Yahtzee!

Christian Zhou-Zheng

June 9, 2023

1 Introduction

1.1 Abstract

Yahtzee is a dice game in which the goal is to score certain combinations of the five game dice. Its element of randomness, degree of player choice, and dynamic scoring system make it a prime candidate for probability analysis when considering its relative simplicity. We analyze a couple of penultimate-round scenarios in Yahtzee and determine the optimal strategy for each, considering the impact on the expected value of the final round.

1.2 Rules

Each round of a Yahtzee game contains one to three rolls and begins by rolling all five dice. The player may choose to reroll any selection of the five dice on the second roll, and again on the third. After three rolls, or after the player is satisfied with their dice, they must score the dice in one of thirteen categories, each of which is scored differently. No category may be scored twice; as such, the player may be forced to score a hand in a row for which that hand is invalid. In this case, the player is simply awarded 0 points. The game ends when all thirteen categories have been scored. The thirteen categories are as follows:

Table 1: Score Sheet

Category	**Points**	**Condition**
Ones	Sum of 1s	None
Twos	Sum of 2s	None
Threes	Sum of 3s	None
Fours	Sum of 4s	None
Fives	Sum of 4s	None
Sixes	Sum of 6s	None
Upper Bonus	35	Total of above rows ≥ 63
3 of a Kind	Sum of all dice	At least three dice are the same
4 of a Kind	Sum of all dice	At least four dice are the same
Full House	25	Sets of three and two matching dice
Small Straight	30	Four dice are in increasing order
Large Straight	40	Five dice are in increasing order
Yahtzee	50	All five dice are the same
Chance	Sum of all dice	None

2 Scenario 1: Fours vs. Full House

Suppose the score sheet looks as follows, and your dice are such that you cannot score anywhere on this round. You must choose between the 4s row and the Full House row to place a 0 in. Which should you choose?

Ones	3
Twos	4
Threes	12
Fours	
Fives	5
Sixes	0
Upper Total	24
Upper Bonus	0
3 of a Kind	17
4 of a Kind	23
Full House	
Small Straight	30
Large Straight	40
Yahtzee	0
Chance	14
Lower Total	124
TOTAL	148

To answer this, we look at two factors: the probability of scoring in each remaining row, and the expected value of each choice. The answer will predominantly depend on the latter, but the former must also be factored in to determine the risk of each.

We first calculate the fours row, and begin with the first roll of the final round. Curiously, the distribution of probabilities for rolling a given number of 4s follows a binomial distribution: specifically, the expansion of $(\frac{5}{6}+\frac{1}{6})^5$. This is because each die is independent of the others, and each has a $\frac{1}{6}$ chance of being a 4. Conceptually, this makes sense: there is $\binom{5}{0}=1$ way to have 0 dice be 4 (the other rolls are insignificant), $\binom{5}{1}=5$ ways to have 1 die be 4, and so on. This forms the sequence of coefficients, and then the increasing exponent on the $\frac{1}{6}$ term represents the increasing number of dice that are 4s. We then calculate the probabilities of having all numbers of 4s on the first roll:

Table 2: Fours Row First Roll Probabilities

# of 4s	Calculation	Probability
0	$\binom{5}{0}\left(\frac{5}{6}\right)^5\left(\frac{1}{6}\right)^0$	$\frac{3125}{7776}$
1	$\binom{5}{1}\left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right)^1$	$\frac{3125}{7776}$
2	$\binom{5}{2}\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)^2$	$\frac{1250}{7776}$
3	$\binom{5}{3}\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)^3$	$\frac{250}{7776}$
4	$\binom{5}{4}\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^4$	$\frac{25}{7776}$
5	$\binom{5}{5}\left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right)^5$	$\frac{1}{7776}$

Note that each has a denominator of 7776: this is 6^5, and the total number of possible rolls. Naively, we could stop here, but we must also factor in the remaining two possible rolls. Since we will never re-roll a die that is already 4, there are 4 succeeding scenarios to analyze: the cases of having 1, 2, 3, and 4 dice remaining. We just analyzed the case for 0 4s, and so the calculations are the same; at 5 4s, we will never re-roll, and the player will always score immediately. It is obvious that the optimal strategy is to keep re-rolling to get as many 4s as possible.

Conveniently, the four second-roll calculations needed also follow a similar binomial distribution! The probability of rolling a 4 is still $\frac{1}{6}$, but the number of dice remaining has changed, meaning the exponent changes accordingly. These are shown generally in Table 8 in Section 2.1; it applies to all dice values, not just 4.

With all the possible probabilities calculated for all three rolls, we can now simply draw out a diagram of all possible routes to score each number of 4s, calculate the probability along each route, and sum up these probabilities to determine the total probability of reaching each amount of 4s. There are too many paths to write them out or brute force it by hand, so a simple drawing is provided on paper and the results were computed by a Java program. It finds the total probabilities for every possible number of 4s, multiplies each probability by the number of points it scores (0, 4, ...) and sums this up to get the total expected value (**EV**). The results follow:

Table 3: Fours Row Expected Value

# of 4s	Score	Probability	EV
0	0	6.49%	0
1	4	23.63%	0.9452
2	8	34.40%	2.752
3	12	25.04%	3.0048
4	16	9.11%	1.4576
5	20	1.33%	0.266
TOTAL		100%	**8.4224**

As such, we see that the expected value of the fours row is **8.4224** points (roughly 2 dice), with a relatively low risk of scoring nothing, at only 6.49%. Next, we analyze the full house row.

The full house row is a fixed-value row, as opposed to the fours row, which is variable-value. As such, we can skip the last step of summing all possible scores, but the calculations are made slightly more complex in return. There are 300 total valid ways to roll a full house on the first try (citation: APCSA lab), meaning the chance of rolling one first try are $\frac{300}{7776}$. If we fail on the first roll, there are five possible succeeding scenarios: three of a kind (four of a kind counts in this category), two pairs, one pair, and no pairs. See the table below for the probabilities of each.

Table 4: Full House First Roll Probabilities

Scenario	**Probability**
Full House	$\frac{300}{7776}$
Three/Four of a Kind	$\frac{1350}{7776}$
Two Pairs	$\frac{1800}{7776}$
One Pair	$\frac{3600}{7776}$
No Pairs	$\frac{720}{7776}$

Two pairs are the simplest: we simply roll the last die until it matches one of the two pairs. We have a $\frac{1}{3}$ chance of doing that the second roll, and again on the third, making for a $\frac{5}{9}$ total chance from here.

With three/four of a kind, keep the three and re-roll the other two. There is a $\frac{1}{6}$ chance of getting a pair on each of the two remaining rolls, making for a $\frac{11}{36}$ total chance from here.

With one pair, re-roll the other three. This has several paths: instantly, you can roll 3 of a kind (chance of $\frac{1}{36}$) or 2 of a kind and 1 that matches the other (also $\frac{1}{36}$), for a $\frac{1}{18}$ chance of getting it on the second roll. Alternatively, you can roll another pair to get to two pairs, with a probability of $\frac{5}{12}$, or one or two matching the existing ones (3 matching is a valid full house), with a probability of $\frac{11}{36}$; the remainder is a $\frac{1}{3}$ chance of only having one pair again. We show these calculations in Table 5, along with the probabilities required for a full house on the third and final roll.

With no pairs, we're back to square one and we look at the first table to repeat the process, with slightly adjusted probabilities on each due to the removal of a roll. We show the final probabilities in Table 6 (the probabilities of getting a full house on the second and third roll combined).

Table 5: One Pair on First Roll to Full House

Scenario	**Probability**
3 of a Kind	$\frac{1}{36}$
2 of a Kind and 1 matching	$\frac{1}{36}$
2 Pairs	$\frac{5}{12} * \frac{1}{3} = \frac{5}{36}$
1 or 2 matching	$\frac{11}{36} * \frac{1}{6} = \frac{11}{216}$
1 Pair	$\frac{1}{3} * \frac{1}{18} = \frac{1}{54}$
TOTAL	$\frac{57}{216}$

Table 6: No Pairs on First Roll to Full House

Full House	$\frac{300}{7776}$
Three/Four of a Kind	$\frac{1350}{7776} * \frac{1}{3} = \frac{450}{7776}$
Two Pairs	$\frac{1800}{7776} * \frac{1}{6} = \frac{300}{7776}$
One Pair	$\frac{3600}{7776} * \frac{1}{18} = \frac{200}{7776}$
No Pairs	$\frac{720}{7776} * \frac{300}{7776} = \frac{1000}{279936}$
TOTAL	$\frac{118000}{279936}$

Table 7: Overall Full House Probabilities from First Roll

Scenario	**Probability**
Full House	$\frac{300}{7776}$
Three/Four of a Kind	$\frac{1350}{7776} * \frac{5}{9}$
Two Pairs	$\frac{1800}{7776} * \frac{11}{36}$
One Pair	$\frac{3600}{7776} * \frac{57}{216}$
No Pairs	$\frac{720}{7776} * \frac{118000}{279936}$

We can now add these up, as in Table 7, to get the final probability of getting a full house, approximately 36.6%. Since a full house is always worth 25 points, the expected value is therefore **9.15** points.

In comparison, the fours row has an expected value of **8.4224** points. This is an example of a scenario in which expected value is not the only factor that should be taken into account, as the difference in EV is only 0.7276 points. Factors such as risk, standard deviation, and the opponent's situation must also be considered.

Firstly, risk; the probability of scoring nothing going for a full house, at 63.4%, is much higher than the probability of scoring nothing going for fours, at merely 6.49%. The standard deviation also differs, as expected; attempting a full house, being all (25) or nothing (0), has a higher standard deviation than going for fours, which has more closely spaced score values at intervals of 4 from 0 to 20.

Finally, the situation of your opponent must be taken into account: while going for fours appears to be the safer option, if your opponent is greatly in the lead, you will need to take more risks to catch up. Therefore, it may be a better decision to take the full house, with its higher potential value despite its greater risk. Vice versa, if you are in the lead, it will usually pay off better to take the safe option of the fours row, as you have less to gain and more to lose by taking the risk of a full house.

In conclusion, the decision of whether to go for fours or a full house depends on the situation, and cannot be determined by expected value alone; however, the risk vs. reward appears to generally favor fours, despite its lower EV.

2.1 Upper Section Roll Probabilities

Table 8: Probabilities of Rolling n for Amounts of Remaining Dice

Dice Remaining	**# of ns post-roll**	**Calculation**	**Probability**
5	0	$\binom{5}{0}\left(\frac{5}{6}\right)^5\left(\frac{1}{6}\right)^0$	$\frac{3125}{7776}$
5	1	$\binom{5}{1}\left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right)^1$	$\frac{3125}{7776}$
5	2	$\binom{5}{2}\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)^2$	$\frac{1250}{7776}$
5	3	$\binom{5}{3}\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)^3$	$\frac{250}{7776}$
5	4	$\binom{5}{4}\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^4$	$\frac{25}{7776}$
5	5	$\binom{5}{5}\left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right)^5$	$\frac{1}{7776}$
4	0	$\binom{4}{0}\left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right)^0$	$\frac{625}{1296}$
4	1	$\binom{4}{1}\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)^1$	$\frac{500}{1296}$
4	2	$\binom{4}{2}\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)^2$	$\frac{150}{1296}$
4	3	$\binom{4}{3}\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^3$	$\frac{20}{1296}$
4	4	$\binom{4}{4}\left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right)^4$	$\frac{1}{1296}$
3	0	$\binom{3}{0}\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)^0$	$\frac{125}{216}$
3	1	$\binom{3}{1}\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)^1$	$\frac{75}{216}$
3	2	$\binom{3}{2}\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^2$	$\frac{15}{216}$
3	3	$\binom{3}{3}\left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right)^3$	$\frac{1}{216}$
2	0	$\binom{2}{0}\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)^0$	$\frac{25}{36}$
2	1	$\binom{2}{1}\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^1$	$\frac{10}{36}$
2	2	$\binom{2}{2}\left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right)^2$	$\frac{1}{36}$
1	0	$\binom{1}{0}\left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right)^0$	$\frac{5}{6}$
1	1	$\binom{1}{1}\left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right)^1$	$\frac{1}{6}$

3 Scenario 2: Threes vs. Full House

Suppose the score sheet looks as follows, finding ourselves again with an open upper and lower section row, and your dice are such that you cannot score anywhere on this round. You must now choose between the *3s* row and the Full House row to place a 0 in. Which should you choose?

Ones	2
Twos	2
Threes	
Fours	8
Fives	15
Sixes	24
Upper Total	51
Upper Bonus	0
3 of a Kind	17
4 of a Kind	23
Full House	
Small Straight	30
Large Straight	40
Yahtzee	0
Chance	14
Lower Total	124
TOTAL	175

Naively, the Full House row seems like the much better option after seeing the last example - it beat out fours, so surely it beats threes as well. However, in this scenario, notice the upper total - it is 12 points, 4 3s, away from the threshold of 63 points necessary to score the upper bonus, worth **35** points. This requires a roll of 4 or more 3s in one hand, but secures a bonus greater than the value of the full house itself; this could potentially change the optimal strategy!

There is no need to re-analyze the full house probabilities and expected value; we may reuse the EV of **9.15** points from the previous scenario. We also need not recalculate the probabilities for threes, as they follow the same distribution as the fours analyzed previously. However, we must recalculate expected value from the last table in the previous section:

Table 9: Threes Row Expected Value

# of 3s	Score	Probability	EV
0	0	6.49%	0
1	3	23.63%	0.7089
2	6	34.40%	2.064
3	9	25.04%	2.2536
4	12 + 35	9.11%	4.2817
5	15 + 35	1.33%	0.665
TOTAL		100%	**9.9732**

It appears that the EV of the threes row is **9.9732** points, after factoring the upper section bonus, which is greater than the EV of the full house row - again, only **9.15** points! In this case, it is clear that the optimal strategy is to place a 0 in the full house row and attempt to score the 3s row in the final round - the discussion points in Scenario 1 are now moot, as all point to threes being the better option in this case. This scenario goes to show how the optimal strategy can change depending on the state of the game, as well as how Yahtzee's dynamic scoring system can affect that strategy.