8

Lucky streaks

GambleAware

February 2024

Activity introduction

Quick summary

What does it mean when we wish someone good luck? Are we really hoping for luck? Are some people luckier than others?

For some, luck is considered a force that causes good or bad things to happen. However, as mathematicians, we can be sure there is no such force as luck. Games of chance rely solely on probability calculations, and it's important to know these calculations in order to check that gambling machines (such as electronic poker-machines) are operating within expected and fair margins.

However, random numbers don't always look like what we might expect random numbers to look like.

In this lesson, students will generate their own random numbers and then test their randomness. Then, using the principles of binomial probability and the normal distribution, they will test whether some gambling machines using random numbers are operating correctly.

Learning intentions

Students will:

- understand what a random number is
- understand that events will not occur evenly spaced but rather happen in clusters
- understand that 'lucky' streaks are expected and nothing extraordinary.

21st-century skills

Critical thinking

Ethical behaviour

Problem solving

Teamwork

Syllabus outcomes

Mathematics Extension 1 (Year 12)

- **ME12-5** applies appropriate statistical processes to present, analyse and interpret data
- **ME12-6** chooses and uses appropriate technology to solve problems in a range of contexts
- **ME12-7** evaluates and justifies conclusions, communicating a position clearly in appropriate mathematical forms.

Stage 6 Mathematics Syllabus Statements

Students develop an understanding of binomial distributions and associated statistical analysis methods and their use in modelling binomial events. Binomial probabilities and the binomial distribution are used to model situations where only two outcomes are possible. The use of the binomial distribution and binomial probability has many applications, including medicine and genetics.

Topic
Probability

Unit of work
Mathematics Stage 6

Time required
60 minutes

Level of teacher scaffolding
Medium – Support students in their independent calculations, facilitate team work and discussions.

Resources required

- Appendix A: Student worksheet
- Dice – 6 sided – one per student
- Whiteboard

Keywords
Gambling, betting, sports, casino, money, gaming, probability, binomial distribution, combination, mean, variance, random, normal distribution, z-score.

Teacher worksheet

Teacher preparation

Gambling can be a high-risk activity and is a priority concern for young people. Therefore, before conducting the lesson on gambling, it is recommended that teachers and parents read the Facilitator pack. The pack provides teachers and parents with essential information about gambling harm amongst young people and clarifies the nature of gambling-related behaviours and how to approach sensitive topics.

Learning intentions

Students will:

- understand what a random number is
- understand that events will not occur evenly spaced but rather happen in clusters
- understand that 'lucky' streaks are expected and nothing extraordinary.

Success criteria

Students can:

- explain that events will naturally occur in groups
- explain why not all numbers will occur equally in the short term
- identify groups of random numbers.

Teaching sequence

5 minutes - Part A: Introduction to random numbers

30 minutes - Part B: Generate and test random numbers

10 minutes - Part C: Testing 'random' machines

5 minutes - Part D: Reflection

Part A: Introduction to random numbers

Work through this resource material in the following sequence:

Step 1

Discuss:

- What does it mean when we wish someone good luck? Are we really hoping for luck?
- Are some people luckier than others?

Some students will believe that they are a lucky person because they live in a happy family in a nice house etc. We can all understand that and to different extents people in Australia are all very lucky, but we are looking more for a working definition of what luck is.

For some, luck is considered as a force that causes good or bad things to happen.

However, as mathematicians, we can be sure there is no such force as luck. Games of chance rely solely on probability calculations.

Step 2

Discuss:

- Why do sportspeople seem to have 'lucky' nights where everything goes well for them?

A sports person's performance, even though they train hard and perform to a structure, is to some extent governed by probability. They will have a natural fluctuation in performance which is nothing to do with luck.

Issues like injury, confidence, the opponent, and the performance of teammates will also contribute to an individual's performance.

Step 3

Discuss:

- What does a person mean when they say, "That's so random?"

This is often taken to mean "This is so unexpected" or "That is so strange".

It can also be used in a derogatory fashion as in, "are you going to take some random person on Twitter's word for that?"

It can also be used to represent an unknown or odd person as in, "I was online talking to randoms."

In mathematics, 'random' means a process of selection in which each item of a set has an equal chance of being chosen. For example, the outcome when we roll a six-sided dice once is random, because every number has an equal chance of appearing.

Random numbers are used for a wide variety of tasks – gaming machines, gambling games, encryption, finding the area of a difficult shape or even checking for fraud in a tax return.

Part B: Generate and test random numbers

Step 1

Before beginning this task, discuss:

- What do you expect from a set of random numbers?

Step 2

Ask students to create a set of random numbers from their head. Students write down a series of 50 digits between 1 and 6.

Step 3

Now ask students to create a set of random numbers using a six-sided die. Students roll the dice 50 times and record the number that occurs each time or use a dice simulator like calculator.net/dice-roller.html (and change the number of dice to 100). Note the frequency of each outcome.

So for both steps, students should have generated a list of 50 numbers like this: 2,4,1,4,3,1,1,6,3,4,5,6 ....

Step 4

Compare the 2 data sets.

Discuss:

- Which set of figures appears to be more random?
- Do either set of figures match our expectation of what a 'random' set of numbers should look like? Why or why not?
- How would you know if either set of figures is truly random?

You might hear:

- The dice gave me a lot of 6's, so I don't think that is very random.
- The dice gave me 3 5's in a row, so I don't think that is very random.
- The dice didn't give me many 1's, so I don't think that is very random.
- There should be an equal number of each number. Since we had 50 numbers there should be about 8 of each.
- There should be no patterns.

Part B: Generate and test random numbers

You are looking for things like:

- The dice is more random because it is not influenced by us.

So, how can we test for randomness? What would a set of random numbers look like?

Students may suggest that looking at the frequency of each number is a good test for randomness. (This is only true for very large samples. There will be natural variation in the frequency of each digit.)

Step 5

Performing tests on the results.

We do not want to get bogged down here and there is the potential to waste a lot of time on this section. To save time divide the class into 3 so that 3 tests can be done simultaneously.

Test 1

Students in this group will find the frequencies of each number in both samples of 50 numbers they have personally generated.

Once they have found the frequencies, have them add their figures to a table. This table can be found in Appendix A: Student worksheet.

	Frequency for the made up numbers						Frequency for the dice rolling					
Student name	1	2	3	4	5	6	1	2	3	4	5	6
Anna												
Beatrice												
Carli												
Dave												
Edward												
etc												
Total												

Once all students have added their data, calculate the total frequency of each number for each method.

By adding the totals, the variation may even out a bit for the die rolls due to the law of large numbers, but may not for the made-up list.

Part B: Generate and test random numbers

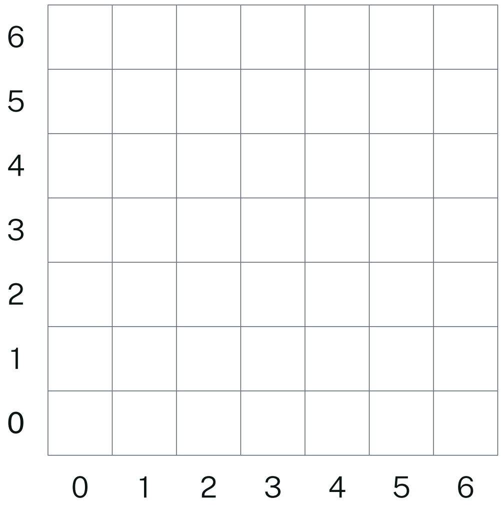
Test 2

Students in this group will use the dice generated list of numbers.

On a grid like the 1 below (also found in Appendix A: Student worksheet), plot each adjacent pair of numbers as a coordinate point.

For example, if the numbers were 2, 4, 6, 2, 8 etc you would plot the points (2,4), (4,6), (6,2), (2,8). Make it clear to students that you are using each number twice, once as the Y coordinate in 1 plot point, and then as the X coordinate in the next plot point.

For the last number, students will need to use the first number as the Y coordinate.



Test 3

Students in this group will do the same as above, but using the numbers that they made up.

Step 6

Ask each group to report on what they notice about their graph.

Typical observations might include:

- There is just a random scattering of points.
- In the made-up data we didn't get any double numbers.

For the made-up figures:

- There are more plotted points in the top left of the grid than the bottom right.
- Some pairs were much more common than others, but others appeared only once or not at all.
- Even pairs like (2,6) and (4,2) were more common.

This is because our brain is used to counting, so we are more likely to think of ascending digits than descending ones when making up numbers.

Note: The intention of this work is to highlight that in random samples there will always be runs of numbers. Knowing which number will be involved, when the run starts and ends, is impossible at the time. You will often hear people say things like 'it was my lucky night', but really, there is no such thing as luck, only probability.

Part C: Testing ‘random’ machines

Step 1

Okay, enough magic, back to the numbers.

Poker machines are designed on the basis of presenting random numbers or results. When poker machines were manual - imagine a dice - it was simple to guarantee that was happening in the way they were designed. But what happens when the machines switched to digital, where we are relying on computers to generate truly random lists of numbers.

As we have seen in the last part, streaks do occur naturally. So how does a machine operator know if their machine is malfunctioning or has been tampered with? A quick search online will bring up cases of operators deliberately setting their machines to return less and players trying to use things like magnets to disrupt the working of a machine, so there does need to be some way of checking.

To understand this, we need to look at the binomial probability and the normal distribution.

Step 2

As the term implies, bi means 2. Binomial probability is concerned with events where there are only 2 possible outcomes: win and loss, or as mathematicians say, win and not win.

With 80% returned in prize money for a poker machine, the chances of winning are 0.8. The odds of not winning are therefore 0.2.

Step 3

Ask students to calculate the chances of winning 3 out of 5 games, which means we need 3 wins and 2 losses. Since ‘and’ means multiply in probability, students might attempt the following.

$$(0.8)^3(0.2)^2$$

However, this calculates the chance of WWWLL in that order only.
We also need to consider WLWWL, WWLLW etc.

Step 4

Remind students that by using combinatorics, we know that the number of ways of different combinations of 3 wins and 2 losses can be found using ${}^5C_3$. (${}^5C_2$ gives the same value)

$${}^5C_3 = \frac{5!}{3!2!} = 10$$

This means there are ten unique ways of arranging 3 W’s and 2 L’s.

Remember: ! means factorial. Example 4! = 4 × 3 × 2 × 1

The chance of winning 3 out of 5 rounds of a poker machine is therefore:

$$10 \times (0.8)^3(0.2)^2 = 0.2048$$

The general rule is:

$$(r \text{ wins from } n \text{ games}) = nC_r \quad p^n \,.\, (1-p)^{(n-r)}$$

Step 5

Ask students to find the probability of winning:

a. 2 games out of 4

b. 1 game out of 3

c. 3 games out of 6

Answers

a. $Pr(2 \text{ wins from } 4 \text{ games}) = 4C_2 .\ 0.8^2 \,.\, (1-0.8)^{(4-2)}$ = 6 × 0.64 × 0.04 = 0.1536

b. $Pr(1 \text{ win from } 3 \text{ games}) = 3C_1 .\ 0.8^1 \,.\, (1-0.8)^{(3-1)}$ = 3 × 0.8 × 0.04 = 0.096

c. $Pr(3 \text{ wins from } 6 \text{ games}) = 6C_3 .\ 0.8^3 \,.\, (1-0.8)^{(6-3)}$ = 20 × 0.512 × 0.008 = 0.08192

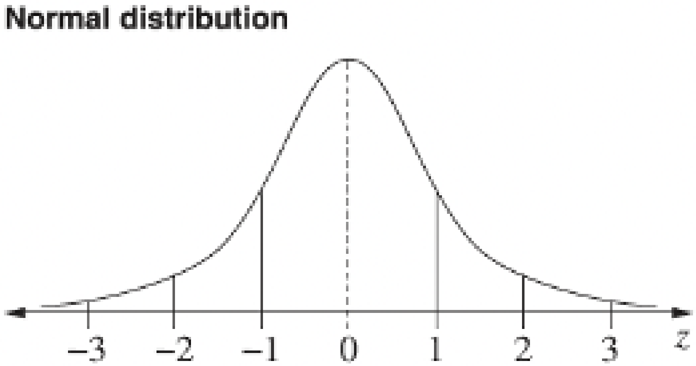
Step 6

However, for larger values of *n* using binomial probability becomes tedious and time consuming, to calculate the probability of each outcome.

Assuming the data is completely random, we can use the normal distribution to test to see if an individual machine is operating within the guidelines for its operation.

We know from studying the normal distribution that 95% of data lies within z-scores of -2 and +2 ie within 2 standard deviations below and above the mean. We assume that a machine that is operating normally will pay out at a level within these values. If a machine is paying out at a rate with a z-score less than -2 there is a good chance it may have been tampered with by the owner to reduce the chance of winning below the expected level and if it is paying out at a z-score of above +2 it may have been tampered with by the customer to increase the chance of winning. By calculating the mean and standard deviation of the individual data set, we can calculate the payout levels equivalent to a z-score of -2 and +2.

Part C: Testing 'random' machines



- approximately 68% of scores have z-scores between -1 and 1
- approximately 95% of scores have z-scores between -2 and 2
- approximately 99.7% of scores have z-scores between -3 and 3

(Source: NESA Mathematics Advanced, Extension 1 and Extension 2 Reference Sheet)

95% of normal data lies within 2 standard deviations of the mean.

Remember p = the probability of a win. (For a poker machine this is 0.8)

For normal data the mean, $\mu = np$

And the standard deviation, $\sigma = \sqrt{variance}$ where $variance = np(1-p)$

Step 7

Independently, students undertake the calculations on Appendix A: Student worksheet to determine whether a machine is operating within a z-score of -2 and +2.

You may wish to work through the first example as a class.

The answers have been included for you below:

Question 1
Test to see if a machine which has returned 85% of the money invested over a 200-game span is operating within a z-score of -2 and +2.

mean, $\mu = 200 \times 0.8 = 160$

standard deviation, $\sigma = \sqrt{200(0.8(1-0.8)}$ = 5.7 to 1 decimal place

The machine in question has returned 85% of 200 = 170.

The question, therefore, is: is this within a z-score of -2 and +2

These limits are $\mu \pm 2\sigma = 160 \pm 2 \times 5.7 = 148.6$ up to 171.4

Our machine has returned 170 which is within normal operating guidelines.

Part C: Testing 'random' machines

Question 2

Test to see if a machine which has returned 85% of the money invested over a 500-game span is operating within a z-score of -2 and +2.

mean, $\mu = 500 \times 0.8 = 400$

standard deviation, $\sigma = \sqrt{500(0.8)(1-0.8)} = 8.9$ to 1 decimal place

The machine in question has returned 85% of 500 = 425.

The limits for z-scores of -2 and +2 are $\mu \pm 2\sigma = 400 \pm 2 \times 8.9 = 382.2$ *up to* 417.8

Our machine has returned 425 which is not within the normal operating guidelines and so the machine may well have been tampered with.

Question 3

Test to see if a machine which has returned 82% of the money invested over a 500-game span is operating within the a z-score of -2 and +2.

mean, $\mu = 500 \times 0.8 = 400$

standard deviation, $\sigma = \sqrt{500(0.8)(1-0.8)} = 8.9$ to 1 decimal place

The machine in question has returned 82% of 500 = 410.

The z-score limits translate into returns of $\mu \pm 2\sigma = 400 \pm 2 \times 8.9 = 382.2$ *up to* 417.8

Our machine has returned 410 which is within the normal operating guidelines.

Question 4

Roulette is a casino game played by betting on the result of a ball being dropped onto a spinning wheel and randomly settling on a number between 0 and 36. 0 is a win for the operator and 1-36 is a win for the players. Calculate the z-score for a roulette wheel that has returned 97% of the money invested over a 1000-game span and determine if it is operating within the normal limits (that is, in the central 95%).

Since there are 37 possible outcomes, 0 to 36, and the only time the operator wins is when a zero is spun the probability of a win for the player is, $p = \frac{36}{37}$ (assuming they bet on every permitted outcome)

mean, $\mu = 1000 \times \frac{36}{37} = 973$

standard deviation, $\sigma = \sqrt{1000(\frac{36}{37}(1-\frac{36}{37})} = 5.1$ to 1 decimal place

The roulette wheel has returned 97% of 1000 = 970.

$$z = \frac{x - \mu}{\sigma}$$

The z-score for the roulette wheel is:

$$z = \frac{970 - 973}{5.1} = -0.59$$

This is well within the accepted range of z-scores between -2 and +2 to cover 95% of the data, so the machine is operating within normal parameters.

Part D: Reflection

Summarise what we have learned through this lesson.

- If events are occurring randomly the data will not be evenly spread and you can expect there to be clusters.
- Events in the gaming world are controlled by long term statistics and probability.

Questions to ponder

- **Random playlists were replaced by shuffle playlists because people did not like the random lists because the same song occasionally happened twice in a row and others never seemed to play. Explain how the shuffle list could be arranged so these problems no longer occur.**
  The shuffle list could still be random with the extra rule that once a song that has been played is excluded for the next 20 plays. Alternatively, once a song is played it is excluded until all songs have been played.

- **Why do regulating authorities, as well as the operator themselves, need to keep a check on whether gaming machines are operating within agreed limits?**
  This allows both the operator and regulator to check the machines are retuning the correct amount of money in prizes. This helps the operator stay within the guidelines by finding faulty machines early (and it also stops any unscrupulous operator cheating).

Extension: Benford's Law (up to 15 minutes)

Step 1

The following set of figures is a section of a real bank statement. (The names have been changed but not the figures). If we just treat the numbers as individual digits, do you think the first 2 digits of each number will be random numbers? We will consider only the non-zero digits.

Deposits	Description	Amount
04/9/2020	Merchant transfer	$5,442.23
06/9/2020	Merchant transfer	$2,482.00
07/9/2020	Merchant transfer	$92,379.04
08/9/2020	Merchant transfer	$3,558.82
09/9/2020	Merchant transfer	$9,859.04
09/9/2020	Merchant transfer	$137.94
11/9/2020	Merchant transfer	$628.05
12/9/2020	Merchant transfer	$6,183.89
14/9/2020	Merchant transfer	$1,315.51
16/9/2020	Merchant transfer	$21,060.29
Electronic credits		
04/9/2020	D.S. Smith	$148.87
05/9/2020	L.Hunter	$107,532.01
07/9/2020	T. Chaiyasit	$45,450.81
07/9/2020	T. Carroll	$11,793.22
09/9/2020	J. Barrendale	$127,765.64
10/9/2020	H. Morgan	$1,324.97
14/9/2020	P. Saengrod	$136,072.12
15/9/2020	L.Thompson	$62,637.02
15/9/2020	G.Norman	$121,390.12

Answer:

Students may say that this is a large group of figures so should be random. Others may wonder why the zero isn't included. (Because the first digit cannot be zero in any whole numbers such as bank account figures, so including a zero, which would never occur, would skew the frequency)

Step 2

Independently, students perform a statistical check on the frequency of the non-zero digits by counting the frequency of each of the first 2 digits.

Answer:

Digit	Frequency
1	12
2	7
3	5
4	4
5	3
6	3
7	0
8	1
9	2

Step 3

Discuss:

- **What do you notice about the figures?**
  This is quite a small data set, only 37 non-zero figures, so we would need a much larger data set to be sure, however, a clear pattern has already emerged with over 50% of the figures being a 1 ,2 or 3.

Step 4

Explain the emerging pattern.

All figures that are counted are not random numbers.

For a 2 to occur a 1 must have preceded it. For a 3 to occur it has been preceded by 1 and 2. For a 9 it has been preceded by 1, 2, 3,...7, 8 which makes each successive digit less likely than the previous one.

Step 5

The actual distribution of numbers is governed by some complex mathematics called Benford's Law. The following graph shows the distribution of leading digits for figures obtained through any counting process, as calculated using Benford's law.

Part D: Reflection

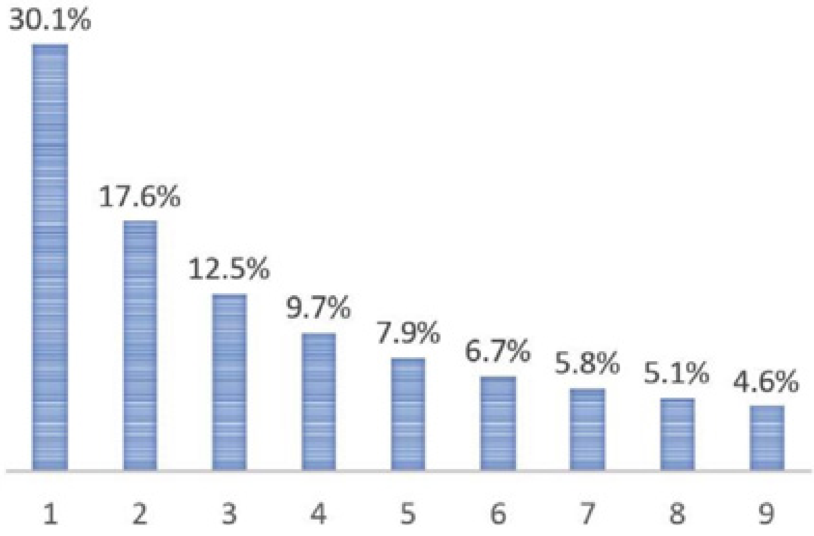


Image sourced from: researchgate.net/figure/Benfords-law-for-the-first-digit_fig1_332699906

Most people do not know about this however the results are used by large companies and agencies such as the Internal Revenue Service (the US taxation department) to find falsified financial records. The Australian Taxation Office makes no comment on whether they use this method or not. They respond only with the fact that a variety of statistical checks are done on tax returns.

Step 6

- **Explain how could Benford's Law be used to check for falsified financial records?**
  By looking at the frequency of the leading digits contained in the report and comparing it to the expected outcome, falsified documents are easy to spot. Anyone unaware of Benford's Law would most probably falsify the report by having a more even spread of digits.

Teacher reflection

Take this opportunity to reflect on your own teaching:

What did you learn about your teaching today?

What worked well?

What didn't work so well?

What would you share?

Where to next?

How are you going to get there?

Appendix A: Student worksheet

Task 1

From your head, generate a set of 50 random numbers between 1 and 6.

Task 2

Using a 6-sided die, generate a set of 50 random numbers.

Task 3

Test 1
Find the frequencies of each number in both samples of 50 numbers you have personally generated.

Frequency for the made up numbers					
1	2	3	4	5	6

Frequency for the dice rolling					
1	2	3	4	5	6

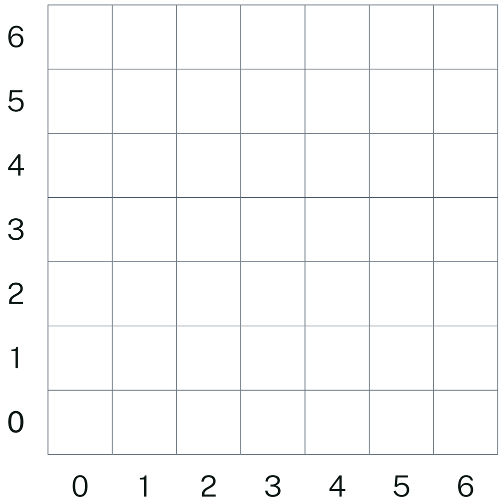
Add your results to the table on the whiteboard.

Test 2

Using your list of **die generated numbers**, plot each pair of numbers as a coordinate point.

For example, if the numbers were 2, 4, 6, 2, 8 etc you would plot the points (2,4), (4,6), (6,2), (2,8). You are using each number twice, once as the Y coordinate in one plot point, and then as the X coordinate in the next plot point.

For the last number, students will need to use the first number as the Y coordinate.

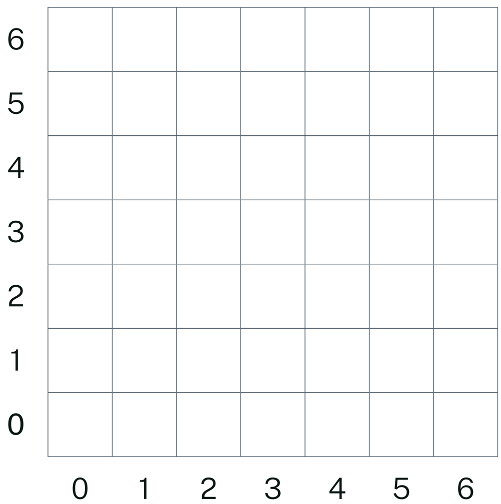


Test 3

Using your list of **randomly generated numbers from your own head**, plot each pair of numbers as a coordinate point.

For example, if the numbers were 2, 4, 6, 2, 8 etc you would plot the points (2,4), (4,6), (6,2), (2,8). You are using each number twice, once as the Y coordinate in one plot point, and then as the X coordinate in the next plot point.

For the last number, students will need to use the first number as the Y coordinate.



Task 4

Question 1

Test to see if a machine which has returned 85% of the money invested over a 200-game span is operating within the 95% confidence limits.

Question 2

Test to see if a machine which has returned 85% of the money invested over a 500-game span is operating within the 95% confidence limits.

Question 3

Test to see if a machine which has returned 82% of the money invested over a 500-game span is operating within the 95% confidence limits.

Question 4

Roulette is a casino game played by betting on the result of a ball being dropped onto a spinning wheel and randomly settling on a number between 0 and 36. 0 is a win for the operator and 1-36 is a win for the players. Test to see if a roulette wheel is operating within the 95% confidence limits if it has returned 97% of the money invested over a 1000-game span.