follow the same approach we used in sections 1 and 2 to find the expected value for x (which I am calling $E(x)$). Say on the batsman's first turn, he does not get out (probability of this happening is $1 - P(x)$). What is the expected value of the runs he gets? Is it just $\left(\sum_{i=1}^6 P(i) * i\right) - P(x) * x$? Well, expected value is defined as $\left(\sum P(i) * i\right)$ and as we are considering the case where the batsman does not play x , we should remove that case (hence the $-P(x) * x$ term). However, we are missing an important detail. As we have already considered that the batsman has not played x , to calculate the expected value of his turn, we need to use the conditional probability (i.e. instead of $P(i)$, we need the probability of the batsman playing i given that he did not play x). Using Bayes theorem, this conditional probability comes out to be $\frac{P(i)}{1-P(x)}$.

Hence we can calculate $E(x)$ by the following equation.

$$(1 - P(x)) * \left(\frac{\sum_{i=1}^6 P(i) * i - P(x) * x}{1 - P(x)} + E(x) \right) = E(x)$$

Notice that the $E(x)$ term inside the bracket represents the expected value of

the runs the batsman gets in all future moves. Manipulating this, we get

$$E(x) = \frac{\sum_{i=1}^6 P(i) * i - P(x) * x}{P(x)} \quad (1)$$

So, to find the optimal play of the bowler, we just need to plug in all numbers from 1 to 6 into $E(x)$ and find the 'W' that gives the minimum expected value. The bowler should then only play this number.

4 Optimal play of batsman for a particular probability distribution of the bowler

Let $P(i)$ represent the probability that the bowler plays i . I want to find out what is the optimal strategy of the batsman. I claim that it will be optimal for the batsman to only play W . Here W is any number from 1 to 6. Again, you can find the proof of this claim at the end of the document. Here W is the number that gives the maximum expected value of runs of the batsman. Let's calculate $E(x)$ for some x between 1 and 6. Say that the batsman does not get out on the first turn(probability is $1 - P(x)$). He gets x runs and then is expected to get $E(x)$ runs for his future plays.

$$(1 - P(x)) * (x + E(x)) = E(x)$$

From this, we get:

$$E(x) = \frac{1 - P(x)}{P(x)} * x \quad (2)$$

So, to find the optimal play of the batsman, we just need to plug in all numbers from 1 to 6 into $E(x)$ and find the 'W' that gives the maximum expected value. The batsman should then only play this number.

Nash equilibrium

Say there are 2 completely rational players playing this game. These players can change their strategy. They play a large number of games with each other and then reach a stage where their strategy does not change. What do I mean by their strategy not changing? I mean that by changing their strategies, neither of the rational players will benefit. In other words, given the batsman probability distribution, the bowler's probability distribution is optimal(in compliance with formula 1), and given this 'optimal' distribution of the bowler, the batsman's distribution is optimal(in compliance with formula 2)(as both players have played a large number of games with each other, I am assuming that each player knows the others distribution). This situation is called a **Nash equilibrium**. My goal in the following section is to find the Nash equilibrium distribution for the bowler and the batsman(in the first inning).

NOTE: Notice that in sections 3 and 4, I mentioned that it is optimal to play only one number always. However, if 2 or more numbers give the optimal (either max or min) expected value, you can choose any probability distribution such that there is 0 probability for the numbers that are not optimal. For example, say 1, 3, and 5 give the optimal values. The following is an optimal distribution

x	1	2	3	4	5	6
P(x)	0.3	0	0.3	0	0.4	0

Now we need to notice an extremely important detail. Let us assume that some distribution function $P_1(x)$ represents the Nash equilibrium distribution of the batsman. Say that by using formula (1) on P_1 , the bowler gets < 6 numbers that lead to the min expected value of runs. This means that the bowler will never play 1 or more numbers. But if this is the case, the batsman will gain by changing his strategy to only playing the numbers the bowler does not play (this will lead to infinite runs). **Hence this is not a Nash equilibrium!** A similar argument can be used to prove that any probability distribution P_2 of the batsman that leads to the bowler playing < 6 numbers is not a Nash equilibrium distribution.

Hence, for P_1 and P_2 to be Nash equilibrium distributions for the bowler and batsman respectively, applying formula (2) on P_1 should lead to all the numbers from 1 to 6 having the same expected value. Similarly, applying formula (1) on P_2 should lead to all numbers from 1 to 6 having the same expected value. (Let us call this **claim Nash**)

5 Nash equilibrium distribution for the bowler

Let the probability distribution function of the bowler be P . Using claim Nash and applying (2)

$$\frac{1 - P(1)}{P(1)} * 1 = \frac{1 - P(2)}{P(2)} * 2 = \dots = k$$

In general, we can say that for some $x \in \{1, 2, 3, 4, 5, 6\}$

$$\frac{1 - P(x)}{P(x)} * x = k$$

Hence $P(x) = \frac{x}{x+k}$

We know that $\sum_{i=1}^6 P(i) = 1$

This means that

$$\sum_{i=1}^6 \frac{i}{i+k} = 1 \tag{3}$$

Using a graphing calculator, $k \approx 16.78$

(k is the expected value of runs in Nash equilibrium)

By using $P(x) = \frac{x}{x+k}$, the Nash equilibrium probability distribution for the bowler is (approximately) as follows:

x	1	2	3	4	5	6
P(x)	0.0562	0.1065	0.152	0.1925	0.23	0.263

6 Nash equilibrium distribution for the batsman

Let the probability distribution function of the batsman be P. Using claim Nash and applying (1). (Let $\sum_{i=1}^6 P(i) * i$ be y)

$$\frac{y - P(1) * 1}{P(1)} = \frac{y - P(2) * 2}{P(2)} = \dots = E$$

However, as this distribution is against the Nash equilibrium distribution for the bowler which gives the same expected value no matter what the distribution of the opponent (all numbers give expected value k), $E = k$.

$$\frac{y - P(1) * 1}{P(1)} = \frac{y - P(2) * 2}{P(2)} = \dots = k$$

In general, we can say that for some $x \in \{1, 2, 3, 4, 5, 6\}$

$$P(x) = \frac{y}{k+x} \quad (4)$$

We know that $\sum_{i=1}^6 P(i) = 1$

Therefore $\sum_{i=1}^6 \frac{y}{k+x} = 1$

Now for a moment, let us look at (3). If we add k and subtract k in the numerator, we get:

$$\sum_{i=1}^6 \left(1 - \frac{k}{i+k}\right) = 1$$

Manipulating this, we get

$$\sum_{i=1}^6 \frac{k}{i+k} = 5$$

Comparing this with $\sum_{i=1}^6 \frac{y}{k+x} = 1$, we get $y = \frac{k}{5}$.

Now using (4), we can get the (approximate) Nash equilibrium probability distribution of the batsman.

x	1	2	3	4	5	6
P(x)	0.1887	0.1787	0.1696	0.1615	0.1541	0.1473

7 The big flaws in my argument

We are done, right? Unfortunately no. I hate to break it to you but whatever I have done is wrong. First, we will try to understand what we have achieved. We have calculated 2 distributions, one for the batsman and one for the bowler. The distributions are designed such that no matter what the opponent's distribution is, the expected value of runs of the batsman will always be constant $= k$. However, this is not very useful as even against a non-optimal player we cannot get more than k runs in expectation(while bowling, we cannot reduce his expected value of runs below k in expectation). Hence, to gain an advantage, we should not use the Nash equilibrium distributions. Instead we should use (1) and (2) to gain an advantage over our opponent. But this too leads to problems as then our play is deterministic and a sneaky player might precalculate our moves and get an advantage.

However, this is not the main mistake I made. My goal in the previous 6 sections was to maximize the expected value of runs. **However, maximizing expected value of runs != maximizing probability of winning.** Why is this so? Let's assume that we have a distribution that gives us a 0.1 chance of getting 10000000 runs and a 0.9 chance of getting 0 runs. This gives us an extremely high expected value of runs but a very low(approximately 10%) chance of victory.

For now, we need to be satisfied with this. Later on, we will try to explore how to maximize the probability of winning.

Proof of the sus claim

Let us assume that there is a sequence of infinite numbers that gives the optimal expected value against a particular probability distribution(Basically the i th number of the sequence represents the number to be played on the i th turn). Let us say that the first number of the 'optimal' sequence is W . If the batsman is not out, then the player will proceed to the second number of the sequence. What should this number be? As the goal remains the same(to give the optimal value of the expected value of runs), the second number should also be W . This same logic can be applied to all the numbers of the infinite sequence.