CASE STUDY

# Lottery Strategy: An activity for new undergraduate students to introduce and reinforce introductory statistical concepts

Salim Neil Khan, School of Computing, Electronics and Mathematics, Coventry University, Coventry, UK. Email: aa9803@coventry.ac.uk

## Abstract

Most university mathematics courses involve the students studying a statistics module in their first year. However, depending on which modules they took at A-level, they arrive at university with varying degrees of interest and ability in statistics. This article presents a classroom activity that introduces and reinforces introductory probability concepts to help prepare and engage the students for the statistics that they will encounter on their course. In the activity, the students consider and contrast two different strategies for selecting numbers for a lottery, in order to conclude which is best under which circumstances. It comprises a mixture of experimentation using a lottery machine, analysis using probability theory, and simulation using computers.

**Keywords:** Practical Activity, Lottery, Probability, Statistics, Simulation.

## 1. Background

At Coventry University (CU), as with many other universities, the students on the various mathematical related degrees all need to take a statistics module in their first year. However, when studying mathematics A-level, some may have taken mechanics and / or decision maths modules and hence encountered very little statistics. According to Cole (2015), of those students in the UK who sat the Edexcel exam board mathematics A-level in 2013, 79% had chosen the Statistics 1 module, with only 19% also choosing Statistics 2; thus 21% didn't study any statistics. Indeed some may be particularly averse to the subject. Hence, at CU an activity was introduced in induction week to help prepare and engage the students for the statistics that lay ahead. It needed to be fun and interactive for those who lacked interest in the subject and also for those who had met some of the topics before. The aim was to introduce and reinforce concepts such as combinatorics, mutual exclusivity, probability distributions, expected values, variances, tree diagrams, and conditional probability, and also create opportunities for further discussion on decision trees and utility functions which they would meet in another module.

## 2. Activity

### 1.1. Introduction

At the beginning, it is always a good idea to obtain an indication of the prior statistical experience of the students, as this can vary from year to year. This will, to a certain extent, help to determine the prospective pace of the session and detail of the content. A simple show of hands for those who have done the Statistics 1 or Statistics 2 A-level modules will suffice.

The activity starts with asking the students if they or any other family members play the lottery, if they have any particular strategies for picking the numbers, and whether they have won anything. This can lead to a brief discussion on different people's strategies and their respective merits. For example choosing combinations such as 1, 2, 3, 4, 5, 6, or numbers in the central column of the ticket are likely to result in lower payouts as they are much more popular (Cox, Daniell and Nicole, 1998).

Using PowerPoint or some other display method, the following format for a simple '15 ball lottery' is then introduced:

- 3 balls are drawn from 15;
- It costs £1 a play to pick 3 numbers;
- If 2 balls match you win £5, if 3 match you win £180.

## 1.2. Experimentation Using a Lottery Machine

The class should be split nominally into groups of four. Clearly it is highly unlikely that the number of students is an exact multiple of four, in which case some 'imaginary friends' are added to the groups. It should be ensured that there are an even number of groups. The usual class size at CU for this activity has been around thirty-five students, so this means that there are usually around ten groups.

Each group (of four) are then told that they have four plays of the lottery i.e. they need to make four selections of three numbers. Half of the groups are told to follow Strategy 1 (S1) and the other half to follow Strategy 2 (S2) as follows:

- S1- Pick 12 different numbers e.g. (1,2 ,15), (3,6,10), (4,7,14), (9,12,13);
- S2- Pick 4 numbers and repeat them e.g. (2,4,7) (2,4,10) (2,7, 10), (4,7,10).

They discuss within their groups and record their selections. Mini whiteboards work well for this. It is important to circulate during this to ensure that the correct directions have been followed, as invariably there will be some students who need clarification.

They are then asked which of the two strategies they think will be the most profitable. The most common reply has been that they are equally profitable. I have then responded by predicting that S1 will win the most money in total.

The lottery draw then takes place. At CU we purchased a bingo machine for around £20 (they are available from many toy outlets), but a simple bag will do.

Figure 1. Lottery Machine and Visualiser

A visualiser was used to transmit the action to screens around the classroom; this is recommended in a large room for maximum involvement. Additionally a lively bingo style commentary can enhance engagement.

When the balls have been drawn, the total winnings for all the groups using S1 are compared with the total winnings for all the groups using S2. Usually S1 will come out on top, and the students will be keen to know how I successfully predicted the outcome. Note that S1 doesn't always fare better than S2 here- I have run this activity six times and S1 has 'won' on five of them. If the prediction isn't correct and S2 wins, then they will still be keen to know why it was predicted that S1 would win.

### *1.3. Analysis Using Probability Theory*

The various calculations for the outcomes and their respective probabilities are then shown as follows:

Strategy 1:

$$\mathrm{P(Match\ 3\ Balls)} = \frac{1}{\binom{15}{3}} \times 4 = \frac{4}{455} \qquad \text{Win £180.}$$

$\binom{15}{3}$ is the number of possible combinations for drawing the 3 balls. It is multiplied by 4 because the four plays are mutually exclusive. This gives an opportunity to introduce to or remind the students of the concepts of combinatorics and mutual exclusivity. Similarly,

$$\mathrm{P(Match\ 2\ Balls)} = \frac{\binom{3}{2}\times\binom{12}{1}}{\binom{15}{3}} \times 4 = \frac{144}{455} \qquad \text{Win £5.}$$

$\binom{3}{2}$ is the number of ways that two of the selected numbers in a play could match the 3 drawn balls, and $\binom{12}{1}$ is the number of ways that the remaining selected number in a play could match the 12 balls not drawn.

Strategy 2:

$$\mathrm{P(Match\ 3\ Balls)} = \frac{\binom{4}{3}}{\binom{15}{3}} = \frac{4}{455} \qquad \text{Win £180 + £5 x 3 =£195.}$$

$\binom{4}{3}$ is the number of ways that the 3 drawn balls could match the four individual numbers chosen over the four plays.

The higher win amount has often been a surprise for many students. Because the numbers are repeated, if one of the plays matches 3 balls, then the other three plays in the group will all match 2 balls.

$$\mathrm{P(Match\ 2\ Balls)} = \frac{\binom{4}{2}\times\binom{11}{1}}{\binom{15}{3}} = \frac{66}{455} \qquad \text{Win £10.}$$

Again, if one of the plays matches 2 balls, another play will match 2 balls.

The figures are then collated, introducing the idea of probability distributions. These are shown in Table 1 and Table 2 below.

Table 1. Strategy 1 Theoretical Outcomes with Probabilities

| **Group Winnings (£)** | **0** | **5** | **180** |
|---|---|---|---|
| p | $\frac{307}{455}$ | $\frac{144}{455}$ | $\frac{4}{455}$ |

Table 2. Strategy 2 Theoretical Outcomes with Probabilities

| **Group Winnings (£)** | **0** | **10** | **195** |
|---|---|---|---|
| p | $\frac{385}{455}$ | $\frac{66}{455}$ | $\frac{4}{455}$ |

The students are then asked to recall the formula for expected value, $E(X) = \Sigma xp$, and consequently calculate the expected values for the two distributions as follows;

$$\mathrm{E(S1\ Winnings)} = 5 \times \frac{144}{455} + 180 \times \frac{4}{455} = £3.16.$$

$$\mathrm{E(S2\ Winnings)} = 10 \times \frac{66}{455} + 195 \times \frac{4}{455} = £3.16.$$

Thus both strategies expect to win the same on average, backing up what many students thought earlier. However, the key is that the two distributions have different variances. Recalling the formula for Variance, $Var(X) = \Sigma x^2 p - [E(X)]^2]$,

$$\mathrm{Var(S1\ Winnings)} = 5^2 \times \frac{144}{455} + 180^2 \times \frac{4}{455} - 3.16^2 = 283.$$

$$\mathrm{Var}(S2\ \mathrm{Winnings}) = 10^2 \times \frac{66}{455} + 195^2 \times \frac{4}{455} - 3.16^2 = 339.$$

This tells us that the amount won using S2 is more variable than with S1.

The values from Table 1 and Table 2 can be represented on a tree diagram, as shown in Figure 2.

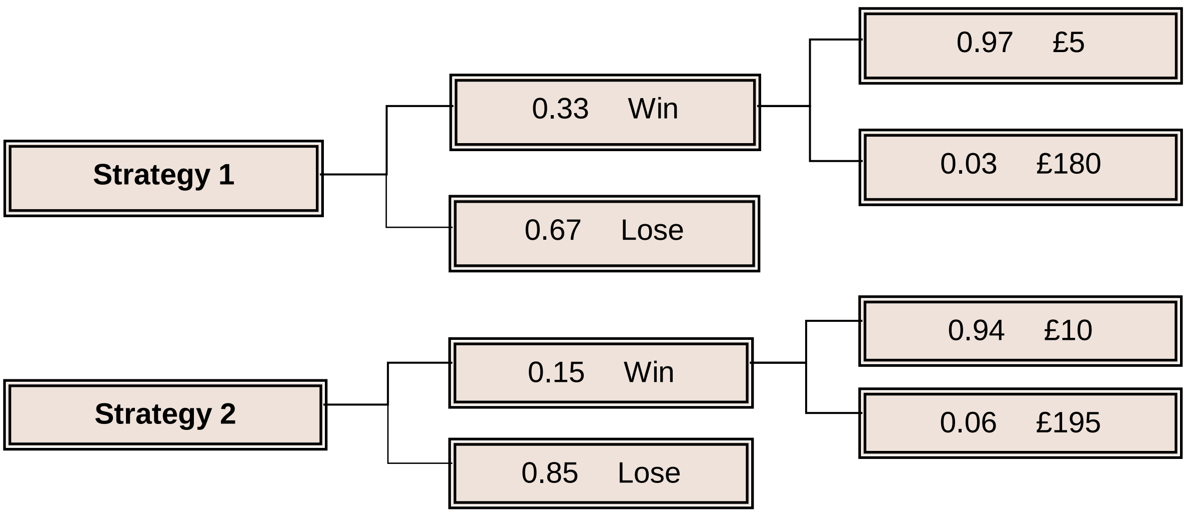


Figure 2. Outcomes and Associated Probabilities for Each Strategy

All of the students will have met probability trees at GCSE Maths. At CU this was a good opportunity to introduce the concept of a decision tree, which the students would meet in another of their modules.

For S1, the value of 0.33 above is obtained by adding $\frac{144}{455}$ and $\frac{4}{455}$ from Table 1.

0.97 is the probability of winning £5 given that you have won something. This is an example of conditional probability which can be explained intuitively as

$$P(£5|Win) = \frac{P(Win\,£5)}{P(Win)} = \frac{144}{144+4}.$$

It could also be pointed out that it is actually using Bayes' Theorem, $P(A|B) = P(A \cap B)/P(B)$.

The values for S2 in the tree are calculated similarly.

From the tree diagram it can be seen that with S1 you are more likely to win than with S2. However, if you win, you will expect to win more with S2 than with S1. Hence which strategy you choose should depend on your attitude to risk. If you want a higher chance of winning something you should choose S1. If you want a smaller chance of winning something, but a higher chance of winning a large amount you should choose S2. This could be phrased as;

- If you are Risk Averse, choose Strategy 1 i.e. select different numbers;
- If you are a Risk Seeker, choose Strategy 2 i.e. repeat the numbers.

This can lead onto discussions of attitudes to risk and some possible factors. For example, Powell and Ansic (1997) contend that females are less risk seeking than males in financial decision making. In turn, the idea of a utility function can be touched upon, which the students at CU will meet in another module.

Going back to why it was predicted that S1 would usually come out on top in this particular session, it is due to the fact that we have such a small sample size e.g. only five groups applying each strategy. As the overall chance of winning anything with S2 is so low, in the short term S1 will win more often than S2 (although of course this is not guaranteed!).

### 1.4. Computer Simulation

Having discussed the shortcomings of drawing conclusions based on small samples, the idea of using a computer to simulate a large sample of lottery draws is introduced. Hence a spreadsheet was created (shown in Figure 3) for the students to do this. (This Excel file is available upon request from the email address at the start of the article).

At CU the session is conducted in a computer lab so each student individually generates their own set of results. They randomly select four sets of numbers using Strategy 1 and four sets of numbers using Strategy 2 by clicking on the 'Select Numbers' tabs. They can then simulate as many draws of the lottery as they want by continually clicking on the 'Do Lottery Draw' tab. The total winnings for each strategy are displayed in the 'Running Total' boxes, and the average winnings are shown on the graph. In my experience, the students have found the interface user friendly and hence no clarification on what to do has been needed. However the lecturer should circulate during this part to ensure the students think about what the values and the graph are telling them.

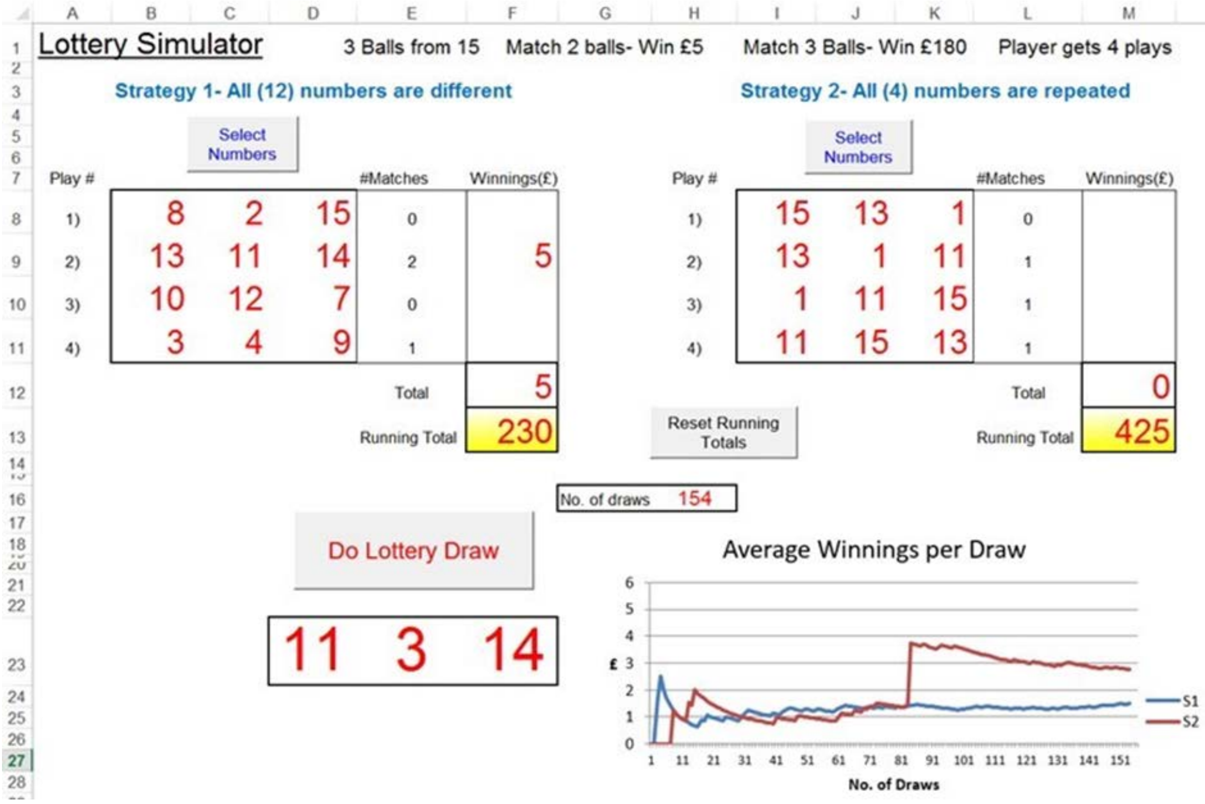


Figure 3. Simulation of Strategy 1 and Strategy 2

Once finished, the findings of the students are shared and discussed. From the graph, it is often seen that both strategies' winnings will eventually tend towards the expected value of £3.16 calculated earlier theoretically. Also, once the lines have settled down (after the first 50 plays say), the fluctuations in the S2 line tend to be larger than the fluctuations in the S1 line supporting the theory that S2 will give a higher variance.

## *1.5. Using the Lottery Machine Again*

To finish off, armed with all their new found knowledge, the students now play another lottery, with the chance of winning an actual prize! The format is different from earlier, as follows:

- 2 balls are drawn from 10;
- £1 a play to pick 2 numbers;
- 1 match wins £1, 2 matches wins £16.

For this the students should be split into groups of 3, so they get 3 plays per group.

They can choose any strategy they want- S1, S2, divine inspiration, anything. (For extra information, the lecturer could also choose to display the probability distributions for S1 and S2 winnings in this version of the lottery). The group that wins the most money wins the prize (a few sweets will normally be enough motivation).

As before, the lottery machine and visualiser are used, and judicious pauses can build the suspense. Once the 2 balls have been drawn, the prize is given to the group with the highest total winnings. If there is a tie, another draw is conducted just with those winning groups. The winning group are then asked what strategy they used. Interestingly, in this case, S2 would probably have been best as the aim is not just to win something, but to win the most.

### *1.6. Wrapping Up*

The findings are then summarised, e.g. Risk seekers should repeat the numbers; the risk averse should select different numbers. Also the students are reminded of the statistical topics / terminology encountered.

And of course, it is important for the lecturer to point out that they are not condoning gambling. Clearly it can be seen from the first lottery activity that for an outlay of £4 they only expect to get £3.16 back! An expected loss is typical in the playing of a lottery. Indeed in the UK National Lottery you can only expect a 48% return (The National Lottery, 2015). So it should really only be played for fun, and if affordable. Nevertheless, I always mention to the students that if they do play the lottery and win a million, I would be happy to receive half the money due to the expert advice I've given them during the session!

## 3. Adaptability

### *3.1. Varying the Parameters*

The activity could be repeated with different numbers of balls. If so, the associated winning monetary amounts can be calculated as follows:

Let n be the number of balls in the draw, d be the number of balls drawn (and hence the number of selections per play), and r be the % rate of return (i.e. the % of stake money paid out in winnings).

Let us assume that there is a jackpot prize, J, for matching d balls, and a second prize, S, for matching d-1 balls. Considering one play, using the same reasoning as for Section 2.3;

$$\mathrm{P(Winning\ J)} = \frac{1}{\binom{n}{d}}, \quad \mathrm{P(Winning\ S)} = \frac{\binom{d}{d-1}\times\binom{n-d}{1}}{\binom{n}{d}} = \frac{\mathrm{d(n-d)}}{\binom{n}{d}}.$$

It can be seen that winning S is $\mathrm{d(n-d)}$ times more likely than winning J. Hence for a 'fair' game it would not be unreasonable to set J at $\mathrm{d(n-d)}$ times the value of S. This would then result in the expected total payouts in second prizes equalling the expected total payouts in jackpot prizes.

Assuming that it costs £1 for a play, then the expected total amount paid out in prizes per play is $£r/100$. Thus the expected amount paid out in jackpots per play is $£r/200$. As $\mathrm{P(Winning\ J)}$ is $1/\binom{n}{d}$, it follows that

$$J = £\frac{r\binom{n}{d}}{200}.$$

Hence,

$$S = £\frac{r\binom{n}{d}}{200\mathrm{d(n-d)}}.$$

S and J should then be rounded to an appropriate level of precision e.g. the nearest £1 or £5 etc.

Illustrating this for the values for the first draw described in Section 2.2 earlier:

$n = 15$ and d= $3$. I had set r = 80%.

So, $J = \frac{80\binom{15}{3}}{200} = 182, \quad S = \frac{80\binom{15}{3}}{200\times 3(15-3)} = 5.05$, which were rounded to £180 and £5.

### 3.2. *Different Audiences*

As outlined, at Coventry University this activity has been used with incoming first year mathematics undergraduates. Also it has been used as an outreach activity for local school students in year 12 currently studying maths A-level. The idea is to give them an idea of what a university lesson might be like, to get them to think above and beyond a topic that they are currently studying, and to engage them so that they are encouraged to study in Higher Education. For those students, in the activity I tend to go into less detail on the theory, as overall they tend not be quite as confident in the subject matter as the new undergraduates. For both types of cohort, I have found that the feedback has been overwhelmingly favourable, with many students commenting that they have enjoyed the session. Additionally, I consider that similarly by removing some of the probability theory, the session could be adapted for non-mathematicians for whom some of the ideas regarding risk could be relevant and interesting e.g. finance students, psychologists.

## 4. References

Cole, J.S., 2015. A-level Mathematics Options – Views of Secondary-level Teachers. In: IMA (Institute of Mathematics and its Applications), *8th IMA Conference on Mathematical Education of Engineers*, Loughborough, United Kingdom, 20 April 2015, IMA. Available at: http://pure.qub.ac.uk/portal/files/15608345/JCole_IMA_paper_2015.pdf [Accessed 19 May 2016].

Cox, S.J., Daniell, G.J and Nicole, D.A., 1998. Using Maximum Entropy to Double One's Expected Winnings in the UK National Lottery. *Journal of the Royal Statistical Society Series D*, 47(4), pp.629-641. Available at: http://math2.uncc.edu/~imsonin/Lottery.pdf [Accessed 19 May 2016].

Powell, M. and Ansic, D., 1997, Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology*, 18(6), pp.605-628. Available at: http://darp.lse.ac.uk/PapersDB/Powell-Ansic_(JEP_97).pdf [Accessed 19 May 2016].

The National Lottery, 2015. *Lotto Online Game Procedures*. [online] Available at https://www.national-lottery.co.uk/games/lotto/game-procedures#int_fund [Accessed 19 May 2016].