# Investigating the Winning Strategies for Mexico

Abdulmajeed Alharbi, Jake Denton, Bethany Johnson, Joshua Rustage, Gil Segev, George Southgate

2nd December 2020

## Abstract

This report aims to devise the best strategy for winning one version of the dice game Mexico; the big question is, when should you take another roll, and when should you not? The overall conclusions and main findings of the game analysis are that, whilst there is a percentage of sheer luck involved (as with all dice games), there are strategies that can be played by specific players in order to increase the likelihood of a win. With the use of theoretical concepts and simulation, optimal thresholds for scores in rounds of 2-9 players are found, which reduce the probability of elimination substantially (by almost 10% for player 1 in the two-player game). The ideas behind the derivation of the thresholds can be stretched to cover the strategies for other players in the game, and beyond this to a three-dice, two-player extension.

# Contents

# Introduction

Mexico is an elimination style dice game, played with two dice. In the game, the first player rolls the dice, and their two separate values are concatenated, with the highest throw coming first followed by the other roll. For example, a roll of 3 and then a 5 makes 53 (we always assume players maximise their score). The first player to roll may choose to roll the dice up to 3 times in total, in hope of achieving a high score (their final score is based on their most recent roll). Each subsequent player can then roll the dice, but not taking more rolls than the first player did. After each round, the player who achieves the lowest score is eliminated, and the player with the highest score is the first to roll in the next round. The winner of this game is the final player remaining.

In a traditional game of Mexico, non-palindromic scores (where the results on each die are distinct) are the lowest set of scores, followed then by palindromic scores, and finally the maximum score is 21 (referred to as achieving 'Mexico'). In this way, the minimum score one can obtain is 31, 11 is ranked one score higher than 65 and similarly, 66 is higher than 55, with 21 beating the lot. The game addressed in this report is a derivative of the traditional game, in which the 'largest score wins', so that 11 becomes the smallest score, and 66 is the new maximum.

A few important rules regarding the game need to be clarified to ensure each player has an equivalent chance of winning from the offset. At the beginning of each game, the allocation of player 1 is decided fairly, for example by drawing from a hat. Each round proceeds clockwise around the 'table' of players. If there is a tie for last place in any round (with more than two players), all those with the minimum score are eliminated. If there is a tie in a two-player game (or in the final round of a $n > 2$ player game), the round is instantly replayed.

This report attempts to formulate the best strategy to minimise the chance of elimination in each round of the game using various mathematical techniques. The methods used begin with calculation of summary statistics then a strategy for player 1 arising from the median is investigated. This is compared with an optimised strategy, found using simulations, for the two-player game. These ideas are then extended to higher player games, the theory this time coming from a binomial model. Again, the strategy derived is pitted against the optimised simulated approach. Additionally, procedures for other players in each round are investigated before extensions of our game to three dice are considered. The report closes with a conclusion of the findings along with some suggestions for further work and how the techniques used could be utilised in researching similar logic-based games.

# The Investigation

Consider a game of Mexico (having our adjusted set of rules) with $n$ players. A few assumptions must be made before the game can be analysed. Firstly, assume that players always maximise their score, that is they are aware of how to form the largest two-digit number possible with two numbers, and always set their score equal to this maximum. Secondly, the dice used in the game are fair, where the probability of rolling any number from one to six is $\frac{1}{6}$. Throws of the dice are assumed to be independent, and from this every roll and the score obtained from rolls are assumed to be independent. On top of this, an additional assumption is made about the decisions of general players which can be adapted into code: players that haven't been assigned a specific strategy will always roll until they either beat the lowest score achieved by the players that came before them (and so ensure their safety), or until they run out of rolls, in which case they have the new lowest score.

If these assumptions are fulfilled, we can consider the distribution of scores. There are 36 possible results for any roll of two dice, but there are only 21 possible scores in the game. These are tabulated below:

| Possible Values | Ways of achieving |
|---|---|
| 21 | Two ways |
| 31, 32 | Two ways |
| 41, 42, 43 | Two ways |
| 51, 52, 53, 54 | Two ways |
| 61, 62, 63, 64, 65 | Two ways |
| 11, 22, 33, 44, 55, 66 | One way |

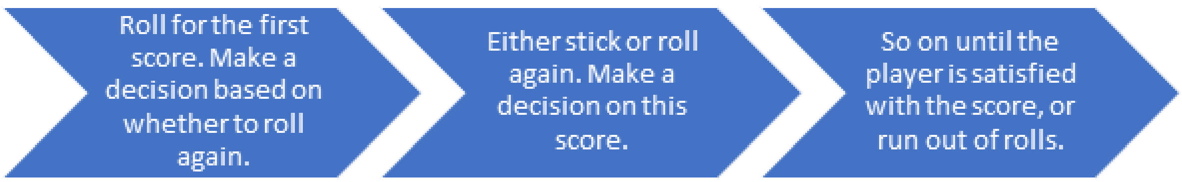
TABLE 1: WAYS OF ACHIEVING POSSIBLE SCORES IN THE GAME

From Table 1 above, it is clear that the distribution of scores has a positive skew, due to there being more scores in the 60's (6 potential values) than in the 10's (1 potential value) and 20's (2 potential values).

Let A denote the set of palindromic values, and B the non-palindromic values. The random variable $X$ can be defined, where $X$ is the score that can be obtained in the game. $X$ has probability mass function (using table 1 and assumption that dice are fair) given by:

$$P(X = x) = \begin{cases} \frac{1}{36} & x \epsilon A \\ \frac{1}{18} & x \epsilon B \\ 0 & Else \end{cases}$$

With this known, consider the process any 'logical' player should go through on their turn:




The most important thing in this process, as mentioned briefly in the introduction, is the criteria by which a player decides to stick or roll again. The individual who has the greatest power in this decision in any round is the first player to roll, as the number of rolls they stick on becomes the maximum number of rolls any subsequent player can take. This first player is referred to as player 1 in this report.

*So, what criteria should player 1 use to decide whether to roll again or not?*

The criteria should depend on a few things: the number of players, the roll number player 1 is currently on and the maximum number of rolls they can take (how many rolls they have left). What impact do these variables have?

To investigate this properly, some statistics need to be calculated.

## Summary Statistics

The median m satisfies $P(Score \leq m) = 0.5$. Using what is known about the probabilities of each score, the median score for one roll is found to be 51. For the second roll and any further roll, the probability of scoring less than or equal to m in the specified number of rolls is required to be at least 0.5.

Consider the second roll: $P(Score \leq m \text{ in } 2 \text{ rolls}) = \left(P(Score \ \leq m \text{ in } 1 \text{ roll})\right)^2 = 0.5$ since the first and second rolls are independent.

When suggesting a threshold $k$ for player 1, an important quantity is $P(\ Score > k\ ) = p$. These probabilities can be used to find the approximate median m for a specified number of rolls, and for this reason the independence of rolls is used as above to tabulate the following likelihoods (for example, $P(Roll > 21 \text{ in } 3 \text{ rolls}) = (1 - (P(Roll \ \leq 21))\text{^}3\ ) = (1 - (1/12)\text{^}3\ ) \approx 0.99942$):

| k | (1 roll) ANY ROLL > $k$ | (2 rolls) ANY ROLL > $k$ | (3 rolls) ANY ROLL > $k$ |
|---|---|---|---|
| 11 | 97.222% | 99.923% | 99.998% |
| 21 | 91.667% | 99.306% | 99.942% |
| 22 | 88.889% | 98.765% | 99.863% |
| 31 | 83.333% | 97.222% | 99.537% |
| 32 | 77.778% | 95.062% | 98.903% |
| 33 | 75.000% | 93.750% | 98.438% |
| 41 | 69.444% | 90.664% | 97.147% |
| 42 | 63.889% | 86.960% | 95.291% |
| 43 | 58.333% | 82.639% | 92.766% |
| 44 | 55.556% | 80.247% | 91.221% |
| 51 | 50.000% | 75.000% | 87.500% |
| 52 | 44.444% | 69.136% | 82.853% |
| 53 | 38.889% | 62.654% | 77.178% |
| 54 | 33.333% | 55.556% | 70.370% |
| 55 | 30.556% | 51.775% | 66.510% |
| 61 | 25.000% | 43.750% | 57.813% |
| 62 | 19.444% | 35.108% | 47.726% |
| 63 | 13.889% | 25.849% | 36.148% |
| 64 | 8.333% | 15.972% | 22.975% |
| 65 | 2.778% | 5.478% | 8.104% |
| 66 | 0.000% | 0.000% | 0.000% |

**TABLE 2: APPROXIMATE LIKELIHOODS FOR SCORING GREATER THAN SCORE K**

So, the approximate median for two rolls is 61 and for three rolls the median becomes 62. For the mean, the distribution of scores is used to compute this for the first roll.

$E[First\ roll\ score\ k] = \frac{1}{18}\sum(Non - palindromic\ scores) + \frac{1}{36}\sum(Palindromic\ scores)$
$= 47.25$

As expected, the mean score is less than the median due to the positive skew of the scores (there are 6 scores above 60 but only 3 scores below 30). It's worth noting the closest possible score to this mean is actually 44. For the means of higher rolls, suppose the player only rolls again if they score less than the mean score 47.25. Then, using the theorem of total probability:

$$
\begin{aligned}
E[Second\ roll\ score\ k] \\
&= P(Roll < 47.25) * 47.25 \\
&+ \frac{1}{18}\sum(Non - palindromic > 47.25) \\
&+ \frac{1}{36}\sum(Palindromic > 47.25) = \frac{4}{9} * 47.25 + \frac{1}{18} * 525 + \frac{1}{36} * 121 \\
&\approx 53.528
\end{aligned}
$$

## Strategy for 2-player round with up to 3 rolls

What do these statistics suggest? If player 1 is in a two-player game, they have a higher chance of winning than the other player if they can roll greater than the median 51 on their first roll and then stick. On the other hand, considering the mean score of 47.25, if the player stuck on a score of 44 player 2 has a 55.566% chance of winning. This might suggest that player 1 should not stick on 44, but it will be seen later that it *is* advantageous to include 44 in the criteria for sticking.

*What should player 1 do if they cannot roll at least their chosen threshold score on their first go, when they are allowed to roll up to three times, in a game with 2 players?*

The answer is obvious, roll again! However, it poses another question, when should the player stick on their second roll? Consider the player using the median as the threshold, recall then that 61 is the median for two rolls. This is a big jump between the first and second roll thresholds, so one can tell straight away that player 1 has a lower chance of winning if they do decide to roll again. If they decide to roll a third time, the median increases to 62, and if player 1 could roll a fourth time the median would again increase to 63, the jumps here are much smaller; one might interpret this as player 1 might as well use their third roll and fourth roll if they have used their second roll. How often can player 1 achieve a score higher than the median if they can have three rolls maximum?

Using the theorem of total probability/independence of rolls and results in table 2:

$$
\begin{aligned}
P(Player\ 1\ scores\ > median) \\
&= P(Score > 51\ on\ 1st\ roll) + P(Score \leq 51\ 1st) * P(Score \geq 61\ 2nd) \\
&+ P(Score \leq 51\ 1st) * P(Score < 61\ 2nd) * P(Score \geq 62\ 3rd) \\
&= 0.73965
\end{aligned}
$$

This means that almost 3 in 4 times, player 1 can give themselves a higher likelihood of winning using the median score as a threshold than player 2, where they can take three rolls maximum.

To see what proportion of games player 1 wins with this median strategy, 300000 games were simulated. As well as this, a second competing strategy was developed, dubbed the simulated strategy. This strategy came from simulating many games where player 1 used different threshold scores each time, and the thresholds that won player 1 the most games for rolls one and two (namely 44 and 52 respectively) were combined. This alternative optimises the number of games player 1 wins (assuming sufficient games were simulated), so another 300000 games were simulated and the chances of winning calculated from that. The graph below presents the results:

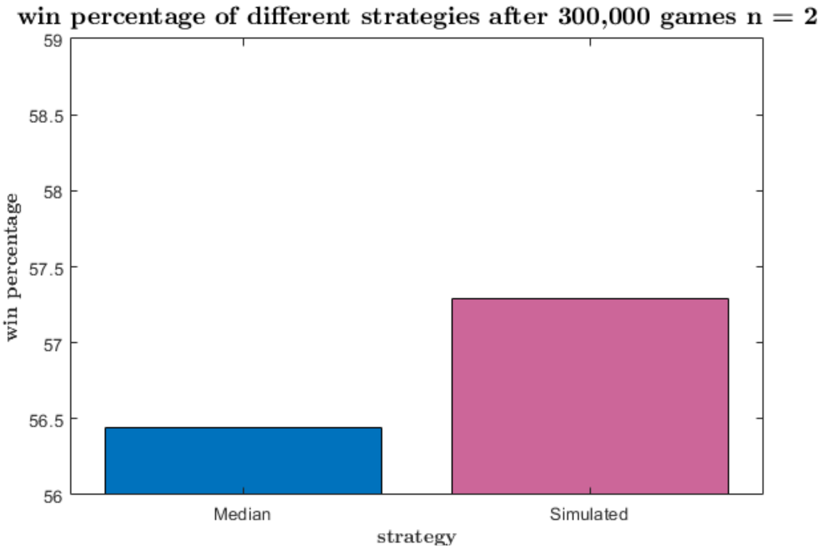



**FIGURE 1: MEDIAN VS SIMULATED STRATEGY FOR 2**

The plot shows that the median strategy led to ≈56.4% chance of winning, whereas the simulated strategy leads to ≈57.4% chance. There is only 1% in it, so it is worth comparing the two strategies to try and explain this. Firstly, the simulated first roll threshold is 44, two scores less than the median threshold of 52. For the second roll, simulated is 52 whilst median is 61 (this is a big difference as there are three possible scores in between). One possible explanation for the difference is that it's beneficial on the first roll to stick slightly lower than the median as player 1 in this case still has a decent chance of winning (at least $\frac{4}{9}$) plus their chance of obtaining the score in the first roll increases by $\frac{1}{12}$, and if they took a second roll instead then player 2 has double the number of chances to beat player 1's score. It will be observed that this explanation is sufficient, considering the theoretical probabilities of winning with each strategy.

*Theoretical probabilities of winning*

Let's split the probability into parts: (1) is where player 1 sticks on the first roll (score 52 or above) and player 2 scores less; (2) will be where player 1 rolls again and sticks on their second roll (score 61 or above) and player 2 scores less both times; (3) will be where player 1 rolls three times due to not meeting the thresholds in rolls 1 and 2, then player 2 rolls less three times.

(1): $\sum_{k=52}^{66} P(Player\ 1\ rolls\ k\ on\ 1st\ roll) * P(Player\ 2\ rolls\ t < k) = \frac{469}{1296}$

(2): $\frac{1}{2}\sum_{k=61}^{66} P(Player\ 1\ rolls\ k\ on\ 2nd\ roll) * \big(P(Player\ 2\ rolls\ t < k)\big)^2 = \frac{9715}{93312}$

(3): $\frac{25}{72}\sum_{k=21}^{66} P(Player\ 1\ rolls\ k\ on\ 3rd\ roll) * \big(P(Player\ 2\ rolls\ t < k)\big)^3 \approx 0.07808$

A few observations can be made about these probabilities. Firstly, it is clear that the chance of player 1 winning decreases massively with every additional roll they take. Secondly, these probabilities do not consider when the two players tie, in the simulations the players immediately restart the round when they tie. As a result, there will be additional probabilities of winning arising from this situation. It is clear this has been done from (3), as scoring 11 is not considered in the summation.

The total of the above is 0.54408 to 5 decimal places, so the strategy does increase player 1's chances of winning. However, the observed percentages in the graph are higher, one reason for this is that the median strategy in the graph includes sticking on a first roll score of 51. It is simple to amend the terms above for this by starting the summation at 51, which adds $\frac{1}{18} * \frac{4}{9}$ to (1). The multiplication factor in (2) and (3) also changes as these multiplication factors represent scoring less than the threshold on each roll, becoming $\frac{4}{9}$ which replaces ½. The probability that ensues from this adjustment is 0.5485, so the probability has increased as we hypothesised! The additional term added to (1) has higher impact on the overall probability than the lower multiplication factor in (2)/(3), which supports what was suggested about the odds of winning dramatically decreasing whenever player 1 re-rolls, so they should stick on the median and potentially even lower (as in the simulated strategy).

As mentioned previously, the difference in the probabilities calculated and those in the graph could be due to the situation where the players tie and replay the round. Calculating these probabilities is complex if player 1 takes more than one roll, so the probability of them winning after a tie with a first roll stick is found, then it can be discussed whether or not the probability adds enough for this explanation to be sufficient.

$$\Big(\sum_{k=51}^{66} P(Player\ 1\ rolls\ k\ on\ 1st)P(Player\ 2\ rolls\ k\ exactly)\Big) * P(Player\ 1\ wins\ replay\ with\ no\ ties) = \Big(\frac{1}{18^2} * 9 + \frac{1}{36^2} * 2\Big) * 0.5485$$

This probability is equal to 0.01608, which when added to the probability for winning gives 0.56458, which very closely resembles the graph. This is comforting, as clearly the probability of winning from ties does make up the difference.

Lastly, consider the thresholds obtained through simulation, 44 and 52. Proceeding with the methods above, the probability of winning without ties of 0.5585. Again, this is comforting as its higher than the median strategy just as in the graph, but clearly after ties are considered these probabilities will be quite close, with the simulated strategy prevailing.

It can therefore be concluded that the best strategy for player 1 in a 2-player game is to stick on the first roll if they score 44 or above and stick on the second roll if they score 52 or above.

## Strategies for rounds with more than 2 players

In the previous section, it was found that when player 1 stuck near the median score they could increase their chances of winning. This naturally implies that when there are more players (n players say), player 1 could settle for an even lower score and still have a good probability of being safe. This section proceeds with trying to find out what this lower score should be, and possible implications of sticking at certain thresholds.

Suppose it is given that player 1 has scored greater than or equal to a score (which we will denote $k$ and will depend on the roll number player 1 is on).

This $k$ could be chosen by working backwards from a probability p, where:

$p = P(Any\ player\ scores \geq k)$

Let there be a new assumption about how the other players behave. The assumption is that other players aim to beat player 1's score, and will either re-roll until they do so, or until they run out of rolls. This might seem like an odd assumption to make, as in real life a player might only roll until they are safe from elimination (i.e. until they beat the lowest score scored by anyone in the round so far) or otherwise depending how many players come after them. Despite this, rolling until they beat player 1's score would not be a terrible tactic so long as player 1 had an achievable score, and in fact the maximum that the minimum score can be for a round is that of player 1 in this situation.

Suppose now that player 1 scores exactly $k$. This means they've passed on a probability of exactly $p$ for any individual player to beat or match their score, and if player 1 instead had achieved a score greater than $k$ they'd have passed on a lower probability to the other players than $p$ (and are thus even safer for the round).

Let $N$ be the number of players that score less than player 1 (and so score less than k). Denote the event that a player scores less than player 1 as a 'success', each success then occurs with probability $1 - p$. When there is at least one 'success', player 1 is safe from elimination, and when there are $n - 1$ successes, player 1 has the highest score and thus keeps control as the first player to roll in the next round.

Each player in this situation can be thought of as an independent Bernoulli variable, so summing these variables gives:

$$N \sim Bin(n - 1, (1 - p))$$

i.e. a binomial variable with size $n - 1$ (no. of players excluding 1) and success probability $1 - p$.

The random variable $N$ can now be analysed to suggest which $p$ player 1 should use.

## Analysing the binomial model

$$E[N] = (n-1) * (1-p) \qquad Var[N] = pE[N] = (n-1) * p * (1-p)$$

It should be in player 1's interest to ensure that the mean number of players scoring less than them is at least one. Solving the inequality $E[N] > 1$ for $p$ gives $p < \frac{n-2}{n-1}$.
Unfortunately, this is not the end of the story for player 1, as pesky variance exists to ruin things. This variance means that even when the mean is greater than 1, many rounds may still result in player 1's elimination. One way that player 1 may attempt to strengthen their defence is by forcing the one standard deviation below the mean to be greater than 1.

This translates to the inequality $E[N] - \sqrt{Var[N]} > 1$. Using the expressions for these statistics in terms of $p$, the following inequality is obtained:

$(n-1) * (1-p) - \sqrt{(n-1)p(1-p)} > 1$ (*)

This inequality is difficult to solve thanks to the awkward square root. Once one has rearranged for the standard deviation, squaring the inequality is not reversible, however the resulting expression is somewhat useful when you consider the smallest root:

$$p < \frac{[(2n-3)(n-1) - \sqrt{(3-2n)^2(n-1)^2 - 4n(n-1)(n-2)^2}\,]}{2n(n-1)}$$

This is the inequality that satisfies (*).

A 'mental check' that this $p$ is about right is that as the number of players $n$ approaches infinity, the right-hand side approaches 1 as the leading term on the numerator and denominator is $2n^2$. This is the equivalent of saying when there are many players, player 1 can stick on any number they roll (as long as its greater than the minimum 11) and be safe, as other players will always roll less eventually in the round. This can be visualised by plotting the expression on the left-hand side of the original inequality along with a reference line at $y = 1$:

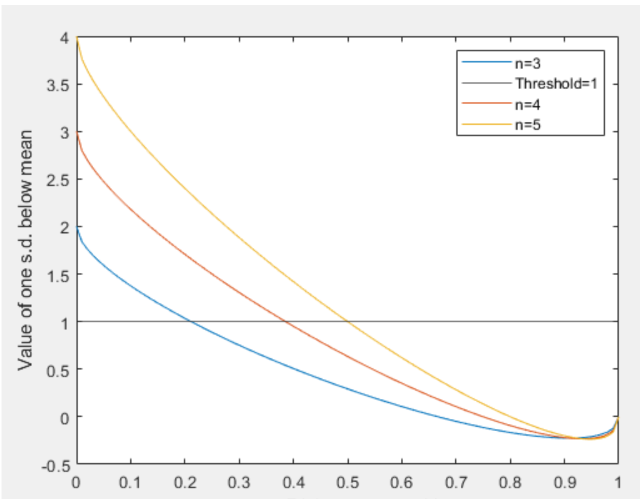



FIGURE 2: GRAPH SHOWING HOW SAFETY VALUE CHANGES WITH P FOR DIFFERENT N

It is clear that as the number of players increases the value of $p$ where the inequality becomes an equality also increases. This means with more players, player 1 can ensure one standard deviation below the mean is above 1 more easily as we might expect. One issue that is clear for these low player rounds is that $p$ is very small so it might be very difficult for player 1 to achieve the score they need.

For example, for three players, the formula given above tells us the equality value is $p \approx 0.21132$. How often can player 1 achieve this? Clearly the threshold score $k$ will be very high, and indeed using the table of likelihoods in the first section of the investigation gives:

| Roll Number | Minimum score required to achieve $p$ |
|---|---|
| 1 | 63 |
| 2 | 65 |
| 3 | 66 |

TABLE 3: MINIMUM SCORES TO OBTAIN P FOR ROLLS 1-3

Player 1 achieves this over their 3 rolls with probability:

$$\begin{aligned} &P(Rolls\ \geq 63\ on\ roll\ 1) + P(Rolls < 63\ on\ roll\ 1)P(Rolls \geq 65\ on\ roll\ 2) \\ &\qquad + P(Rolls < 63\ on\ roll\ 1)P(Rolls < 65\ on\ roll\ 2)P(Rolls\ 66\ roll\ 3) \\ &\qquad = \frac{7}{36} + \frac{29}{36} * \frac{1}{12} + \frac{29}{36} * \frac{11}{12} * \frac{1}{36} \approx 0.28209 \end{aligned}$$

This means player 1 cannot even achieve this 1 in every three rounds, so clearly using this criterion is over the top for this many players.

It is mentioned above that $p$ converges to 1 as the number of players goes to infinity. This means that surely this strategy will become achievable after a certain large number of players are involved.

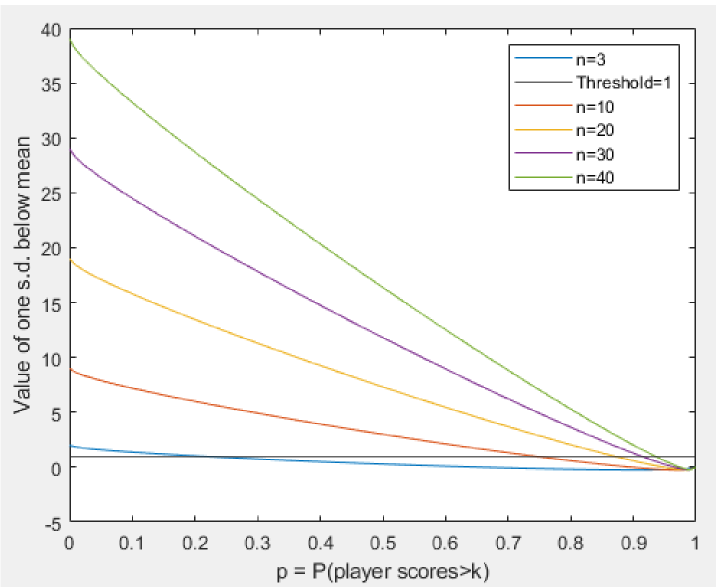



FIGURE 3: GRAPH SHOWING HOW SAFETY VALUE CHANGES WITH P FOR LARGE N

Indeed, the plot above shows that if there are even 10 players in a round, this strategy is an achievable one for player 1. Still, a massive assumption was made in saying that the players roll to beat player 1, and in fact when games are simulated, the strategy other players used was simply to beat the lowest score achieved before them in a round. This means the binomial model is often an underestimate for the true number of players that score less than player 1, and indeed only holds until a player in the round does score less than them. As such, simply making the mean of the binomial model greater than one might be sufficient to ensure the safety of player 1, as this will be investigated later by using this strategy as a competitor against the simulated thresholds.

Other important quantities arising from the binomial model include:

$P(N = 0) = p^{n-1}$ *and* $P(N = n - 1) = (1 - p)^{n-1}$

The first is the probability that all players beat the score of player 1 so player 1 is eliminated, whilst the second is the probability that player 1 beats all other players by obtaining the highest score, and so gets to roll first in the next round.

Can these help us decide the best course of action for player 1 in the three-player game?

In the $n = 3$ player game, player 1 might prioritise maintaining control for the final round as we know that being the first player to roll in the final round is a major advantage. This translates to maximising $(1 - p)^{n-1}$ which of course requires $p$ to be as close to zero as possible, and so reduces the likelihood that player 1 can achieve the threshold in their three rolls. Player 1 must try to find a balance between maximising this formula and being able to achieve the score required to maximise it.

For example, they might try setting $(1 - p)^2 = 0.4$, which results in $p = 0.3675$ and $p^2 = 0.13506$. Using the definition of $p$ and the table of likelihoods to find the $k$-values then using the method on the previous page, we find the probability that player 1 can obtain the scores in their three rolls is $\approx$0.53736. Clearly this is not very high, and since the probability of keeping control is quite low even when the scores are achieved, this is not a good strategy for player 1.

One might ask the question: should player 1 just try and survive a round with three players as best they can, not aiming to maintain control for the final round?

For a game in real life, it really depends on the strategies of the other players. If the other players know what to do to give themselves the advantage as player 1 in the final round, then player 1 should take some measured risks. On the other hand, if the other players are amateurs with a lack of knowledge in probability, maintaining control is not quite as important as player 1 still has a solid chance of winning the final round.

All of the criteria considered for the threshold to be derived from have disadvantages depending on the number of players. The main disadvantage is that it is difficult for player 1 to achieve a score which gives them a significant advantage, and in attempting to do this they risk rolling badly a third time and being eliminated.

There must be thresholds which have the best balance between being easily achievable for player 1 and also minimising their risk of elimination (maximising chances of winning). Let simulations be used to find these thresholds.

## Strategies from simulations

Earlier in this report, it was found that for the two-player game the thresholds which maximised player 1's chances of winning are 44 for roll one and 52 for roll two. These thresholds were found by simulating many two-player rounds with different strategies and identifying which strategy led to player 1 winning the most games. This method can be extended to obtain the optimised simulated strategy for higher player rounds:

| Number of players $n$ | First roll threshold | Second roll threshold |
|---|---|---|
| 3 | 43 | 51 |
| 4 | 41 | 44 |
| 5 | 41 | 43 |
| 6 | 33 | 43 |
| 7 | 33 | 43 |
| 8 | 33 | 42 |
| 9 | 32 | 42 |

TABLE 4: OPTIMISED THRESHOLD VALUES FROM SIMULATION

It is known from the previous analysis (see figures 2 and 3) that the thresholds would decrease with increasing n, but by this much is quite surprising! The aim is to find a strategy backed up by theory that might explain the simulation, and with this objective in mind return to the mean of the binomial distribution. Recall that in order for the number of players $N$ scoring less than player 1 to be at least 1, it follows that $p < \frac{n-2}{n-1}$. Consider using $p = \frac{n-2}{n-1}$, so that the mean is exactly 1, then for each number of players $n$ we obtain (from the definition of $p$ and the table of likelihoods) the following thresholds:

| Number of players $n$ | $p$ | Roll 1 threshold | Roll 2 threshold |
|---|---|---|---|
| 3 | ½ | 52 | 61 |
| 4 | 2/3 | 43 | 54 |
| 5 | 3/4 | 41 | 52 |
| 6 | 4/5 | 33 | 51 |
| 7 | ≈0.833 | 32 | 44 |
| 8 | ≈0.857 | 32 | 43 |
| 9 | ≈0.875 | 31 | 43 |

TABLE 5: THRESHOLD VALUES FOR MEAN=1 STRATEGY

The two methods can now be compared. Interestingly, the mean strategy has lower thresholds for roll 1 when $n = 7,8,9$, whilst the roll 2 thresholds are always at least one score higher than the simulated strategy. As $n$ increases, the two methods seem to converge.

With the threshold values known for these n, 300000 rounds were simulated for each $n$. The results (note that the best strategy minimises elimination percentage) for $n = 3,6,9$ can be found below (see appendix for $n = 4,5,7,8$):

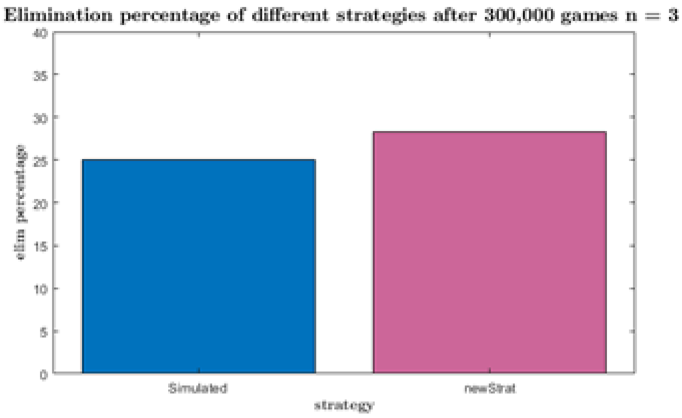



FIGURE 4: COMPARING STRATEGIES FOR 3-PLAYER ROUND

When $n = 3$, there is ≈3% difference between the elimination chances of the methods, with the simulated strategy providing a 25% probability of elimination. If the chance of elimination with no strategy is distributed equally between the players ($\frac{1}{3}$), the simulated strategy decreases this by $\frac{1}{12}$.

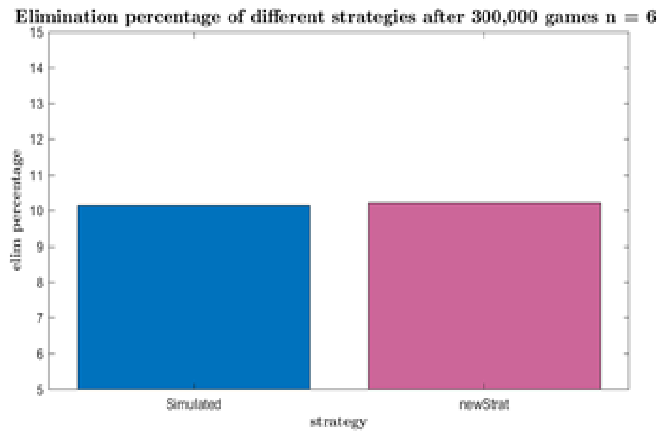



FIGURE 5: COMPARING STRATEGIES FOR 6-PLAYER ROUND

When $n = 6$, you can barely tell the difference between the methods, however the simulated strategy comes out on top with elimination chance of 10.143% vs 10.2163% for the mean strategy. This makes sense as they have the same threshold for roll 1 (33) and the second threshold is only two scores higher for the mean strategy (51).

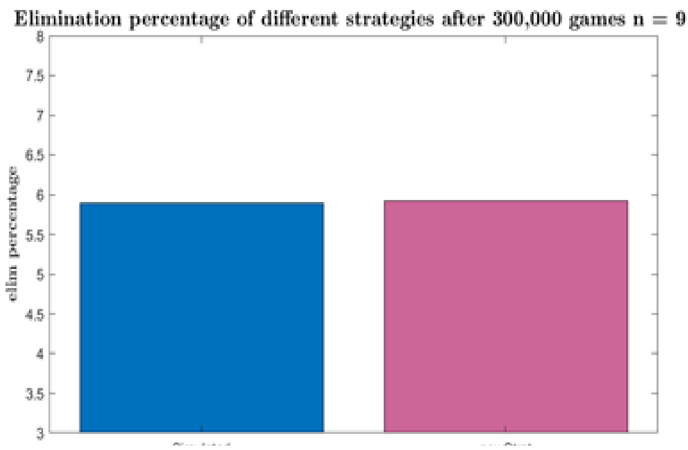



FIGURE 6: COMPARING STRATEGIES FOR 9-PLAYER ROUND

Again, for $n = 9$ the strategies are even closer, the simulated strategy now having an elimination probability of 5.8933% vs 5.9160 for the mean strategy. Clearly the theoretical strategy described above is very near to being the optimum!

Strategies for other players can now be considered.

## Strategies for other players

This section is concerned with strategies for other players, with particular emphasis on player 2. A good strategy will be related to the minimum score achieved by players so far in the round, the number of players left to roll, and the number of rolls a player can take (which is determined by the choices of player 1). The first observation worth making is that the only viable strategy for the final player in any round is to roll until they beat the minimum score. The strategies can be considered in turn on a case-by-case basis, conditioning on the decisions of player 1.

*Case 1: Player 1 sticks on their first roll*

In this situation, regardless of the number of players, every player only gets to roll once (i.e. there are no decisions to make). There is no strategy that can be implemented to increase the odds of survival.

*Case 2: Player 1 sticks on their second roll*

Now, each player has a decision to make -- stick on roll 1 or roll a second time. The player should stick if they score above the minimum achieved so far in the round, or if they score sufficiently high so that at least one of the remaining players should score less than them. But what is a sufficiently high score?

To illustrate what is meant by 'sufficiently high', consider the $n = 3$ round. Here, player 2 must either beat player 1's score or achieve a score high enough to beat player 3 to survive. Suppose that player 1 scores highly, say above 63, then player 2 may as well ignore this when considering re-rolling as chances are, they will not beat 63 or above in one roll. They should focus instead on obtaining a score that is difficult for player 3 to beat in their two rolls. This essentially reduces the round to a 2-player round, and the simulated threshold we found previously for two rolls is 52. Generally, for an $n$-player round, if player 1 scores highly and player 2 decides to ignore them, they could use the simulated roll 2 threshold for the $(n-1)$-player round as their sticking factor for roll 1.

| No. of players (n) | 1st roll threshold for player 2 (ignoring player 1) |
|---|---|
| 3 | 52 |
| 4 | 51 |
| 5 | 44 |
| 6 | 43 |
| 7 | 43 |
| 8 | 43 |
| 9 | 42 |

**TABLE 6: THRESHOLDS FOR PLAYER 2 IF PLAYER 1 SCORES HIGHLY**

The effect of using these thresholds when player 1 scores 61 or above is shown below. The plot shows, as expected, that the thresholds reduce player 2's chances of elimination by a significant margin (can draw exact comparisons with $(n-1)$-player rounds).

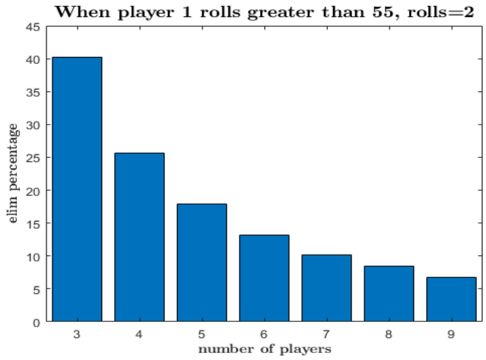
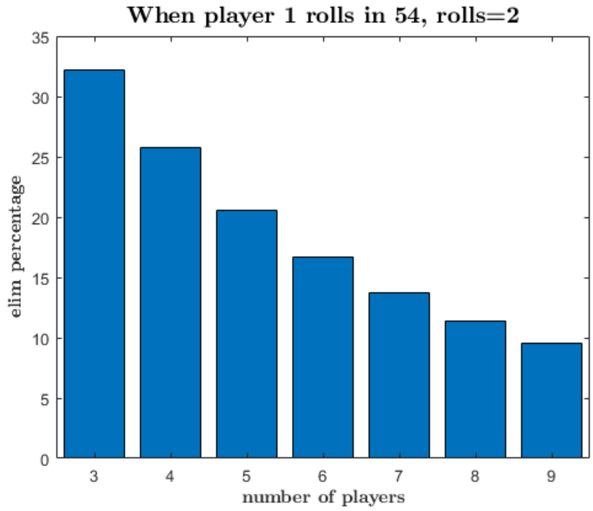



**FIGURE 7: ELIMINATION PERCENTAGE WITH OUR STRATEGY FOR WHEN PLAYER 1 SCORES**

Now, suppose instead player 1 does not score as highly, but they stick on roll 2 (e.g. 51, 52, 53 but this could be any score if player 1 is only able to take two rolls maximum). How should this effect the threshold for player 2? In a three-player game, the threshold suggested above is 52, which is quite high and should ensure that player 2 can beat player 1's score quite often. This is especially significant for a 3-player round as when player 2 beats player 1 they're in the running to being the first player to roll in the final round, which is known from our previous results to provide a great advantage. On top of this, if player 1 has the lowest score, player 3 will concentrate on beating this as opposed to getting the highest score. In this way, when the player 1 score is below the threshold it could still be of benefit to keep 52 as player 2's threshold, and when player 1 scores slightly above 52 it may be worthwhile for player 2 to increase the threshold to match that score.

In higher player games, the thresholds suggested above can be improved. Recall that these thresholds are based upon those assigned to player 1, and they were slightly lower than expected to account for the number of rolls player 1 passes on to the players for the rest of the round (more rolls increased chance of elimination for player 1). So, in this respect, when player 1 obtains an achievable score (e.g., scores of 54 or below say), player 2 should take this into consideration and adjust their threshold accordingly. In a game with many players, player 2 should try not to become the new lowest score, as in the simulations that makes them the target for all other players. Additionally, considering the above threshold for a 9-player game, if player 2 rolls 42 on the first roll they will stick, but if they decided to roll again, they will often roll higher than 42 (the median for one roll is 51). Sticking on 42 could stop player 2 from obtaining a high score overall and, if they score less than player 1, makes them the target for other players. This has an adverse effect on the chance of survival. This is visualised with the figures below:



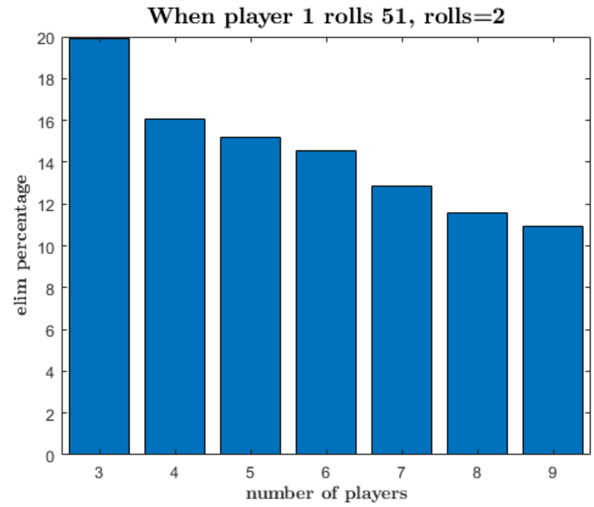



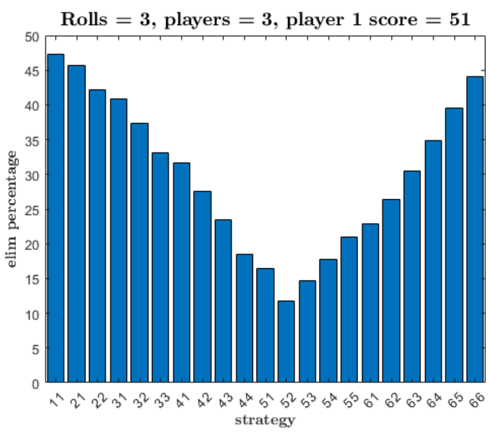

**FIGURE 8/9: PLOTS SHOWING ELIMINATION PERCENTAGE WITH TABLE 5 THRESHOLDS WHEN PLAYER 1 ROLLS 54 OR 51 IN 2 ROLLS**

The plot on the left shows that the risk of elimination goes down when player 1 rolls 54 in a 3-player game, as the threshold for player 2 is already 52 so there will be many rounds where player 2 can beat player 1's score. On the other hand, for $n = 9$ the risk of elimination goes up as the threshold of 42 means that quite often player 2 will stick on their first roll and become the target for other players. The effect of the 42 threshold becomes even more clear in the right plot where the risk of elimination goes up above 10%. Since 51 is achievable in two rolls 75% of the time, a threshold of 42 is clearly too low to minimise the chance of elimination.

Indeed, if many games are simulated with different strategies and the elimination percentage associated with using each (see appendix for plot) is obtained, it is found that 51 is the optimum threshold, the same score as player 1. For two rolls, when player 1 scores lower, player 2's threshold moves towards their score, and when they score higher the player 2 threshold moves towards that found in the table (where player 1 is ignored).

*Case 3: Player 1 takes three rolls*

If player 1 takes three rolls, one can expect the effect on the threshold for player 2 to be even greater as they now have three opportunities to beat player 1's score. Some plots for situations in a three-player round are presented below:



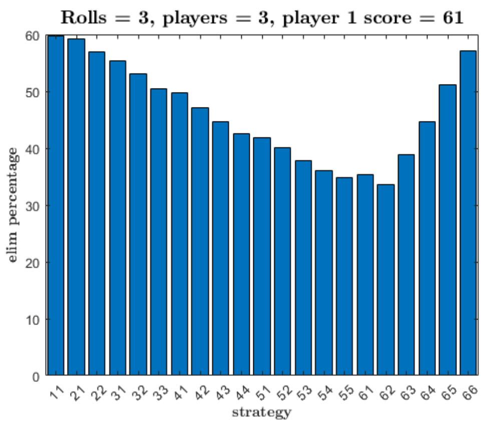




**FIGURE 10/11: PLOTS SHOWING HOW PLAYER 2 THRESHOLD IN 3-PLAYER GAME SHIFTS TOWARDS PLAYER 1'S SCORE**

These graphs show that this effect is far greater if player 1 takes three rolls, and indeed if player 3 gets 51 on their third roll, player 2 can use a threshold of 52 to get a probability of elimination of <15% (which makes sense as 51 is the median for one roll meaning player 2 will be able to beat this in 87.5% of rounds). If player 1 rolls 61, setting the player 2 threshold to 62 minimises elimination chance. This makes sense as in three rolls (six tosses of a dice), considering a geometric distribution of probability $\frac{1}{6}$ of success (where a 6 is rolled), the average number of tosses for this success is 6, and when a 6 is obtained there is 5/6 chance the other throw in that roll is not 1.

Extending what has been found to other players gives us a general strategy: for player $k$ in an $n$-player game with $k \neq n$, if the $k-1$ players before them all get high scores, use the threshold values for an $n-k$ round. Otherwise, consider the minimum score so far, if it is achievable (can achieve it in more than 75% of rounds say) in 2/3 rolls set the threshold one above this score, otherwise set the threshold just below the minimum score.

## Extension – The three-dice game

One possible extension for our game is the addition of dice. In this section, three dice are considered instead of the two dice used earlier in the report. A table of likelihoods produced in a similar way to table 2 (see The Investigation) is shown in the appendix, where an assumption akin to that for two dice is made that scores are always maximised.

The median values are 543 for one roll, 641 for two rolls and 644 for three rolls and so consequently, if the strategy of sticking on the median is used, player 1 would not roll again if they achieved these values or higher. The probability that player 1 will achieve any of these median scores on their go is 0.72. This is similar to the strategy for two dice but slightly less, showing the strategy in this case is less achievable.

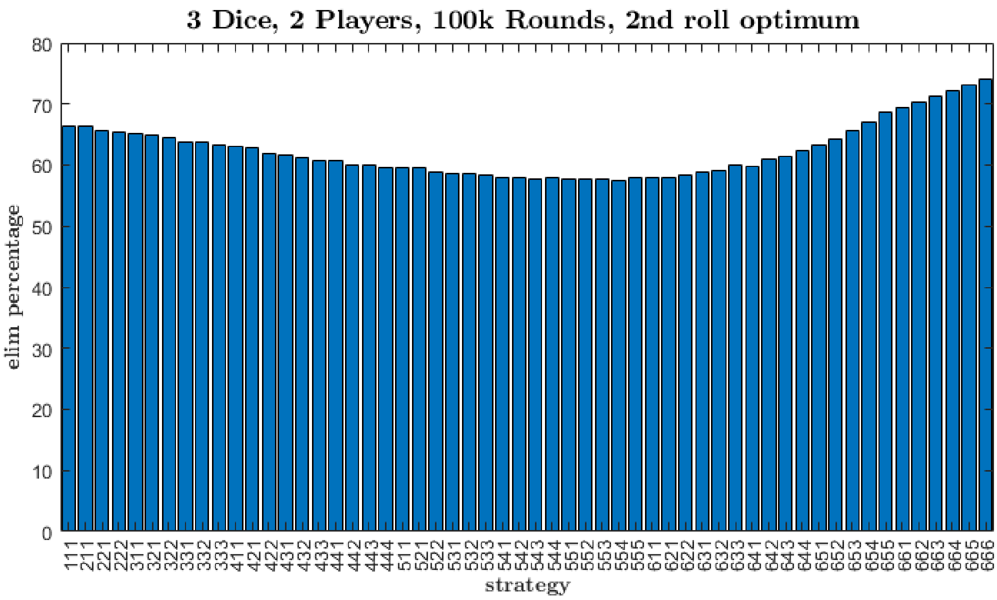



**FIGURE 12: INVESTIGATING SECOND ROLL THRESHOLDS IN THE THREE-DICE TWO-PLAYER GAME**

To find the sticking threshold for the second roll, a scenario where player 1 has already rolled once and therefore can only take two or three rolls was simulated. This means only the second threshold is relevant and one can run each possible option 100,000 times and see which one offers the lowest chance of eliminating player 1.

The result is shown in figure 12, from which two observations can be made: the best second roll threshold is 554 and it has a chance of eliminating player 1 in 57.6450% of the 100,000 games. Before the first roll plots are simulated, a guess can be made of what the best first threshold will be because of this. If taking a second roll causes player 1 to lose in 57.6450% of games, they should be picking a first roll threshold that they can achieve relatively easily to avoid these odds. This threshold can be found using the table (see appendix), it can be observed that 541 gives you a 55.09% chance of losing which is not amazing but it provides better odds than rolling again.

The simulations are summarised in the figure below. Player 1's second sticking threshold is set at 554 and 100,000 games are run at each possible first sticking threshold to see the result. The best threshold strategy is 541 as suggested above, and the combination of these thresholds gives a win percentage of 59.24%. So, the best strategy for player 1 is 541 for the first roll and 554 for the second.

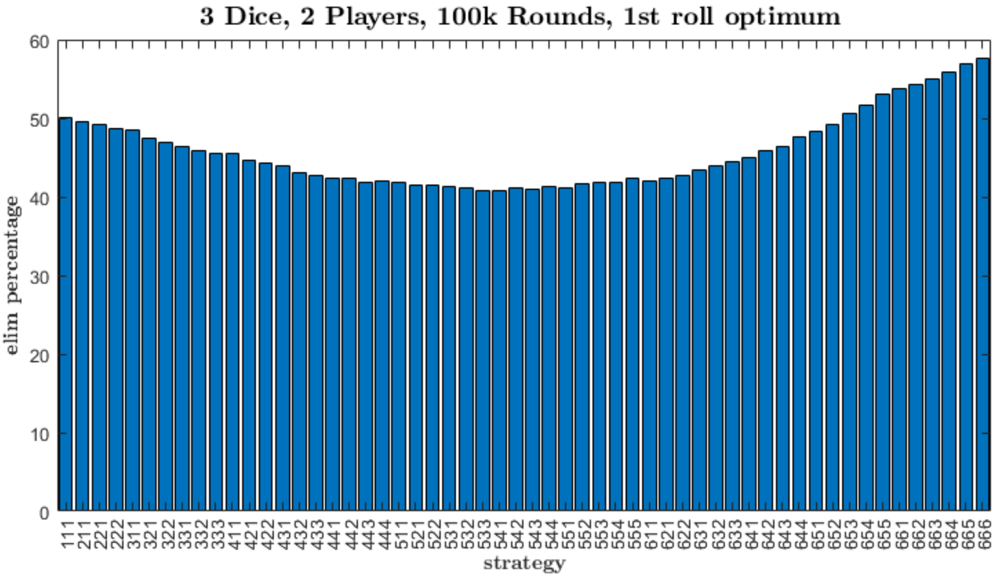



**FIGURE 13: INVESTIGATING FIRST ROLL THRESHOLDS FOR THREE-DICE TWO-PLAYER GAME**

In both games, referring to two dice and three dice, the optimised threshold values that maximises the chances of winning found by simulating lots of games are lower than the median values from the tables.

Further research could also be conducted by investigating how the strategy changes with the traditional rules of Mexico, where higher scores numerically are not as awarding as palindromic scores and a score of 21. This would of course affect the sticking threshold values. Besides this, biased dice or many-sided dice would add another level of complexity and chaos which would be interesting to investigate.

# Conclusion

The main purpose of this investigation was to determine the optimal strategies for players in a game of 'Mexico', that is, the best choices of when to 'stick' or 'roll again'. This is influenced heavily by the number of players in the game, the maximum number of rolls that one can use and the number of dice. The main bulk of this report considered the "three-roll, two-dice" game and the strategies that each player in the game should take based on when they roll in the round.

To tackle the problem, the most simplistic version of the game was considered first, in which there are just two players. Having identified that player 1 has the greatest influence on the game, this basic game was used to explore a number of potential strategies for player 1 using fundamental theories of probability. This led to using the median scores as thresholds for rolls one and two. To shed some light on the efficacy of this theoretical strategy, thousands of two player games were simulated in which player 1 uses many different threshold scores and a strategy was identified that maximised the number of rounds won by player 1. It was discovered that in fact the optimal thresholds for player 1 to stick with are 44 and 52 for their first and second rolls respectively. Using this strategy, player 1 can alter the odds of winning the game to ≈57.5% in their favour.

The report then moved on to considering games with $n$ players, in which case it was assumed that all players following player 1 strive to beat player 1's score. The idea of using a binomial variable to model the number of players scoring less than player 1 was analysed. The general first and second roll thresholds for player 1 to use in an $n$-player game were determined by setting the probability that any player scores higher such that the mean of the binomial model is at least 1 (average number of players scoring less than player 1 is at least 1). To validate this theory, once again thousands of simulations of $n$-player rounds (for $n \in \{1,2,\ldots,10\}$) were produced, where player 1 uses differing thresholds. Crucially, as the number of players $n$ increases, the theoretical and simulated thresholds converge, showing that the "mean=1" method is a good strategy for higher player games.

Finally, a general strategy was devised for all the other players in an $n$-player round by conditioning on the score of player 1 when they roll more than once. For players other than those rolling first or last, the strategies were found to be more complex. It was found that for player $k$ in an $n$-player game with $k \neq n$, if the $k-1$ players before them all get high scores, $k$ can use the threshold values for an $n-k$ round. Otherwise, if the minimum score so far is deemed achievable, they should set their threshold above that score; if the minimum is not deemed achievable then they set their threshold just below the minimum score.

It is worth noting that in this report only simulated strategies for games with up to 10 players were considered, which seems sensible because it is hard to imagine that many dice games in reality are played with more than 10 players. However, with the general strategy devised, it would be quite simple to extrapolate the strategy to games where $n$ is greater than 10. Clearly, this is a potential area for future analysis. In the simulations, it was also assumed that players do not consider the incentive of achieving the highest roll in a round in order to roll first in the next round. In reality, the players rolling later in a particular round may use the fact that they have seen all previous scores to their advantage, and attempt to be the highest scorer of the round if they deem this to be achievable. Furthermore, the way ties were dealt with in our simulations may not be how ties are handled traditionally, other methods could be considered such as coin tosses, further dice rolls or rock-paper-scissors. Extensions of our standard game to three dice were also considered and optimised thresholds by simulation found. Obviously, even more dice could be introduced to the game, biased dice could be used and the maximum number of rolls for player 1 could be increased. All of these are ideas for further research, and excellent strategies could be found based upon the methods used in this investigation.

This report seems to stand alone on the analysis of 'Mexico', especially since a modified set of rules have been used. However, there have been several published papers on other logic-based dice games. For example, 'Pig' has been analysed extensively in "Dice Games and Stochastic Dynamic Programming" by Henk Tijms (2007) and also by Neller and Presser in "Optimal Play of the Dice Game Pig", where the use of value iteration is incorporated in their investigation. The key motivation of this report stems from the work of Morris (1994) which studies the mathematical theory of games, with a goal of advising people of the best tactics to adopt in some popular games. Haigh *et al* (2000) looked at the application of the concepts of probability, Markov Chains and decision theory, and the techniques used by them have influenced the methods for finding the best strategy in this report. The papers mentioned here also present some interesting alternative methods to finding optimal strategies in games and are therefore worth considering as further reading beyond this report.

Ultimately, it must be stressed that a player's chance of winning is heavily influenced by their assumptions of the rationality of other players. For instance, if there are some players involved who have little to no concept of probability, competing against players with greater expertise, then it would be natural to assume that those experienced players need to take less risks and could quite possibly reduce their threshold values without disadvantaging themselves. Essentially, although strategies have been found that maximise a player's chance of survival/winning, Mexico is still a game of luck, and the proportion of games a player with a certain strategy wins only reveals itself when many games are played. This highlights the limitations of the strategies in this report, where the assumptions made on human behaviour in a game of 'Mexico' will rarely hold.

# References

Haigh, J., & Roters, M. (2000). Optimal strategy in a dice game. *Journal of Applied Probability*, *37*(04), 1110-1116.

Krishnamachari, R. (2011) *Dynamic programming – Probability, gambling and a die*. Thoughts. (2020). Retrieved 19 November 2020, from https://tkramesh.wordpress.com/2011/04/17/dynamic-programming-probability-gambling-and-a-die/.

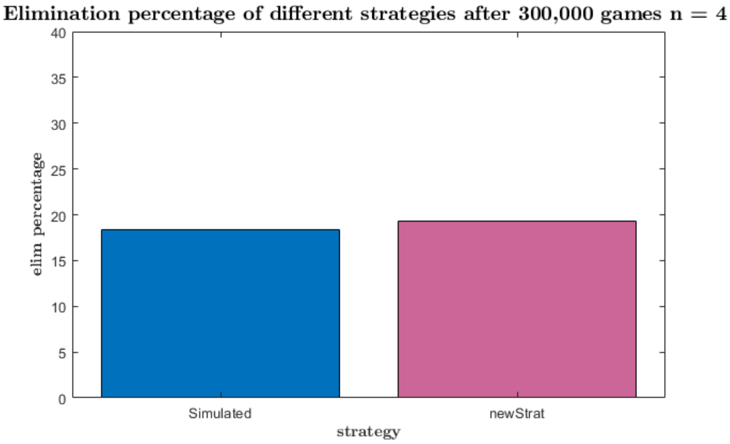
Morris, P. (1994) Introduction to game theory. Springer-Verlag, New York.

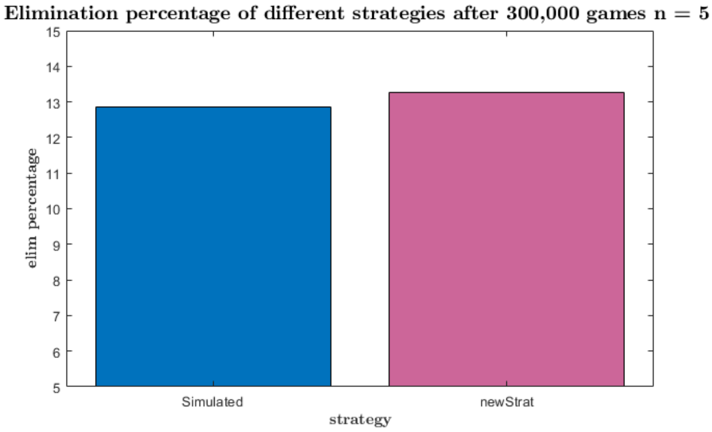
Neller, T. W., & Presser, C. G. M. (2004). Optimal play of the dice game pig. *The UMAP Journal*, 25(1), 25–47. http://www.amstat.org/publications/jse/v14n3/datasets.kern.html

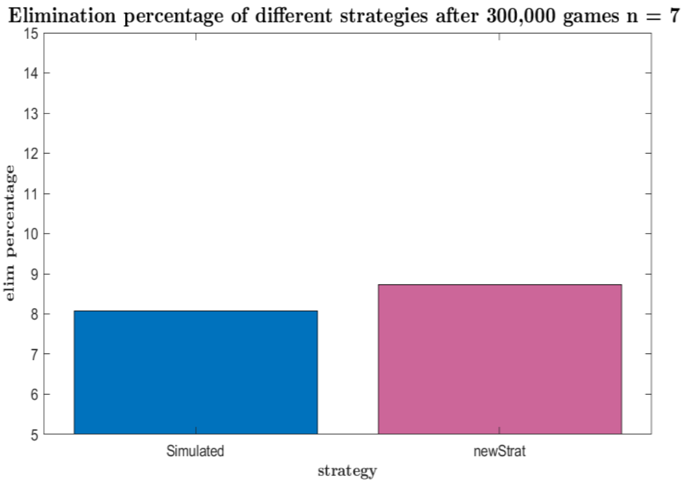
Tijms, H. (2007). Dice Games and Stochastic Dynamic Programming. https://www.researchgate.net/publication/255585804_Dice_Games_Stochastic_Dynamic_Programming

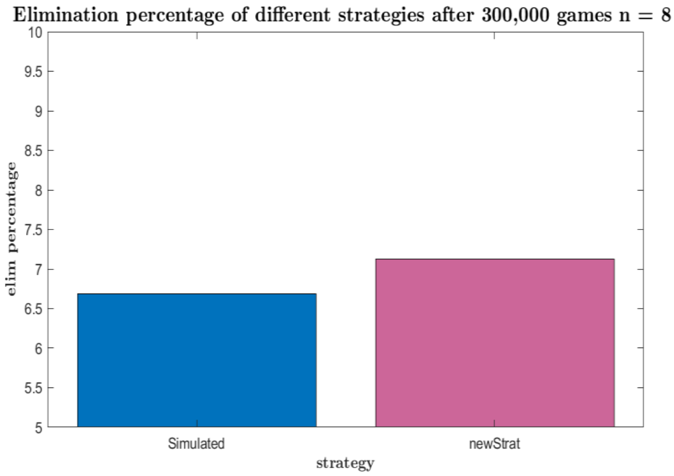
# Appendix

The following plots show the comparison between the elimination percentage using the optimised simulated strategy versus the binomial mean equals one strategy for $n = 4,5,7,8$.




The table on the next page shows the likelihoods of scoring greater than a certain score within one roll, two rolls and three rolls, in the game with three dice.

| Probability of rolling | Score | Chance of being beaten on 1 roll | Chance of being beaten on 2 rolls | Chance of being beaten on 3 rolls |
|---|---|---|---|---|
| 0.00463 | 111 | 0.99537037 | 0.999978567 | 0.999999901 |
| 0.013889 | 211 | 0.981481481 | 0.999657064 | 0.999993649 |
| 0.013889 | 221 | 0.967592593 | 0.99894976 | 0.999965964 |
| 0.00463 | 222 | 0.962962963 | 0.998628258 | 0.999949195 |
| 0.013889 | 311 | 0.949074074 | 0.99740655 | 0.999867926 |
| 0.027778 | 321 | 0.921296296 | 0.993805727 | 0.999512488 |
| 0.013889 | 322 | 0.907407407 | 0.991426612 | 0.999206168 |
| 0.013889 | 331 | 0.893518519 | 0.988661694 | 0.99879268 |
| 0.013889 | 332 | 0.87962963 | 0.985510974 | 0.998255951 |
| 0.00463 | 333 | 0.875 | 0.984375 | 0.998046875 |
| 0.013889 | 411 | 0.861111111 | 0.980709877 | 0.997320816 |
| 0.027778 | 421 | 0.833333333 | 0.972222222 | 0.99537037 |
| 0.013889 | 422 | 0.819444444 | 0.967399691 | 0.994113833 |
| 0.027778 | 431 | 0.791666667 | 0.956597222 | 0.990957755 |
| 0.027778 | 432 | 0.763888889 | 0.944251543 | 0.98683717 |
| 0.013889 | 433 | 0.75 | 0.9375 | 0.984375 |
| 0.013889 | 441 | 0.736111111 | 0.930362654 | 0.981623478 |
| 0.013889 | 442 | 0.722222222 | 0.922839506 | 0.978566529 |
| 0.013889 | 443 | 0.708333333 | 0.914930556 | 0.975188079 |
| 0.00463 | 444 | 0.703703704 | 0.912208505 | 0.973987705 |
| 0.013889 | 511 | 0.689814815 | 0.903785151 | 0.970155579 |
| 0.027778 | 521 | 0.662037037 | 0.885781036 | 0.96139822 |
| 0.013889 | 522 | 0.648148148 | 0.876200274 | 0.956440837 |
| 0.027778 | 531 | 0.62037037 | 0.855881344 | 0.945288288 |
| 0.027778 | 532 | 0.592592593 | 0.834019204 | 0.932378194 |
| 0.013889 | 533 | 0.578703704 | 0.822509431 | 0.925223881 |
| 0.027778 | 541 | 0.550925926 | 0.798332476 | 0.909436343 |
| 0.027778 | 542 | 0.523148148 | 0.772612311 | 0.89156976 |
| 0.027778 | 543 | 0.49537037 | 0.745348937 | 0.871495528 |
| 0.013889 | 544 | 0.481481481 | 0.731138546 | 0.860590357 |
| 0.013889 | 551 | 0.467592593 | 0.716542353 | 0.849085049 |
| 0.013889 | 552 | 0.453703704 | 0.701560357 | 0.836963528 |
| 0.013889 | 553 | 0.439814815 | 0.686192558 | 0.82420972 |
| 0.013889 | 554 | 0.425925926 | 0.670438957 | 0.81080755 |
| 0.00463 | 555 | 0.421296296 | 0.665102023 | 0.806193301 |
| 0.013889 | 611 | 0.407407407 | 0.648834019 | 0.791901641 |
| 0.027778 | 621 | 0.37962963 | 0.615140604 | 0.761244634 |
| 0.013889 | 622 | 0.365740741 | 0.597715192 | 0.744847136 |
| 0.027778 | 631 | 0.337962963 | 0.561706962 | 0.709833775 |
| 0.027778 | 632 | 0.310185185 | 0.524155521 | 0.671755429 |
| 0.013889 | 633 | 0.296296296 | 0.504801097 | 0.651526698 |
| 0.027778 | 641 | 0.268518519 | 0.464934842 | 0.608609746 |
| 0.027778 | 642 | 0.240740741 | 0.423525377 | 0.562306305 |
| 0.027778 | 643 | 0.212962963 | 0.380572702 | 0.512487775 |
| 0.013889 | 644 | 0.199074074 | 0.358517661 | 0.486220164 |
| 0.027778 | 651 | 0.171296296 | 0.313250171 | 0.430887874 |
| 0.027778 | 652 | 0.143518519 | 0.266439472 | 0.371718992 |
| 0.027778 | 653 | 0.115740741 | 0.218085562 | 0.308584919 |
| 0.027778 | 654 | 0.087962963 | 0.168188443 | 0.241357052 |
| 0.013889 | 655 | 0.074074074 | 0.14266118 | 0.206167759 |
| 0.013889 | 661 | 0.060185185 | 0.116748114 | 0.169906792 |
| 0.013889 | 662 | 0.046296296 | 0.090449246 | 0.132558077 |
| 0.013889 | 663 | 0.032407407 | 0.063764575 | 0.094105538 |
| 0.013889 | 664 | 0.018518519 | 0.036694102 | 0.0545331 |
| 0.013889 | 665 | 0.00462963 | 0.009237826 | 0.013824688 |
| 0.00463 | 666 | 0 | 0 | 0 |