Spring 2003

Hazard: The Scientist's Analysis of the Game.

Kaloian Petkov

Biographical Information for Kaloian Petkov:

During my senior year of high school, I took an honors course in literature that was probably the single most excruciating experience in my education up to that time. My passion is for science – mathematics, physics, computer science, etc. Rarely do I discuss Ibsen or analyze Stevie Smith's poems outside of an English class, yet the lessons I've learned in writing have become of paramount importance. When I write about science, mathematics in this particular case, I strive to balance my language. It has to be precise (to satisfy scientific requirements), but not to the point of boring my audience. Writing this paper required a great deal of effort and persistence, but it was all worthwhile. As a CS/Math major, I'm particularly pleased to win a writing competition!

HAZARD

The Scientist's Analysis of the Game

Kaloian Petkov

Mathematics of Games and Gambling
Prof. Edward Packel

Hazard
The Scientist's Analysis of the Game

Hazard was one of the most popular dice games of the seventeenth and eighteenth centuries. It created and destroyed fortunes in various gambling establishments and provided entertainment for the army as well (Bell, 1960). In its current iteration, the game is still among the most popular – a slightly simplified version of Hazard is now known as Craps. The casino version – Casino Craps – adds a number of side bets.

The rules of Craps varied depending on where the game was played. In France, for example, players bet against a bank, whereas in England the game was between the players. Since my analysis of the game is primarily over single-player scenarios, I will use the French version.

The game starts with the bank placing an initial bet on the table. The player, who is called a "caster", rolls two dice to determine his or her MAIN point (the sum of the two dice becomes the MAIN point). The caster rolls the dice until their sum falls between 5 and 9; no other values are allowed for the MAIN point. The second stage involves rolling for the CHANCE point. If the caster is so unfortunate as to roll a 2 or a 3, also known as CRABS, he loses the bet outright, no matter what the MAIN point is. He also loses if the CHANCE is 12 with a MAIN of 5, 7 or 9, and if the CHANCE is 11 with a MAIN of 5, 6, 8 or 9. This kind of loss is called an OUT. Since everything should be in balance, there is also the NICK (or immediate winning). If the caster matches his CHANCE point with the MAIN, he wins the stakes. Another way to get a NICK is to roll a 12 when the MAIN is a 6 or an 8, or to roll 11 for the CHANCE when the MAIN happens to be 7. I believe this particular NICK is the origin of the "Come seven, come eleven" saying. If the caster does not get an OUT or a NICK at this point, he keeps rolling until he duplicates his CHANCE, whe4n he wins, or until he gets the MAIN, when he loses all the money on the table.

The game implements a relatively easy betting system. Once the caster places her original bet, other players and/or the bank cover it, and the caster establishes his MAIN and CHANCE points. The floor is then open to any side bets at specified odds, either with the caster or against her. Again, the opponents must cover every bet. The caster can increase the stakes at any time before rolling the dice.

The rules of Hazard are complex, but Table 1 summarizes all the ways that the Caster can win or lose up to the point when he rolls for the CHANCE point.

Table 1: The rules for OUTs and NICKs

Main	Chance	Caster Gets
9, 7 or 5	12	OUT
9, 8, 6 or 5	11	OUT
	2 or 3	CRABS = OUT
6 or 8	12	NICK
7	11	NICK
any number	same as MAIN	NICK

Let's start by limiting the game to the first two rolls, or to the MAIN and CHANCE rolls. Assume that we are in France in the 18th century, where the Caster would play against a bank. Given that the analysis spans over the first two rolls only, I also assume that the rest of the game does not contribute to the expected value. Later analysis will cover the whole game. Now, instead of making counting errors in my head, I will use a table for the counting of how many ways a given roll can happen.

Table 2: All possible sums of two dice

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Table 3: the MAIN and CHANCE rolls

MAIN	CHANCE to win	CHANCE to lose	p(win)	p(lose)	p(go on)	Bet	expected value
5	5	2,3,11,12	4/36	6/36	26/36	X	-2/36*X
6	6,12	2,3,11	6/36	5/36	25/36	X	1/36*X
7	7,11	2,3,12	8/36	4/36	24/36	X	4/36*X
8	8,12	2,3,11	6/36	5/36	25/36	X	1/36*X
9	9	2,3,11,12	4/36	6/36	26/36	X	-2/36*X
overall			**5.83/36**	**5.08/36**	**25.09/36**	**X**	**0.75/36*X**

At first glance, the game might seem unfair, because there are always more ways to lose than to win, but this is not the case – there are more ways to roll a 7, for example, than to roll a 3. At this point I might be tempted to simply average the columns but that would be wrong since the five MAINs are not equally likely to appear. The last row contains the weighted averages of the probabilities and the expected value. Table 4 lists the "weights" or the probabilities of each MAIN.

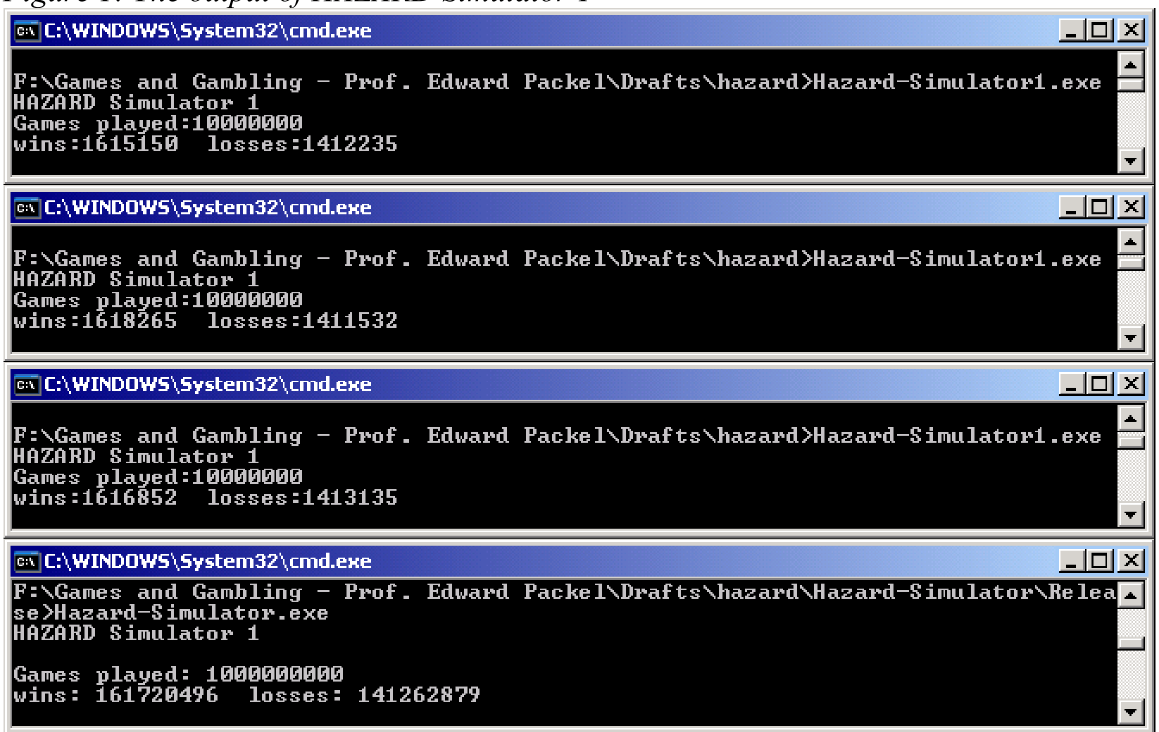
Table 4: The probability distribution for the MAIN roll

MAIN	weight	probability
5	4	4/24
6	5	5/24
7	6	6/24
8	5	5/24
9	4	4/24

There is a tiny but positive expected value for the game so far! If that was all there was to the game, the gaming establishments would hate it. This might be a very good place to play a few million games and see if the calculations really predict the chances of winning. Since playing by hand is rather slow, I will rely on my programming skills and make a very simple simulator that will play exactly the same game 10,000,000 times.

Calculating from Table 3 shows that the simulator should win 10,000,000* 5.83/36 = 1,619,444 times and lose 10,000,000*5.08/36 = 1,411,111 times. As Figure 1 shows, the output of HAZARD Simulator 1 supports the calculation.

Figure 1: The output of HAZARD Simulator 1



The above calculation may provide some practical insight on the game and its rules, but it is not very exciting. As Table 3 shows, in 25 of every 36 cases, the caster does not win or lose immediately, and if he does, the stakes are not too high. The casual player may hope for a 7 on the MAIN roll, but experienced players know that the big bets come after the CHANCE roll. However, big bets are not necessarily good if the probability of winning them is small. Given that gambling houses have always flourished, my initial feeling is that the game has an overall negative expected value.

For the calculation of the overall expected value of the game, I will assume that the player always bets a specific amount. Furthermore, he or she bets before every roll except the CHANCE and never runs out of money. I also assume that the game is played with fair dice.

Let's start by analyzing a single case. The player bets \$X before rolling for the MAIN and CHANCE, and then bets the same amount before each of his rolls if he does not win or lose right away. The MAIN is 5 and the CHANCE is 10. At this point, the player raises the stakes to 2X and starts rolling. My trusty friend Table 2 shows that there are 4 ways to get a 5 and lose because this is the MAIN point, 3 ways to get a 10, and 29 ways to continue the game. If the player does not win or lose, the same reasoning applies to the next roll, but this time the stakes are higher. And so it goes again and again. Eventually, the probability of continuing will become negligible and will not contribute to the expected value.

$$\begin{aligned} E(5,10) = \quad & 3/36\ (2X) + 4/36\ (-2X) + \\ & 29/36 * \{3/36\ (3X) + 4/36\ (-3X) + 29/36 * [3/36\ (4X) + 4/36\ (-4X) + \ldots\ ]\ \} \end{aligned}$$

I have a feeling this will be a nightmare to calculate in its current form. Maybe some algebra will help.

$$\begin{aligned} E(5,10) = \quad & [3/36\ (2X) + 4/36\ (-2X)] + 29/36 * [3/36\ (3X) + 4/36\ (-3X)] + \\ & (29/36)^2 * [3/36\ (4X) + 4/36\ (-4X)] + (29/36)^3 * [3/36\ (5X) + 4/36\ (-5X)] + \ldots \end{aligned}$$

This looks much better. The next step is to write the expression in summation notation.

$$E(5,10) = \sum_{n=0}^{\infty} (29/36)^n\,[3/36\ (2+n)X - 4/36\ (2+n)X]$$

Finally, I can type this expression in any Computer Algebra System, such as Mathematica™, Derive™, or even a TI-89 calculator.

∞

$$E(5,10) = \sum_{n=0} (29/36)^n [3/36\,(2+n)X - 4/36\,(2+n)X] = -43/49 *(X)$$

Table 5 shows the results of this calculation for all other cases. For each cell, I recalculated the probabilities of winning, losing and continuing, and then typed the new summation expression into my TI-89. The shaded boxes indicate a combination of a MAIN and a CHANCE that has led to an OUT or a NICK already.

Table 5: Expected values

		CHANCE										
		2	3	4	5	6	7	8	9	10	11	12
Main	5			-43/49 X		5/9 X	23/25 X	5/9 X	0	-43/49 X		
	6			-11/8 X	-5/9 X		47/121 X	0	-5/9 X	-11/8 X		
	7			-5/3 X	-23/25 X	-47/121 X		-47/121 X	-23/25 X	-5/3 X		
	8			-11/8 X	-5/9 X	0	47/121 X		-5/9 X	-11/8 X		
	9			-43/49 X	0	5/9 X	23/25 X	5/9 X		-43/49 X		

Table 6 contains the middle section of Table 5 (rows 5 through 9). It is clear that each pair of a row and column contains the same numbers, but one has the values with a negative sign. When I take the weighted average of this part of the data, everything will cancel and, therefore, the expected value is 0.

Table 6: Canceling expected values

		CHANCE				
		5	6	7	8	9
Main	5		5/9 X	23/25 X	5/9 X	0
	6	-5/9 X		47/121 X	0	-5/9 X
	7	-23/25 X	-47/121 X		-47/121 X	-23/25 X
	8	-5/9 X	0	47/121 X		-5/9 X
	9	0	5/9 X	23/25 X	5/9 X	

Table 7 contains the only portion of Table 5 that is relevant to the calculation.

Table 7: Expected values that did not cancel

		CHANCE	
		4	10
MAIN	5	-43/49 X	-43/49 X
	6	-11/8 X	-11/8 X
	7	-5/3 X	-5/3 X
	8	-11/8 X	-11/8 X
	9	-43/49 X	-43/49 X

Since the different combinations of MAIN and CHANCE points do not happen with the same probability, I cannot simply average the sum of all cells. Table 8 has the weight of each cell, which is the probability of having the cell's particular MAIN and CHANCE points. For example, there are 4 ways to get a MAIN of 5 and 3 ways to get a CHANCE of 4. The CHANCE cannot be 2, 3, 11, 12, the MAIN has to be between 5 and 9 and the CHANCE cannot equal the MAIN. For the first cell, the probability is 4/24 * 3/26 = 1/52. (4 ways to roll MAIN of 5; 24 possible MAINs; 3 possible ways to roll a CHANCE of 3; 26 possible CHANCEs, which excludes 2,3,11,12 and 5).

Table 8: Probability distribution

		CHANCE	
		probabilities	
		4	10
MAIN	5	1/52	1/52
	6	1/40	1/40
	7	1/32	1/32
	8	1/40	1/40
	9	1/52	1/52

To find the weighted average of Table 7, I will multiply each value with its corresponding weight in Table 8 and then sum everything. As a result,

Expected Value = -23,633/76,440 X.

I am not done yet. Since this is the expected value for the second part of the game only, I have to plug it into my original discussion of the game. I assumed that the expected value would be 0, and clearly -23,633/76,440 X ≠ 0.

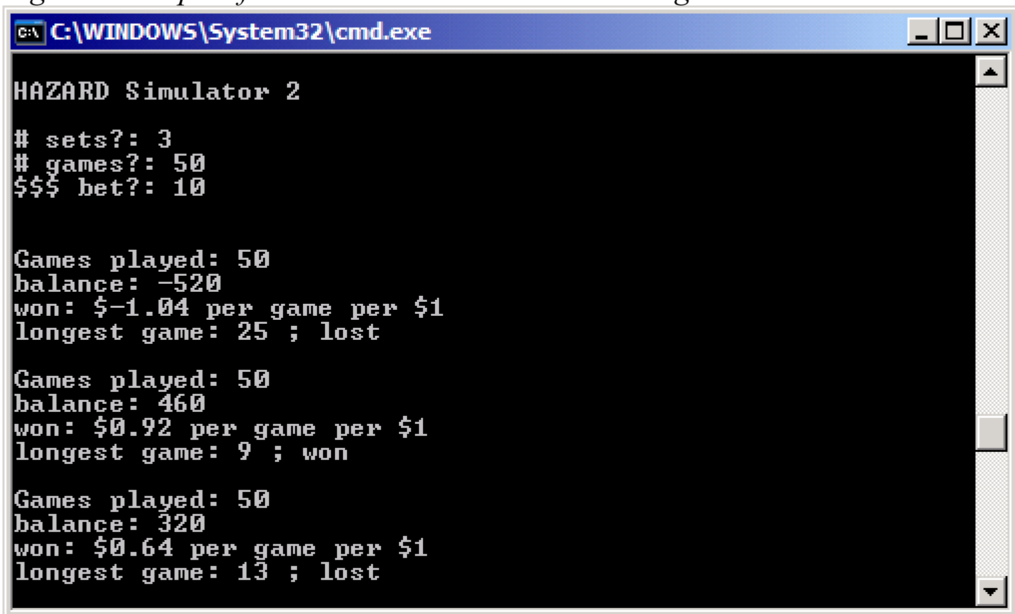
Total Expected Value = 5.83/36 [X] - 5.08/36 [X] – 25.09/36 [23,633/76,440 X] =

= **<u>-0.1946 X</u>**

In terms of betting, the result is that a $1 bet will produce a loss of 19.46¢, which translates to a house edge of 19.46%. I realize that the mathematics involved in the analysis became rather complicated after the introduction of the infinite series. Since I cannot possibly play a few million games to verify the calculation, I will again turn to technology for help.

This time the simulation is significantly more complex, since it keeps track of betting and the balance of the player. Out of curiosity, I added some code that keeps track of the lengthiest games. The output in Figure 4 contains three sets of fifty games each. The program plays the games and outputs the final balance of the player, the winnings per $1 bet, the roll number of the longest game, and whether the player won that game. The initial balance of the player is $0, and he bets $10 before the MAIN. After the CHANCE roll, he bets $10 each time before he rolls.

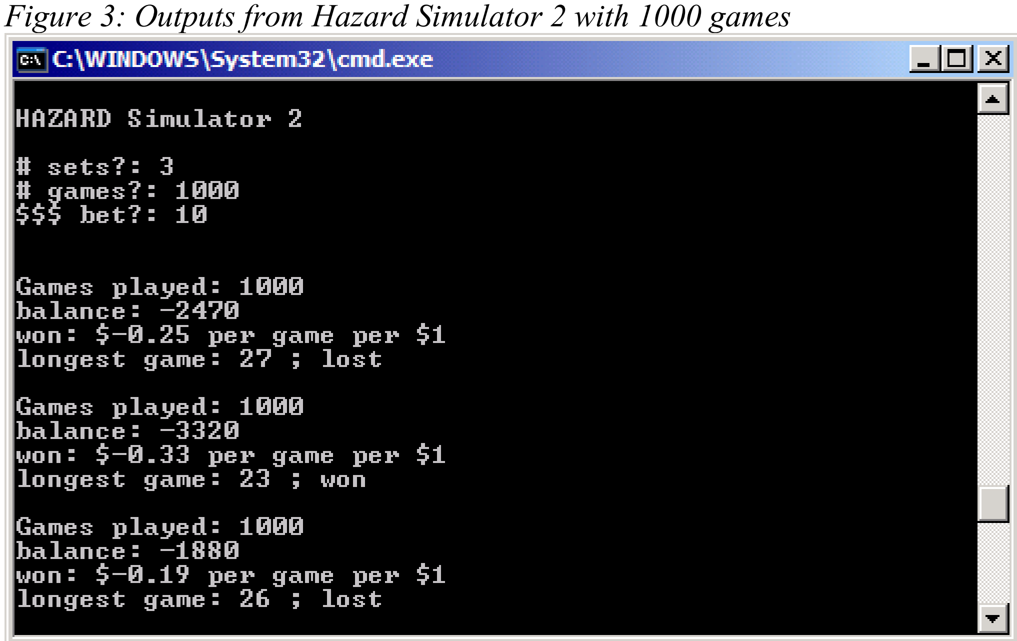
Figure 2: Outputs from Hazard Simulator 2 with 50 games



It seems possible to win a good amount of money over relatively few games. Notice that some games can become lengthy – in the first set, the longest game was 25 rolls. Usually, people place huge side bets along the player (or against him), and a game as long as that would have increased the stakes tremendously.

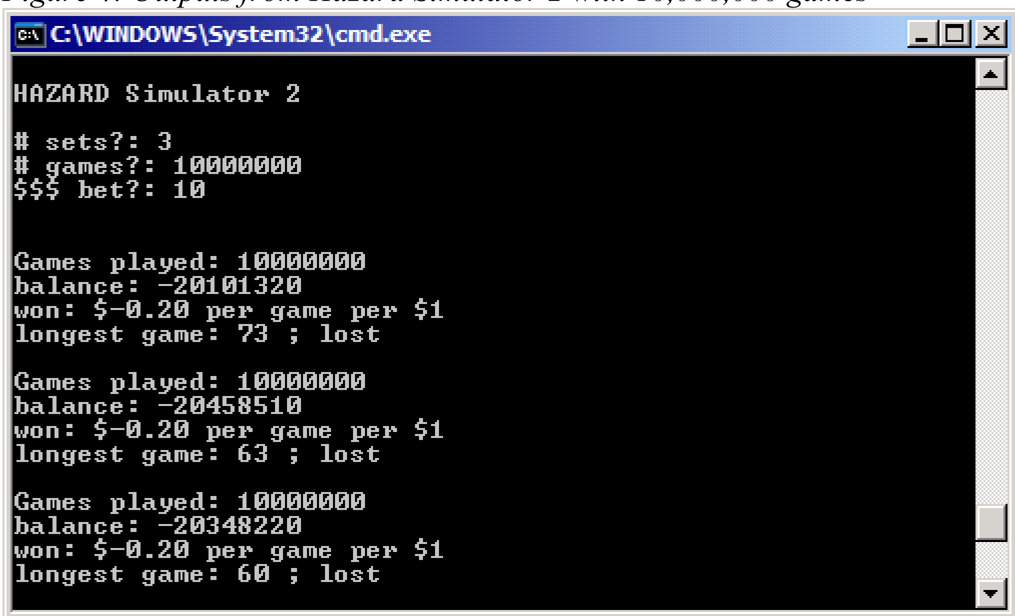
The next simulation is of three sets of a thousand games each.

Figure 3: Outputs from Hazard Simulator 2 with 1000 games



The situation starts looking rather grim for the bettor. As the number of games increases, the average winning of the game becomes negative. At this point the house edge is somewhere between 19% and 33%. Of course, this could be just a coincidence. To minimize the effects of chance, the next simulation plays three sets of 10 million games each.

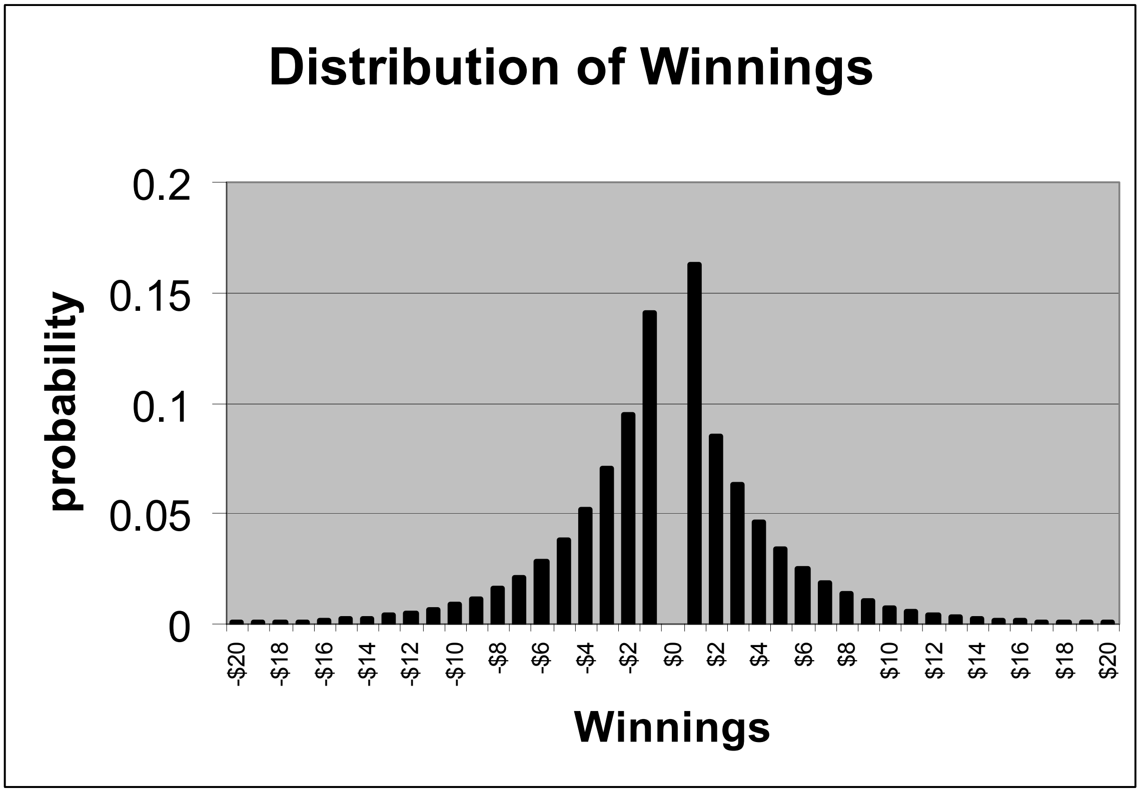
Figure 4: Outputs from Hazard Simulator 2 with 10,000,000 games



At this point I can safely assume that the house edge is approximately 20%, which agrees with the mathematical analysis.

My third and final computer simulation will play 100,000,000 games with all $1 bets and use the result to generate a distribution of the results. The output of "Hazard Simulator 3" is a text file with 1001 number, where each number represents the number of times every payoff occurred. The data covers payoffs between -$500 and $500, although I rarely saw a payoff above $80. Then the data goes into an Excel spreadsheet and finally into a bar graph. As you can see, the graph shrinks the domain even more; the rest of the probabilities are negligible and only ruin the nice peak in the middle.

Figure 5: Winnings distribution



In my initial analysis, I found that the game had a positive expected value over the MAIN and CHANCE rolls. As you can see in the graph, the player is winning the $1 from these rolls more often than losing it. However, the rest of the bars on the winning side are shorter than their counterparts on the left side, which is evidence for the negative expected value of the game.

For the most part, the results are not surprising – a game that led to the prosperity of gaming houses must have had a decent house edge. What I cannot explain is the slight edge that the player had during the first two rolls. It might have been a hook to attract bettors. Would it ever attract me enough that I play? Not a chance!

Bibliography

Bell, R. C. (1960). Board and Table Games from many Civilizations. Oxford University Press.