Sophia Oku

Solo Project

Project Group 130

Optimizing the Game of Yahtzee

ISYE 6644 Simulation and Modeling for Engineering and Science

1. Introduction

Yahtzee is a dice game consisting of strategy and luck. According to the Wikipedia article on Yahtzee, the origin of the game is widely attributed to two unknown Canadian couples who traded the rights for the game for 1000 Yahtzee gift sets to Edwin Lowe (Wikipedia contributors, 2025). The game of Yahtzee is believed to be played by 100,000,000 people each year (Cremers, 2002), with an estimated 50,000,000 games sold each year (Wikipedia contributors, 2025). The appeal of the Yahtzee game lies in its simplicity and ease of play compared to other popular board games like chess; the simplicity of Yahtzee remains unmatched. Yahtzee is even simpler than other probability-based games such as poker, blackjack, or even board games like Risk that involve decision-making with hidden information or complex player interaction. Yahtzee is usually played in a group of 2 to 10 players with five 6-sided dice and a scorecard. Each player gets three rolls in a round to throw the dice, with the primary objective of maximizing their score from throwing different combinations in a total of thirteen rounds per player. The reliance on chance and strategy makes Yahtzee a suitable game for probability modeling and simulation, as done by various studies (Pawlewicz, 2011; Sturtevant, 2003; Wiksten, 2015). Over the years, Yahtzee has evolved from a board game based on chance to a game where art and strategy implementation, using various simulation systems to optimize winnings, are used.

The use of simulation to improve winnings and strategy has been employed in multiple probabilistic forecasting contexts such as business strategy, game theory, and war strategy, with more recent years focusing on layering artificial intelligence in simulation programs (Gruetzemacher et al., 2025). Simulation helps to evaluate strategies by quantifying randomness and revealing patterns that cannot easily be analyzed from single analytical formulas. In games like backgammon (Papahristou & Refanidis, 2014), poker (Brandon, 2018), Scrabble (Zook et al., 2019), or Settlers of Catan (Szita et al., 2010), Monte Carlo (Metropolis & Ulam, 1949) simulations have been used to improve gameplay strategies. Similar to the previously mentioned board games, simulation has been explored in multiple research studies to optimize Yahtzee strategies and improve scoring advantage. In most research, this is done by exploring different strategies and rare combinations and evaluating their scores in a simulated gameplay. These simulated interactions then provide an output that is applied in real human-played board games, providing regular game players and researchers with an in-depth understanding of how different strategies affect their winning probabilities.

This study focuses on understanding how various gameplay strategies affect the chances of winning. The main objective is to understand how often specific outcomes occur and to identify which strategies consistently yield the highest average scores. The project then explores the trade-off between risk and reward by examining the winning potential of focusing on the upper section versus the lower section in a Yahtzee game. By simulating 100,000 game plays using a Monte Carlo framework, the model aims to understand the risk and reward associated with focusing on different sections during a Yahtzee game play.

2. The Game: Yahtzee

2.1 Background

Yahtzee is a dice-rolling game where players take turns rolling five dice with the aim of accumulating the highest score based on the Yahtzee scorecard and rules. Each player is allowed to roll the dice up to three times during a turn. During a turn, a player can decide to hold onto dice that are considered of high value and re-roll the remaining dice for a chance at a better outcome. Players are each assigned a scorecard which is divided into upper and lower sections with a total of thirteen categories, and over the course of thirteen rounds, must fill one entry per turn. If a player cannot roll the combination of an unused category on their scorecard, they must enter zero in one of the unused categories. At the end of the game, all scores and bonuses are tallied, and the player with the highest score wins.

2.2 The Sections

A Yahtzee scorecard is divided into two sections: the Upper Section and the Lower Section (Table 1). Each section is subdivided into categories with different points allocated to each category. The number of points awarded usually depends on the probability of achieving that outcome. For example, a player gets higher points from scoring a Yahtzee compared to Aces. The limitation of a player being able to use a category only once makes strategy an important factor in the game of Yahtzee. The following subsections explain each Yahtzee section.

2.2.1 Upper Section Categories

The Upper Section comprises the first six categories in the score sheet, which are simple combinations such as Aces, Twos, Threes, Fours, Fives, and Sixes. The points in the Upper Section are usually the total sum of the dice that matches the value of the category selected. For example, if a player rolls 1, 1, 1, 5, 4 and chooses "Aces," they score 3. The Upper Section also has a bonus category. If the total number of points in the Upper Section is 63 or more, the player receives 35 bonus points.

2.2.2 Lower Section Categories

The Lower Section contains the last seven categories: Three-of-a-Kind, Four-of-a-Kind, Full House, Small Straight, Large Straight, Chance, and Yahtzee. For Three-of-a-Kind and Four-of-a-Kind, the scoring is straightforward: add up the total of all dice that satisfy the category. For example, if a player rolls 2, 3, 3, 3, 5 and chooses "Three-of-a-Kind," the player gets a score of 16 for that category (2 + 3 + 3 + 3 + 5). What makes this category a difficult decision is that a player can choose to use the Upper Section or the Lower Section for the Three or Four-of-a-Kind. The other sections of the Lower Section are much harder to achieve, hence higher scoring. To score a Full House, a player must get a combination of three of one number and two of another, e.g., (3, 3, 3, 2, 2), to get a score of 25 points. For a Large Straight, a player must get five sequential dice with a total score of 40 points achieved in that category. A Yahtzee is when a player rolls all five dice showing the same number (e.g., five 3s) and gets a total score of 50 points. Chance is usually a placeholder for when a roll does not fit any of the other categories and is often seen as a fallback option.

2.2.3 Bonuses

Bonuses are awarded in the Upper and Lower Sections if certain criteria are fulfilled. A player is awarded bonus points for the Upper Section if their total score in that section reaches 63 or more, which is typically achievable by rolling at least Three-of-a-Kind in most of the number categories. 100 bonus points are awarded in the Lower Section for rolling a subsequent Yahtzee after the first Yahtzee was scored, regardless of whether that initial Yahtzee category score was 50 points or zero. In addition to the bonus, the player also gets to fill an open eligible category. For example, if the player rolls five sixes on their second Yahtzee, they are assigned 100 bonus points and must choose to fill the Sixes (Upper Section), Three-of-a-Kind (Lower Section), Four-of-a-Kind (Lower Section), or Chance. In the event that the player has already used all four of the mentioned categories, the player is subjected to the Joker Rule. The Joker rule allows a Yahtzee to be used as a wildcard in the Lower Section (Full House, Small Straight, or Large Straight categories) only if the Upper Section for the number in the Yahtzee roll is already filled with a score of zero. For example, if a player had already scored a Full House in a previous roll and later in the game rolls five 3's (a second Yahtzee), and the "Threes" category in the Upper Section is already filled with a zero, the player can use the five 3's as a Joker to score points in a Small Straight or a Large Straight category if available, even though five of a kind isn't normally considered a straight.

Table 1. Yahtzee scoring categories and description.

Category	Description
Upper Section	
Aces	Count and score only the dice with the number 1
Twos	Count and score only the dice with the number 2
Threes	Count and score only the dice with the number 3
Fours	Count and score only the dice with the number 4
Fives	Count and score only the dice with the number 5
Sixes	Count and score only the dice with the number 6
Lower Section	
Three-of-a-Kind	At least three dice must show the same number, the score is the sum of all dice.
Four-of-a-Kind	At least three dice must show the same number, the assigned score is the sum of all dice.
Full House	Three of one number and two of another number (e.g., 3, 3, 3, 3 2, 3), the assigned score is 25 points.
Small Straight	A sequence of four numbers (e.g., 1, 2, 3, 4), the assigned score is 30 points.
Large Straight	A sequence of five numbers (e.g., 1, 2, 3, 4, 5), the assigned score is 40 points.
Yahtzee	Five of a kind (e.g., 6, 6, 6, 6, 6, 6), the assigned score is 50 points.
Chance	Any combination of dice, the assigned score is the sum of all dice.

The rules defined above show the importance of strategically applying rolls to the right categories to give a winning edge. Should a player prioritize the Upper Section over the Lower Section? Should a player aim at rolling only Yahtzee's? Choosing the right dice to keep, the section to favor, or the category to sacrifice is a critical step in winning a Yahtzee game. This strategic thinking and these choices make Yahtzee more than just a game of luck; it's also a game of strategy and risk management.

3. Literature Review

There are two approaches used in optimizing a Yahtzee game model: theoretical and simulation based. Theoretical approaches in most studies are usually limited by processing ability; hence, they focus more on probability structure or scoring theory. In Cremers (2002), a Markov Decision Process was used to model each game state as a combination of dice values, rerolls left, and unfilled score categories. This allowed the algorithm to evaluate actions like rerolling or selecting a scoring category while accounting for both chance and player decisions. Another study using a theoretical approach, Maynard et al. (2001), prioritized the upper section, leading to a high outcome. The decision tree model from Maynard et al. (2001) prioritized aiming for Sixes, Fives, and Fours as high-value targets with an aim of avoiding zeroing these categories. The average score from this model was 226, lower than the 250 point average obtained from Cremers (2002).

Simulation-based approaches have the upper hand due to their ability to handle complex scenarios and manage the effect of randomness. In Pawlewicz (2011), a simulation model with traits similar to a Monte Carlo model was used to model expected scores per move. The study used a directed acyclic graph to model different states and moves. The states were split into random states and player states, and edges between each state represented the moves, with the edges from the random state having probabilities attached to them and the player state edges having no probabilities attached. Aside from the model's requirement of high processing power, it overestimated the expected scores (estimated EV of 264 when the true max EV is ~254). Felldin and Sood (2012), on the other hand implicitly applied monte carlo simulation through binomial probability equations to calculate expected values for each category per turn. While this simpler model was faster, it failed to consider long-term impacts, often missing the upper section bonus due to undervalued early choices. The binomial model achieved an average score of ~ 222, as opposed to the estimated expected value of 264 from Pawlewicz (2011).

Looking at both the theoretical and simulation-based models, one can deduce that the average performance of a simulated Yahtzee game should be between 221.68 and 264 (Cremers, 2002; Felldin & Sood, 2012; Maynard et al., 2001; Pawlewicz, 2011). These average scores provide the benchmark to which the simple simulation model used in this study will be evaluated while trying to address some of the limitations of the previously discussed models.

Table 2. Average expected score across various literature studies.

Literature Study	Average Expected Score
Cremers (2002)	250*
Felldin and Sood (2012)	222
Maynard et al. (2001)	226
Pawlewicz (2011)	264

* This study did not explicitly state the average expected score; hence the expected score was estimated from the charts provided.

4. Methods

The study uses python programming language specifically collections, intertools and tabulate python libraries to simulate a Yahtzee game based on Monte Carlo probability methods. The model tracked scores and bonuses to measure the performance prioritizing the upper section versus the lower section in a game play. The methods are detailed in the following paragraphs.

4.1 Monte Carlo Simulation

Monte Carlo Simulation is a method of estimating the outcome of a population by repeating an experiment *n* times and averaging the total outcome. The Monte Carlo model applies the law of large numbers, which states that the average of many independent, identically distributed (i.i.d.) random variables will converge to the expected value as the number of samples goes to infinity. So, to estimate the expected value of a function $\mathbb{E}\, f(x)$, random numbers are run through the function, and then the results are averaged to arrive at an estimate of μ (Equation 1). Where $\hat{\mu}$ is the sample mean, an estimate of the true mean μ, X_i are the points, and *n* is the number of samples taken.

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

This makes Monte Carlo an ideal method for estimating the strategy outcomes of a simple Yahtzee game compared to other methods of determining strategy outcomes, such as a Markov Decision Process. However, the model comes with its own limitations, such as slow convergence and the inability to learn from previous games. The Markov Decision Process addresses these limitations by modeling multiple states using Bellman equations and matrices but trades this advanced technique for higher computational processing, a key limiting factor in this study. Regardless of these limitations, the simplicity and scalability of the Monte Carlo model make it very practical for this study.

4.2 Statistics and Random Number Generation.

The method of evaluation was based on basic summary statistics generated by the simulation model. These statistics included the average score, best and worst scores, and average performance across scoring categories. The random number generator used was the built-in generator from Python's random module. This module uses the Mersenne Twister algorithm as its core generator, which produces 53-bit precision floating-point numbers and has a period of $2^{19937} - 1$. This pseudo random number generator ensures that every number within the specified range has an equal chance of being selected, minimizing the appearance of patterns. Although the Mersenne Twister is deterministic in nature, it is considered suitable for the purposes of this simulation study due to its proven extended cycle randomness and long period.

4.3 The Model - A Strategy Simulation

The Yahtzee simulation model used in this study was designed to evaluate the effectiveness of different strategies over full game plays. The model applies the Monte Carlo method by simulating individual dice rolls, rerolls, and scoring decisions based on predefined strategies. For each game, the simulation rolled five dice, identified the best dice to hold based on the current strategy, rerolled the remaining dice up to two times, and then selected the category that would maximize the score for that turn. The model follows all standard Yahtzee scoring rules, with the exception of the Joker Rule, which was omitted due to its complexity and lack of valuable input to the objectives investigated.

Two primary strategies were evaluated: one that prioritized the upper section to maximize the 35-point bonus, and another that prioritized the lower section. Unlike other simulation models such as decision trees, the Monte Carlo approach is memoryless, which increases the risk of wasting high rolls on low-scoring categories or duplicating scoring entries. To address this, each roll was logged into a scorecard to ensure proper category tracking. While the model does not incorporate formal gameplay optimization, which would require advanced understanding of professional-level strategy, it successfully meets the objectives outlined for this study.

5. Results and Discussion

5.1 Upper Section Strategy

The first simulation strategy prioritized the upper section, aiming to capitalize on the 35-point bonus awarded for scoring at least 63 points across the Ones through Sixes categories. The average score achieved using this approach was 176, with a maximum score of 333 and a minimum score of 60. While this score is lower than the average scores reported in other simulation studies (Table 2), this model performed relatively well as the literature models were designed to optimize overall performance across the full game rather than test a single section focus. The average upper section bonus recorded per game was 0.95, suggesting that the success rate for achieving the bonus remains relatively low. It also shows that players that prioritize filling out Ones through Sixes face a high chance of low scores. This statement is rehashed in the category score distribution (Table 4).

The lower section category dominated the score sheet even when prioritizing the upper section during the simulation. The highest performing categories were Small Straight (28.64), Three-of-a-Kind (19.9), and Full house (19.26). In contrast, low value upper section categories like Ones (1.92), Twos (4.47), and Threes (6.74) contributed the least to the overall score.

In summary, these results suggest that focusing on high-risk high-value categories provided a higher winning chance than targeting easily low-risk low-value categories. Even when the primary objective was to prioritize the upper section, the highest contributions to the total average score still came from the lower section.

Table 3. Summary of Total Game Score Statistics Prioritizing Upper Section Categories

Average Score	Best Score	Worst
175.99	333	60

Table 4. Average Score per Category in 100,000 Simulated Games

Category	Average Score
Chance	18.3
Fives	10.89
Four-of-a-Kind	13.12
Fours	9.03
Full House	19.26
Large Straight	16.68
Yahtzee	13.24
Ones	1.92
Sixes	12.85
Small Straight	28.64
Three-of-a-Kind	19.9
Threes	6.74
Twos	4.47
Upper Bonus	0.95

5.2 Lower Section Strategy

The lower section strategy simulation focused on maximizing scores from lower section categories and achieved an average score of 205.46, with a maximum score of 431 and a minimum score of 62. This is higher than the average score from the upper section prioritization model (175.99). However, when compared to the average scores reported in other models (Table 2), the lower section strategy performed similarly, though still slightly lower. It is important to note that prioritizing the lower section bonus increased the average upper bonus from 0.95 to 1.24, which suggests that shifting priority from the upper section to the lower section did not prevent players from getting the bonus scores attached to the upper section. High-value categories like Small Straight, Full House, and Three-of-a-Kind were the three highest scoring strategies. While the lower performing categories like Ones and Twos showed improved performance in this model compared to the Upper Section optimization model.

In summary, targeting lower section categories during gameplay results in a higher overall average score than prioritizing the upper section.

Table 5. Summary of Total Game Score Statistics Prioritizing Lower Section Categories

Average Score	Best Score	Worst
205.46	431	62

Table 6. Average Score per Category in 100,000 Simulated Games

Category	Average Score
Chance	18.12
Fives	10.95
Four-of-a-Kind	16.39
Fours	9.12
Full House	20.21
Large Straight	16.57

Yahtzee	12.66
Ones	2.2
Sixes	12.84
Small Straight	29.05
Three-of-a-Kind	19.37
Threes	6.81
Twos	4.59
Upper Bonus	1.24

5.3 Average Score

The Lower Section optimization model average score of 205.46 outperformed the Upper Section optimization model by approximately 30 points, further proving that aiming for risky high scoring categories give a higher score on average than low risk categories. However, when compared to strategies from previous literature (Table 2) both strategies performed lower. The low performance in the model used in this study was expected as previous studies used more advanced optimization techniques such as Markov Decision Processes (Cremers, 2002), decision trees from expert behavior (Maynard, 2001), or recursive models (Pawlewicz, 2011). Regardless, the results are not too far off from the results in Table 2 and can be improved upon by including the Joker rule and a turn-by-turn memory function.

6. Conclusion

The model used in the study was able to successfully establish that prioritizing lower section categories or strategies increases the average score of a player’s performance as opposed to targeting lower section categories and bonuses. The model underperformed when compared to other models explored in the literature review due to its memoryless function and lack of professional knowledge applied during the model build. However, by integrating the Joker rule and a turn-by-turn memory function, there is a high chance of further improving the model's performance. Despite the shortcomings of the model, it was able to identify high-performing categories and strategies. Future work could explore decision trees or reinforcement learning approaches to further improve the model's performance.

References

Cremers, C. J. F. (2002). How best to beat high scores in Yahtzee: a caching structure for evaluating large recurrent functions. [MA thesis, Technische Universiteit Eindhoven]. In *Master's thesis*. https://www-set.win.tue.nl/~wstomv/misc/yahtzee/beat-high-score.pdf

Felldin, M., & Sood, V. (2012). *Optimal Yahtzee – Strategies and Heuristics.* [Masters Thesis, KTH Royal Institute of Technology]. https://www.csc.kth.se/utbildning/kth/kurser/DD143X/dkand12/Group5Mikael/final/Markus_Felldin_and_Vinit_Sood.pdf

Gruetzemacher, R., Avin, S., Fox, J., & Saeri, A. K. (2025). Strategic Insights from Simulation Gaming of AI Race Dynamics. *Futures*, 103563. https://doi.org/10.1016/j.futures.2025.103563

Maynard, K., Moss, P., Whitehead, M., Narayanan, S., Garay, M., Brannon, N., Kantamneni, R. G., & Kustra, T. (2001). Modeling Expert Problem solving in a Game of chance: A Yahtzee case study. *Expert Systems*, *18*(2), 88–98. https://doi.org/10.1111/1468-0394.00160

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, *44*(247), 335–341. https://doi.org/10.1080/01621459.1949.10483310

Papahristou, N., & Refanidis, I. (2014). Opening statistics and match play for Backgammon games. In *Lecture notes in computer science* (pp. 569–582). https://doi.org/10.1007/978-3-319-07064-3_49

Pawlewicz, J. (2011). Nearly optimal computer play in multi-player Yahtzee. In *Lecture notes in computer science* (pp. 250–262). https://doi.org/10.1007/978-3-642-17928-0_23

Sturtevant, N. (2003). A comparison of algorithms for multi-player games. In *Lecture notes in computer science* (pp. 108–122). https://doi.org/10.1007/978-3-540-40031-8_8

Szita, I., Chaslot, G., & Spronck, P. (2010). Monte-Carlo tree search in Settlers of Catan. In *Lecture notes in computer science* (pp. 21–32). https://doi.org/10.1007/978-3-642-12993-3_3

Wikipedia contributors. (2025, March 22). *Yahtzee*. Wikipedia. https://en.wikipedia.org/wiki/Yahtzee

Wiksten, L. (2015). *Optimal Yahtzee : A comparison between different algorithms for playing Yahtzee*. DIVA. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A810580&dswid=8045

Zook, A., Harrison, B., & Riedl, M. O. (2019). Monte-Carlo Tree search for simulation-based strategy analysis. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1908.01423