

UNIVERSITY OF DUBLIN, TRINITY COLLEGE



DICE INVESTIGATION

ST2004 & ST2353 ASSESSMENT 1

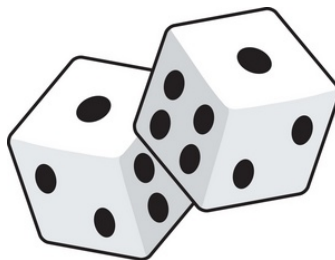
Authors:

Alexandru SULEA 12315152
Edmond O'FLYNN 12304742
Jonathan LESTER 14310385
Ronan CAMPBELL 14324880

Lecturer:

Dr. Brett HOULDING

November 19, 2015

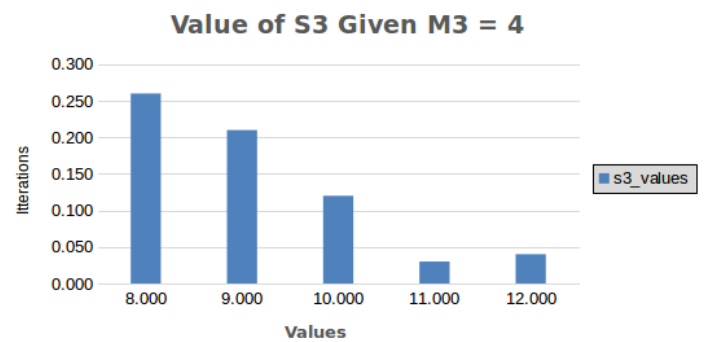
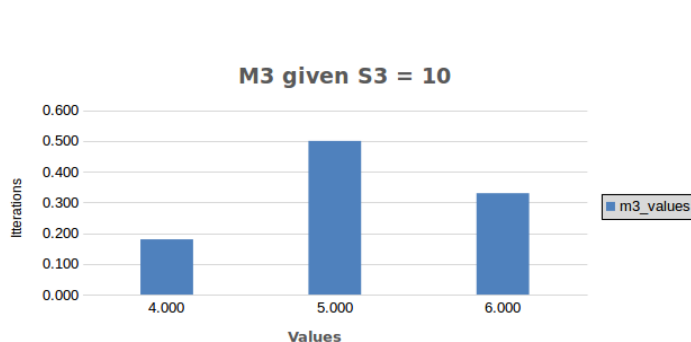
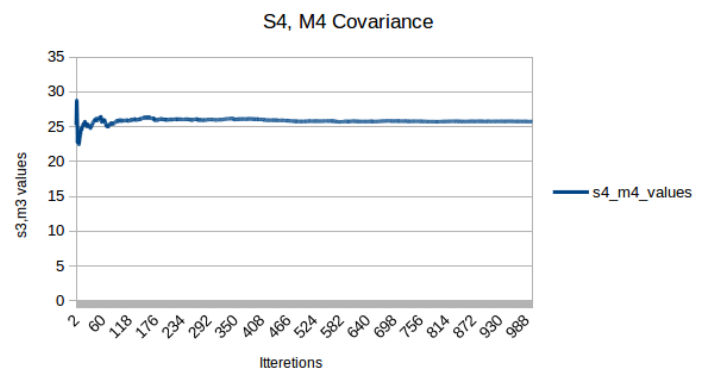
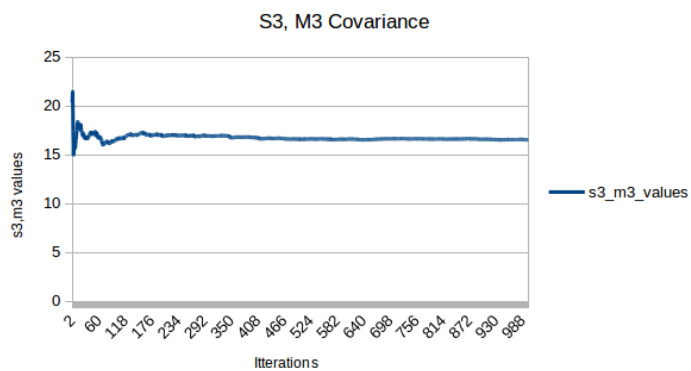
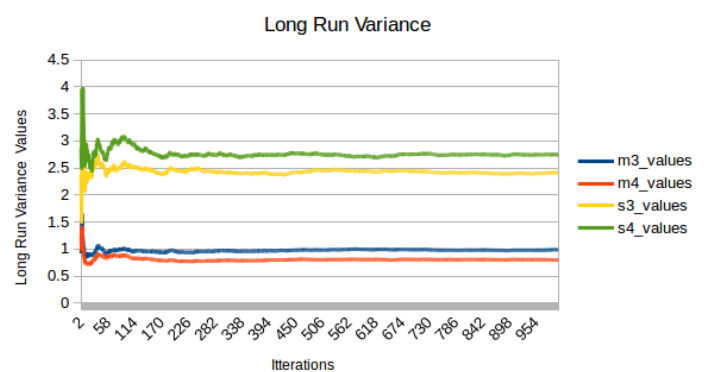
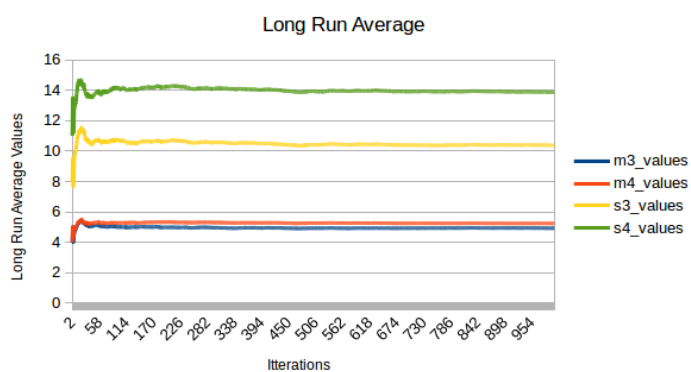
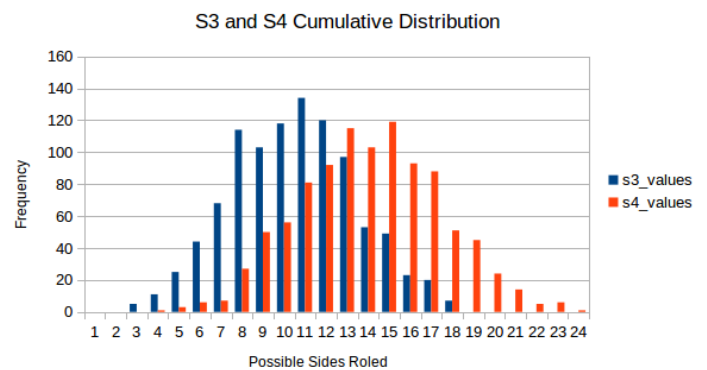
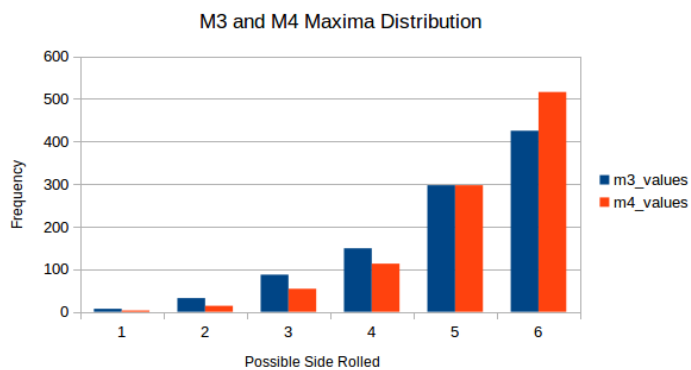


Contents

References	1
1 Q1i: Simulation of Probability Distributions	1
1.1 Graphs	1
2 Q1ii: Probability Distributions	4
2.1 Probability Distribution	4
2.2 Multinomial Theorem	4
2.3 Expected Value and Variance	5
2.4 Covariance	6
2.5 Long Run Expected Values and Variances	6
3 Q1iii: Law of Large Numbers	7
3.1 Converging	7
4 Q2: Martingale-Roulette Investigation	8
4.1 Martingale Strategies	8
4.2 Payoff	8
4.3 Simulation	8
4.4 Conclusion	10

1 Q1i: Simulation of Probability Distributions

1.1 Graphs



No.	Dice				Maxima			Cumulative Sums			
No.	X1	X2	X3	X4	M2	M3	M4	S1	S2	S3	S4
1	4	3	1	6	4	4	6	4	7	8	14
2	1	1	5	2	1	5	5	1	2	7	9
3	4	1	3	6	4	4	6	4	5	8	14
4	6	6	2	2	6	6	6	6	12	14	16
5...	6	5	2	1	6	6	6	6	11	13	14
1000	5	2	5	6	5	5	6	5	7	12	18

Table 1: Part 1.

No.	Long Run Averages				(X-U)^2				Variance			
No.	M3	M4	S3	S4	M3	M4	S3	S4	M3	M4	S3	S4
1	4	6	8	14	0.933	0.59	6.636	0.002	0.97	0.77	2.58	0.04
2	4.5	5.5	7.5	11.5	0.001	0.054	12.788	25.402	0.03	0.23	3.58	5.04
3	4.333	5.667	7.667	12.333	0.933	0.59	6.636	0.002	0.97	0.77	2.58	0.04
4	4.75	5.75	9.25	13.25	1.069	0.59	11.724	3.842	1.03	0.77	3.42	1.96
5...	5	5.8	10	13.4	1.069	0.59	5.876	0.002	1.03	0.77	2.42	0.04
1000	4.966	5.233	10.577	14.045	0.001	0.59	2.028	15.682	0.03	0.77	1.42	3.96

Table 2: Part 2.

There are various versions for The Law of Large Numbers that basically all mean the same thing. As the sample size grows indefinitely, the mean value of the data will converge to the “expected value” for the data.

The Weak Law states that as n approaches infinity, $\varepsilon > 0$ for the probability of $|Xn - \mu| > E$ is 0 whilst the Strong Law states that for some large number N ; $n > N$ for the probability of $Xn = \mu$ is 1. The key difference between them is that the Weak Law leaves open the possibility of an infinite number of tests before convergence but the Strong Law states that convergence is achieved for all large enough n . The Strong Law implies the Weak Law due to a strong convergence guaranteeing a weak convergence but there are situations where there is a weak convergence and not a strong convergence so mathematicians have not yet been able to unify both Laws.

A proof for the weak law involves using Chebyshev’s inequality(which basically states that only $1/k^2$ of a distributions values can be outside k standard deviations from the mean).

Instead of a point estimate, an interval estimate called a confidence interval can be found. For n large the confidence interval for the unknown expected value $u = E(X)$ is given by $X(n) \pm Z(S(n)/\sqrt{n})$. The confidence interval gives the upper and lower confidence bounds for the estimate.

The correlation coefficient of two variables is defined as the covariance of the two variables divided by the product of their standard deviations. It measures the strength and direction of the relationship between two variables. Taking this into account we were able to invert the equation and find the covariance of the two variables using the known values of the variance and correlation.

$$p(x, y) = Cov(X, Y) \sigma_X \sigma_Y$$

$$p(x, y) = E[(X - \mu_x)(Y - \mu_y)] \sigma_X \sigma_Y$$

$$cov(X, Y) = \Sigma \frac{(xi - \bar{x})(yi - \bar{y})}{N}$$

No.	Long Run Variance Averages				Long Run Covar Averages	
No.	Var M3	Var M4	Var S3	Var S4	S3,M3	S4,M4
1	0.97	0.77	2.58	0.04	15.976	31.234
2	0.5	0.5	3.08	2.54	16.739	27.209
3	0.657	0.59	2.913	1.707	16.510	28.625
4	0.75	0.635	3.04	1.77	15.886	28.203
5...	0.806	0.662	2.916	1.424	15.933	28.828
1000	0.878	0.799	2.367	2.735	16.779	25.058

Table 3: Part 3.

Long Run Overall Averages			
M3	M4	S3	S4
4.966	5.232	10.576	14.04

Table 4: Part 4.

Possible values				Frequency of each value			
S3	S4	M3	M4	S3	S4	M3	M4
1	1	1	1	0	0	4	0
2	2	2	2	0	0	31	16
3	3	3	3	4	0	84	53
4	4	4	4	12	0	171	134
5	5	5	5	26	5	291	275
6	6	6	6	47	10	414	517
7	7			59	14	0	
8	8			95	25	0	
9	9			111	42	0	
10	10			128	54	0	
11	11			129	71	0	
12	12			124	98	0	
13	13			92	126	0	
14	14			69	102	0	
15	15			50	106	0	
16	16			29	86	0	
17	17			14	89	0	
18	18			1	65	0	
	19				44		
	20				23		
	21				17		
	22				10		
	23				1		
	24				0		

Table 5: Part 5.

2 Q1ii: Probability Distributions

2.1 Probability Distribution

Three dice can be rolled in a variety of 216 ways ($6*6*6$) over a probability distribution shown below. The possible combinations for each die results in varied probabilities for each number, where a greater number of combinations results in a higher probabilistic outcome, and thus over the long run, results in a variance of probabilities for each dice combination. The probability of each number rolled tends from the value of the number of combinations that can be possibly formed for that combination, with respect to the total number of combinations possible for n dice.

Sum	Dice Combinations	Combinations	Probability
3	1,1,1	1	0.005
4	1,1,2	3	0.014
5	1,2,2 1,1,3	6	0.028
6	1,1,4 1,2,3 2,2,2	10	0.046
7	1,1,5 1,2,4 1,3,3 2,2,3	15	0.069
8	1,1,6 1,2,5 1,3,4 2,2,4	21	0.097
9	1,3,5 1,4,4 1,2,6 2,2,5 3,3,3	25	0.116
10	1,3,6 1,4,5 2,2,6 2,3,5 3,3,4	27	0.125
11	1,4,6 1,5,5 2,3,6 2,4,5 3,4,4	27	0.125
12	1,5,6 2,4,6, 2,2,5 3,3,6 4,4,4	25	0.116
13	1,6,6 2,5,6 3,5,5 3,4,6	21	0.097
14	2,6,6 3,5,6 4,4,6 4,5,5	15	0.069
15	3,6,6 4,5,6 5,5,5	10	0.046
16	4,6,6 5,5,6	6	0.028
17	5,6,6	3	0.014
18	6,6,6	1	0.005
		216	1

Table 6: Probability Distribution of Combinations

2.2 Multinomial Theorem

The multinomial theorem is a *generating function* by which a probability distribution of dice sums may be calculated and expressed. The series is defined by the power of the number, where the exponent itself is a polynomial with respect to the number rolled by the die for any positive integer m and non-negative integer n.

$$(x_1 + x_2 + \dots + x_m)^n = \sum_{k_1+k_2+\dots+k_m=n} \frac{n!}{k_1!k_2!\dots k_m!} \prod_{1 \leq t \leq m} x_t^{k_t}$$

In the case of S_n where n dice of 6 sides have been rolled

$$G(S_n) = \left(\sum_{k=1}^6 x^k \right)^n = (x^1 + x^2 + x^3 + x^4 + x^5 + x^6)^n$$

Such that for multiple dice

$$\begin{aligned} G(S_1) &= x^1 + x^2 + x^3 + x^4 + x^5 + x^6 \\ G(S_2) &= (x^1 + x^2 + x^3 + x^4 + x^5 + x^6)^2 \\ &= x^2 + 2x^3 + 3x^4 + 4x^5 + 5x^6 + 6x^7 + 5x^8 + 4x^9 + 3x^{10} + 2x^{11} + x^{12} \end{aligned}$$

2.3 Expected Value and Variance

Expected value is defined as the product of given value with respect to the probabilistic outcome of that value, as a summation of all the possible values within that set for a randomly distributed variable X:

$$E[X_n] = x_n p_n \quad \text{where} \quad E[X] = \sum_{i=1}^n x_n p_n$$

Given the probability distribution, the expected value is calculated as such:

$$E[X_1] = 1 * \frac{1}{6} + 1 * \frac{2}{6} + 1 * \frac{3}{6} + 1 * \frac{4}{6} + 1 * \frac{5}{6} + 1 * \frac{6}{6} = 3.5$$

In the case of having three dice, this method can be extended to encapsulate a wider range of values over two additional dice:

$$\begin{aligned} E[X_3] = & 3 * \frac{1}{216} + 4 * \frac{3}{216} + 5 * \frac{6}{216} + 6 * \frac{10}{216} + 7 * \frac{15}{216} + 8 * \frac{21}{216} + 9 * \frac{25}{216} + 10 * \frac{27}{216} + \\ & 11 * \frac{27}{216} + 12 * \frac{25}{216} + 13 * \frac{21}{216} + 14 * \frac{15}{216} + 15 * \frac{10}{216} + 16 * \frac{6}{216} + 17 * \frac{3}{216} + 18 * \frac{1}{216} \end{aligned}$$

The rolling of dice in such a way is denoted as an independent event, as the dice rolls in the future are not dependent on previous rolls and do not affect the outcome of values, where essentially the expected outcome of the event is given by:

$$E[X + Y + Z] = E[X] + E[Y] + E[Z]$$

Given the independence of each die roll, the expected value for three rolls is given by:

$$E[X + Y + Z] = 3.5 + 3.5 + 3.5 = 10.5$$

Thus confirming the supposition by applying probabilistic arguments to the real-world. Ergo, variance of a random variable is denoted by the following:

$$\begin{aligned} \sigma^2(X) &= E[(X - \mu)^2] \\ &= E[X^2 - 2XE[X] + E(X)^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ &= E[X^2] - (E[X])^2 \end{aligned}$$

Therefore the variance for the set can be determined from the above equations results in the following using the square of the value rolled from 1-6 a total of 1 time:

$$\begin{aligned} E[X_1^2] &= 1 * \frac{1}{6} + 4 * \frac{1}{6} + 9 * \frac{1}{6} + 16 * \frac{1}{6} + 25 * \frac{1}{6} + 36 * \frac{1}{6} = 15.16 \\ \sigma^2(X_1) &= 15.16 - (3.5)^2 = 2.91 \end{aligned}$$

Given the joint probabilities via S3, S4, M3 and N4; the expected value of M3 is the sum-product of the marginal probability of M3. The same is true for the expected value of M4 with respect to M4 itself. Thus,

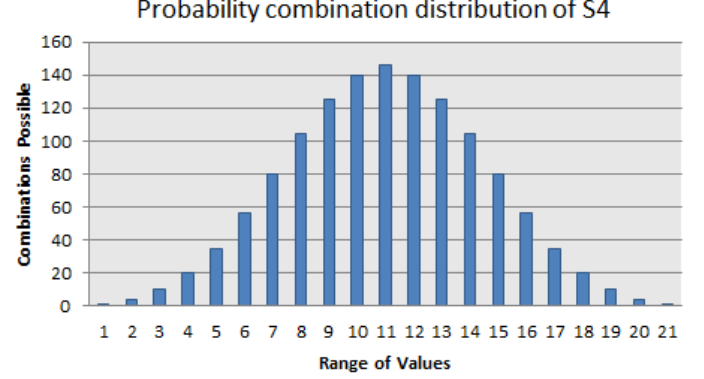
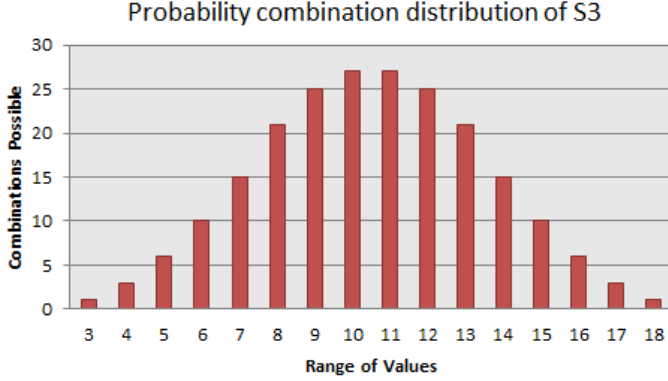
$$E[M_3] = 0.005 + 0.065 + 0.264 + 0.630 + 1.481 + 2.529 = 4.972$$

Similar to one die, calculating the sum of three dice results in a similar manner:

$$\begin{aligned} E[X_3^2] &= 9 * \frac{1}{216} + 16 * \frac{3}{216} + 49 * \frac{6}{216} + 64 * \frac{10}{216} + 81 * \frac{15}{216} + 100 * \frac{21}{216} \\ &+ 121 * \frac{27}{216} + 144 * \frac{27}{216} + 169 * \frac{21}{216} + 196 * \frac{15}{216} + 225 * \frac{10}{216} \\ &+ 256 * \frac{6}{216} + 289 * \frac{3}{216} + 324 * \frac{1}{216} = 119 \\ \sigma^2(X_3) &= 119 - (10.5^2) = 8.75 \end{aligned}$$

For each given roll, there is independence between variables. Therefore the total variance is the sum of the individual variances within the set:

$$\begin{aligned}\sigma^2(S_3) &= \sigma^2(X_1) + \sigma^2(X_2) + \sigma^2(X_3) \\ &= 2.91 + 2.91 + 2.91 = 8.73 \approx 8.75\end{aligned}$$



2.4 Covariance

Covariance in a dataset is a measure of how much two random variables change together within a given time. Covariance within the data sets may be defined as followed

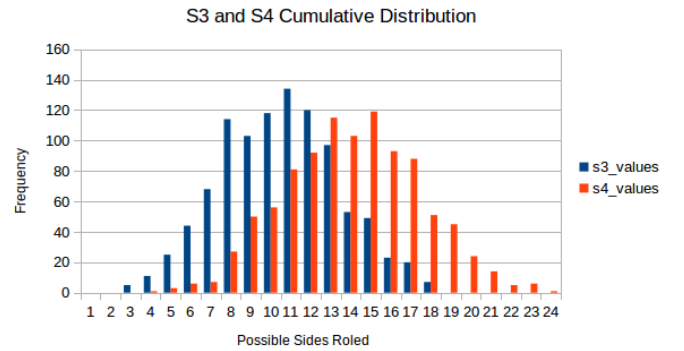
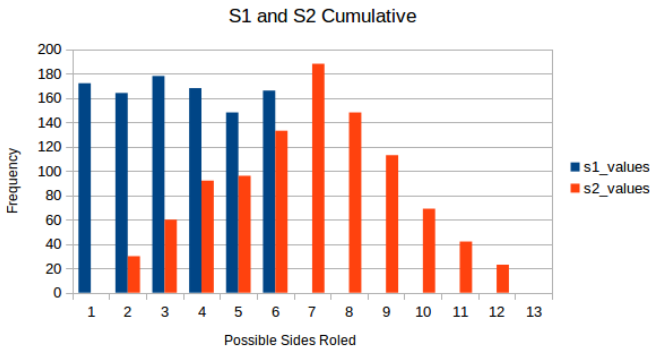
$$\begin{aligned}\sigma(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

Thus calculating $\sigma(S_3, M_3)$ for the joint probabilities of (S_3, M_3) and (S_4, M_4) respectively gives

$$\begin{aligned}\sigma(S_3, M_3) &= 54.764 - (4.972 * (3 * 3.5)) = 2.558 \\ \sigma(S_4, M_4) &= 74.589 - (5.181 * (4 * 3.5)) = 2.055\end{aligned}$$

2.5 Long Run Expected Values and Variances

Following the *Central Limit Theorem* which implies that X_1, X_2, \dots, X_n are all independent random variables, where each random variable within the set pertains to tending to the same expected value, μ , and standard deviation, σ^2 , within the distribution given that the population is sufficiently large with n values of dice summation.



The above charts are S_1 , S_2 , S_4 , and S_8 which demonstrate the effects of tending to the normal distribution curve as n samples approaches infinity. Ergo, the expected value of $E[S_k]$ is therefore shown to be $E[S_k] = K * E[S_1]$, while the expected long run variance $V[S_k]$ is shown to be $V[S_k] = K * V[S_1]$

3 Q1iii: Law of Large Numbers

3.1 Converging

There are various versions for The Law of Large Numbers that basically all mean the same thing. As the sample size grows indefinitely, the mean value of the data will converge to the “expected value” for the data. The Weak Law states that as n approaches infinity, for all $\epsilon > 0$, the probability of $|\bar{X}\bar{n} - \mu| > \epsilon$ is 0, whilst the *Strong Law* states that for some large number N ; $n > N$, the probability of $\bar{X}\bar{n} = \mu$ is 1. The key difference between them is that the Weak Law leaves open the possibility of an infinite number of tests before convergence but the Strong Law states that convergence is achieved for all large enough n . Borel also investigated this phenomenon and produced his own *Law of Large Numbers*:

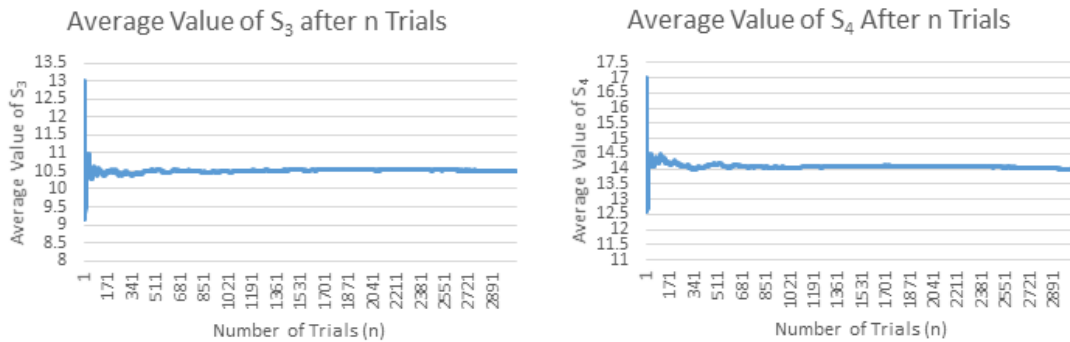
For an experiment repeated independently under identical conditions, the ratio of occurrence of any specified event approximately equals the probability of the event’s occurrence on any trial if it is repeated a large enough number of times.

From looking at section 2.3 we see that $E[S_1] = 3.5$ and it is given that $E[S_k] = k[E[S_1]]$ (see section 2.5). This means that:

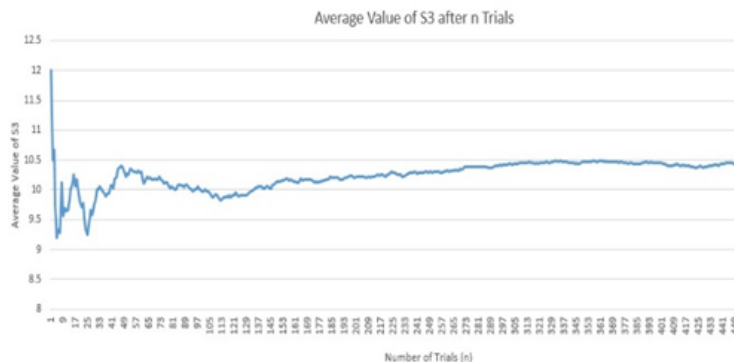
$$E[S_3] = 3[E[S_1]] = 10.5$$

$$E[S_4] = 4[E[S_1]] = 14$$

From looking at the graphs in section 1.2 it is clear that the Average Value of S_3 , after initially fluctuating from the expected value, begins to converge to 10.5. We see a similar occurrence regarding the Average value of S_4 converging to 14. Those graphs clearly illustrate the Law of Large Numbers.



What is evident is the importance of running many simulations as even after 150 trials the average value of S_3 is fluctuating back and forth around 10, as opposed to near the expected value of 10.5. A similar problem can be observed with regards to S_4 . Clearly the initial skew is due to a high value within the first few trials. As you can then see, this is eventually cancelled out with many trials and the overall mean value will converge upon the expected value. The key word is that it “eventually”, not “immediately”, converges to the expected value. It is this misinterpretation that leads to flawed gambling strategies, such as *Martingale strategy* for the roulette wheel which will be analysed in section 4.



4 Q2: Martingale-Roulette Investigation

4.1 Martingale Strategies

Roulette is a casino game of French origin that has become a staple game found in the majority of casinos worldwide. Roulette's popularity is due to the fact that it involves no player skill whatsoever, making it the go-to game of casino "newbies". However this comes with a price, the casino holds a great edge in every spin of the wheel. For this project I will focus on the European version of Roulette which holds a lesser house edge than that of American Roulette, due to the presence of only 1 zero in European Roulette compared to the 2 zeros of American Roulette.

4.2 Payoff

European Roulette	House Edge Calculation	Percentage	Expected Value
House Edge	$P(loss) * (stake\ lost) + P(win) * return$	House Edge	on \$10 Bet
Single # Betting	$\frac{36}{37} * (-1) + \frac{1}{37} * (35)$	-2.70%	-\$0.27
Outside Betting	$\frac{19}{37} * (-1) + \frac{18}{37} * (1)$	-2.70%	-\$0.27

Table 7: Martingale Strategy Applied to Roulette

As you can see this is a very foreboding prospect for gamblers as they are expected to lose money on every bet they place. Naturally gamblers will want to try and offset this house edge and the most common way of attempting to do this is through the use of betting systems to maximise their winnings or minimise their losses. As convincing as many of these betting systems appear to be, none of them can change the edge of the game being played.

One of the most well-known betting strategies is the "Martingale" betting system. Martingale originated in 18th century France where roulette was very popular at the time. Martingale involves betting on even money bets usually in roulette (Red or Black, odd or even) or craps (pass/come bets). After every loss the player doubles their bet until they win, in which case they return to their base unit bet. The idea behind Martingale is that no matter how many times the player loses, a win will net them a profit equal to their base bet. However there are two key problems facing a player that utilizes the Martingale Strategy.

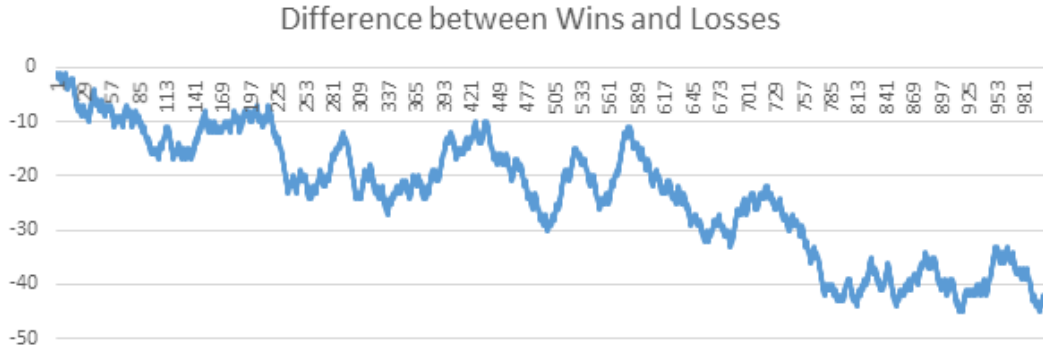
1. Martingale can only be completely effective if the player has infinite money as a long run of losses will leave the player betting astronomical amounts of money.
2. Casinos often have limits on the stakes that can be placed at any table. Then when the player runs on a streak of losses, they get to a point when the casino will not let them increase their bet any further, breaking the system.

4.3 Simulation

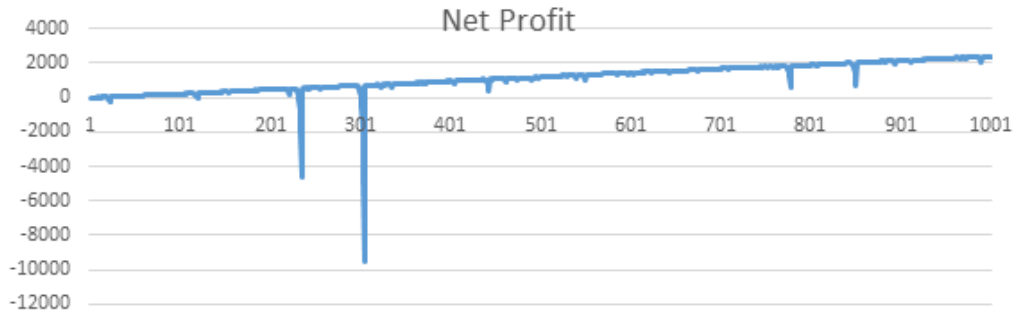
A simulation was carried out using the `RANDBETWEEN()` function to mimic the spin of the roulette wheel 1,000 times. This would equate to roughly 9 hours playing time for a gambler alone at a table.

To simplify the simulation, the bet chosen was odd or even as Excel has a useful modulo function, however we would see the same results had any other even money bet been used. A base unit bet of \$5 was used as this is often times the minimum bet accepted at a roulette table. A running count of net profit was calculated along with a tally of the streaks of wins and losses and the cumulative difference between

wins and losses.



The downfall of Martingale can be seen clearly from the results of the simulation. While the strategy exhibits a returns profile that is deceptively low risk with steady returns, there are sometimes catastrophic sequences that will cause the gambler to run into ruin. At the lowest point on the Net Profit graph, the player was down \$9,535 requiring a bet of \$10,240 to return the player into profit. This is impossible to do in reality as the casinos will only allow a maximum bet of around \$2,000 on outside bets.



This is good example of a *Taleb Distribution* or *Taleb Payoff*. The gambler usually wins a small net reward causing the illusion of a sound strategy. However the gambler's expected value still remains less than zero due to the inherent house edge in the game and the small probability that he will suffer catastrophic loss. The size of the bet increases exponentially and this, combined with the fact that streaks of consecutive losses are more common than expected, can leave the gambler broke.

Given I walk into a casino with \$1,000 and proceed to employ Martingale starting at \$5, it would only take a streak of 7 losses in a row to cause me to be unable to bet enough on the 8th bet to continue with the strategy.

$$\sum_{n=0}^7 5(2^n) = \$1,275$$

The probability of 7 losses in a row when playing European Roulette is as follows:

$$\left(\frac{19}{37}\right)^7 = 0.00942$$

Although this appears quite low, no player expects to walk into a casino and only bet one spin of the roulette wheel. The more bets you place, the more likely the player is to go bankrupt however during the game the player does not see this. Instead the player sees steady returns of his unit bet multiple times per minute. This causes the player to continue betting and eventually run into ruin. Intuitively, we can look at the results of each bet on a game of roulette as a sequence of independent and identically distributed random variables. As expected value is a linear operator, we have:

$$E(b_1 + b_2 + \dots + b_n) = E(b_1) + E(b_2) + \dots + E(b_n)$$

4.4 Conclusion

As the bets are independent in roulette, the expected value of any bet does not rely on any previous bet. As we have calculated at the start of the question, roulette holds an inherent negative expected value on every bet. Therefore our expected value will always be negative as the sum of negative numbers is always negative. From this we can conclude that there is no way to create a profitable betting strategy for roulette due to the independence of its outcomes. The best way to win at Roulette is to avoid playing in the first place.

References

- [1] Jim Kilby, Jim Fox, Anthony F. Lucas, *Casino Operation Management, 2nd Edition*, Wiley (2006).
- [2] Nassim Nicholas Taleb, *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets*, Random House Trade (2005).
- [3] Henk Tijms, *Understanding Probability, 3rd Edition*, Cambridge University Press (2012).