Answers

Page 6

1. Roll a die and if it is a six ignore it and roll again. If the result is a one = 'Gift'.
If it is a two to five = 'No gift'.
Roll a second time and again record if you have a one = 'Gift' or a two to five = 'No gift'.

   Repeat a large number of times. The required probability is the number of times you have got a TWO roll result of 'No gift' divided by the number of rolls.
2. Use a spinner with twenty divisions. Each division is 18°. One of the divisions represents going by car, five of them represent the bus and 4 of them represent walking. The remaining ten divisions represent biking.
3. As there are six possible outcomes the best physical model would be a six-sided die. Roll the die the first time and note the result. Then roll for the second day and if it is a repeated result it is ignored and the die rolled again. For each day if the result is a five or six it represents buying lunch otherwise a lunch from home.
4. Each intersection represents two outcomes so either the toss of a coin or cutting a pack of cards and look at the colour only. This is done twice to see the result. Only one combination represents home.

Page 9

5. Usually 2 or 3 of each number.
Mean is likely to be between 3.5 and 5.5 (Theory = 4.5).
6. Usually 2 of each number.
Mean is likely to be between 5.5 and 8.5 (Theory = 7).
7. Usually 0 or 1 of each number.
Mean is likely to be between 39 and 61 (Theory = 50).
8. Usually 3 or 5 of each number including 10 and 16.
Mean is likely to be between 12 and 15 (Theory = 13).

Page 11

Simulation 1

Mini Lotto
In the description the following points will be needed:

- Two numbers need to be selected first.
- Randomly generate two numbers from 1 to 6 on a calculator Intg(1 + 6 x Ran#) and subsequently pressing the EXE key will generate random numbers from 1 to 6.
- Discard any repeated numbers in the two random numbers.
- Record a ✔ if the two numbers are correct and a ✖ if incorrect.
- Repeat this a large number of times (e.g. 50).

The experimental probability should be between 0 and 0.12.

The theoretical result is given by the number of different ways two numbers can be selected divided by 2 as the order of selection is not important. There are 6 selections possible for the first number leaving 5 for the second number.

$$\text{Ways} = \frac{\text{Select 1st} \times \text{Select 2nd}}{2}$$

$$= \frac{6 \times 5}{2}$$

$$= 15$$

$$P(\text{win}) = \frac{1}{15} \text{ (= 0.067 to 3 dp)}$$

Page 12

Simulation 2

Lego toys
Expected number of visits is between 14 and 15 but results will vary around these figures.

The largest number of visits is 32 and this is a lot more than expected.

Author's results were 15.8 visits.

#	Toy 1	Toy 2	Toy 3	Toy 4	Toy 5	Toy 6	CF
1	✓	✓	✓	✓	✓	✓	17
2	✓	✓	✓	✓	✓	✓	21
3	✓	✓	✓	✓	✓	✓	10
4	✓	✓	✓	✓	✓	✓	16
5	✓	✓	✓	✓	✓	✓	14
6	✓	✓	✓	✓	✓	✓	15
7	✓	✓	✓	✓	✓	✓	12
8	✓	✓	✓	✓	✓	✓	10
9	✓	✓	✓	✓	✓	✓	17
10	✓	✓	✓	✓	✓	✓	11
11	✓	✓	✓	✓	✓	✓	19
12	✓	✓	✓	✓	✓	✓	13
13	✓	✓	✓	✓	✓	✓	12
14	✓	✓	✓	✓	✓	✓	9
15	✓	✓	✓	✓	✓	✓	15
16	✓	✓	✓	✓	✓	✓	17
17	✓	✓	✓	✓	✓	✓	27
18	✓	✓	✓	✓	✓	✓	7
19	✓	✓	✓	✓	✓	✓	32
20	✓	✓	✓	✓	✓	✓	17
21	✓	✓	✓	✓	✓	✓	14
22	✓	✓	✓	✓	✓	✓	17
23	✓	✓	✓	✓	✓	✓	19
24	✓	✓	✓	✓	✓	✓	18
25	✓	✓	✓	✓	✓	✓	17

Page 14

Simulation 3

Doner Kebabs
The results will vary for each simulation but these are the author's results.

Trials 1 – 40.
MM = 22, MS + SM = 18, SS = 0
P(MM) = 0.550
P(MS or SM) = 0.450

Trials 41 – 80.
MM = 12, MS + SM = 27, SS = 1
P(MM) = 0.300
P(MS or SM) = 0.675

Trials 1 – 80.
MM = 34, MS + SM = 45, SS = 1
P(MM) = 0.425
P(MS or SM) = 0.5625

Trials 81 – 120.
MM = 13, MS + SM = 24, SS = 3
P(MM) = 0.325
P(MS or SM) = 0.600

All 120 trials
MM = 47, MS + SM = 69, SS = 4
P(MM) = 0.392
P(MS or SM) = 0.575

At 40 trials in each simulation the result varied for P(MS or SM) = 0.45 to 0.675

By the time there were 80 to 120 trials the variation was lower with P(MS or SM) = 0.575 against the theoretical answer of P(MS or SM) = 0.533 (not required).

Probability of two spicy sauces.
P(SS) = 1 – P(MM) – P(MS or SM)

P(SS) = 0.033

Page 17

Simulation 4

Waiting for the Bus
Use a calculator to generate a random number, n, i.e.

n = int(1 + 60*rand)

that the second bus arrives after the hour.

For each trial generate a random number 1 to 60 from the calculator representing the time after the hour that Diana arrives at the stop. For each number work out which bus she catches and hence her waiting time. Also calculate for your n the probability she catches each bus.

Page 17 Simulation 4 cont...

Example: If the first random number n is 17, then bus A still goes at 3 pm, 4 pm 5 pm etc. and bus B goes at 3:17 pm, 4:17 pm 5:17 pm etc. If Diana arrives at 4:07 pm she would wait 10 minutes until bus B. If she arrived at 4:20 pm she would have to wait 40 minutes until bus A.

Each result for Diana is one trial. The trials are repeated in sets of 50 and the results calculated.

The results will vary for each simulation and the value of n, but these are the author's results.

Trials 1 – 50 with n = 17.
Waiting time = 18.3 minutes
P(Bus B) = 0.20

Trials 51 – 100 with n = 17.
Waiting time = 20.5 minutes
P(Bus B) = 0.20

Trials 1 – 100 with n = 17.
Waiting time = 19.4 minutes
P(Bus B) = 0.20

With a different n of 35 minutes.
Trials 1 – 50 with n = 35.
Waiting time = 13.6 minutes
P(Bus B) = 0.58

Trials 51 – 100 with n = 35.
Waiting time = 16.3 minutes
P(Bus B) = 0.56

Trials 1 – 100 with n = 35.
Waiting time = 15.0 minutes
P(Bus B) = 0.57.

Yes the waiting time depends upon n. If the buses arrived together you would get a different waiting time than if they arrived 30 minutes apart.

If n was 30 minutes then the average waiting time should be 15 minutes. If the buses arrived together (n = 0) the average waiting time should be 30 minutes.

Page 17 Simulation 4 cont...

The proportion of the hour that bus B is the next bus is

$$P(\text{bus B}) = \frac{n}{60}$$

When n = 17 this is

P(bus B) = 0.283 against an observed result of 0.20

When n = 35 this is

P(bus B) = 0.58 against an observed result of 0.57

Waiting time after 50 trials was within 1.4 minutes of the result after 100 trials. In this case 50 trials seemed to be sufficient to estimate waiting time but the more trials the more accurate the result.

Page 20 Simulation 5

Lightning Strike
You expect that a plane will have a lightning strike between 5% and 15% of the time you will only get a crash between 2% and 7% of the time.
Average crashes 4%.
(P = 0.04).

Page 24 Simulation 6

Monty Hall
Probability is initially at

$$P(\text{win}) = \frac{1}{3}$$

as there are three doors and one winning door, but it changes to between 0.58 and 0.75 (based on the author's simulation) if they change after a goat is revealed.

The probability becomes

$$P(\text{win}) = \frac{2}{3}.$$

Page 31 Simulation 7
The Lift
Description
The simulation was done on an Excel spreadsheet. Each trial consisted of randomly generating a floor (from 1 to 10) for the four people going in the lift. Each time two or more of the four people get the same number (get off on the same floor) a 1 is recorded and added to the cumulative total of 'same floors'.

Page 31 Simulation 7 cont...

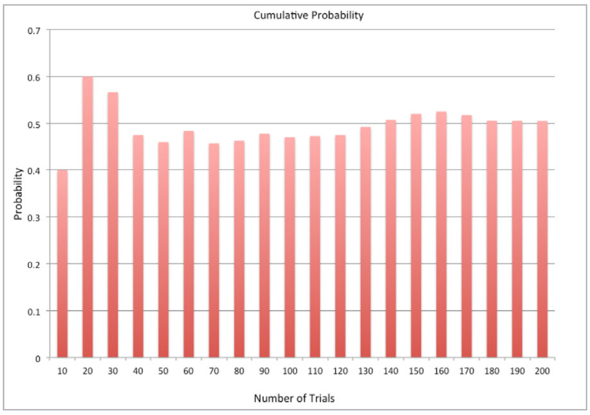
The trial was repeated over 150 times and the probability of two or more of the four people getting off on the one floor was calculated.

To estimate the number of trials needed, a graph of the probability versus the number of trials was done.

Analysis

The probability of two or more getting out at the same floor when four get into a lift with 10 floors is approximately 0.50.

The probability becomes more stable after about 100 trials. It can be as high as 0.53 and as low as 0.48.



The assumptions are that the floor each person is going to exit on is independent of the floor other members of the group will exit on. Also all the floors are equally likely.

The simulation could be extended by looking at the effect of increasing or decreasing the number of people and finding a relationship between the number and the probability. Similarly the number of floors could be varied. As you increase the number of floors you should decrease the probability and increasing the people in the lift will increase the probability.

Page 35 Simulation 8

Beating the Odds

Description

The simulation was done on an Excel spreadsheet. A random number was generated from 1 to 38. If this was less than 19 a $1 win was recorded.

If it was not less than 19 a second through to sixth number was generated. If any were less than 19 the $1 win was recorded otherwise a $63 loss was recorded. We did this a large number of times noting how many times Vanessa won $1 or lost $63.

A single trial is generating up to six random numbers (representing spins of the roulette wheel). If the result of any of the spins is a number under 19 then a $1 win is recorded. If all of the six spins are greater than or equal to 19 then a loss of $63 is recorded.

Page 35 Simulation 8 cont...

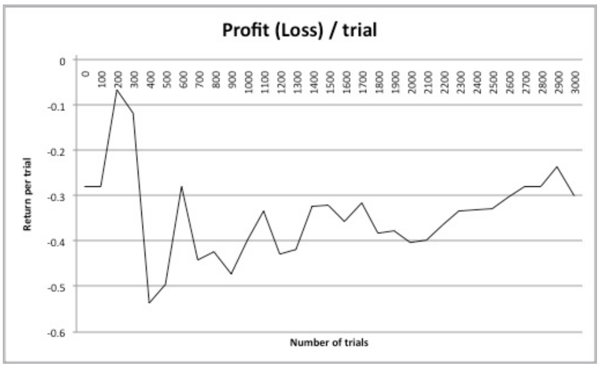
Based on the author's first 100 trials there were 53 of one spin, 19 of two spins, 13 of three spins, 5 of four spins, 4 of five spins and 3 of six spins a total of 188 spins.

Alternatively you would expect about 50 of the 100 spins to give a win first spin, 25 of the remaining 50 to give a win in two spins, about 13 of the remaining 25 would give a win in three spins etc. so an estimate of the number of spins is

$50 + 25 \times 2 + 13 \times 3 + 7 \times 4 + 3 \times 5 + 2 \times 6 = 194$ spins.

Analysis (Author's results)

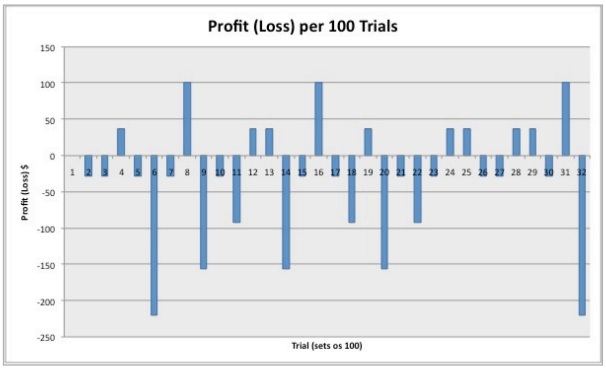
Although sometimes Vanessa can make $100 after 100 trials sometimes she is losing up to $220. The most common result was a loss of $28. After 3000 trials Vanessa had lost $900 so you would expect her to lose.



The loss per 100 trials is $30 or per single trial is a loss of 30¢.

The profit loss per 100 trials varied widely and even after 3000 trials had not settled to a figure. It would suggest at least 3000 trials based on this simulation.

The chance of losing six spins in a row and hence $63 is small and happens infrequently, but the consequence is large so it has a major effect. If we look at the cumulative profit loss per trial after 3000 trials we see that it is not settled. After 2000 trials it is close to a loss of 30¢ per trial. Therefore at least 2000 trials to get an estimate.



A theoretical probability of losing of 0.0212 55846 means the expected return is a loss of 36¢ per trial.

It is assumed that Vanessa does not change her strategy and the roulette wheel is 'fair'.

Page 40

Simulation 9 Marbles on a Pegboard

Description

The simulation could be done physically by tossing 6 coins and counting the number of heads. 0 heads means 0 left turns and 6 right turns so position A, 1 head means 1 left and 5 right turns so position B etc.

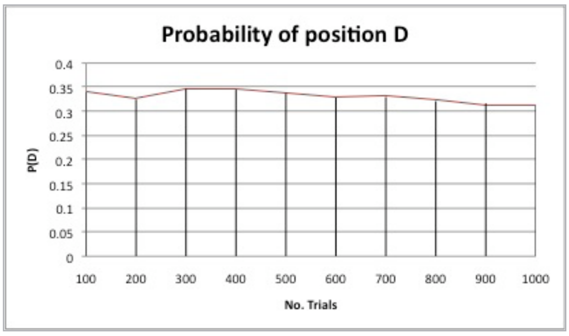
Alternatively it can be done on a spreadsheet with six results of 1 or 0 representing the number of left turns. 0 left turns means position A, 1 left turns means position B, 2 left turns means position C etc.

A single trial is six results (coins or random numbers) determining the number of left turns.

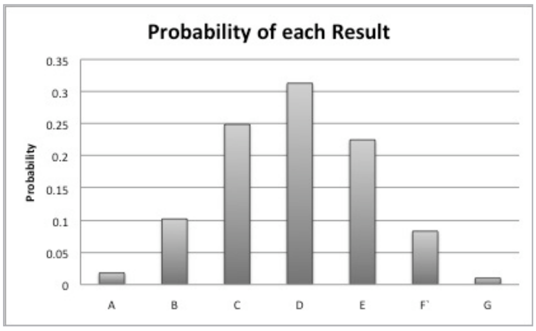
The results are recorded for each trial as a 1 against the position A through to G. The final result is both a total probability and cumulative probability for each position.

Analysis

The probability that the marble will end up in position 'D' is 0.31. The probability varied between 0.3 and 0.35 but stayed close to 0.31 from 900 onwards.



The distribution of results for all positions looked like the normal curve. 'D' is the most probable while 'A' and 'G' are unlikely.



Assumptions include that the marble had a 0.5 chance of going each side of a peg. In reality this may not be exactly correct. Also the initial marble was dropped in the middle.

The result looks realistic in that it is symmetrical about 'D'.

If there were more rows of pegs you would expect the probability distribution to smooth out and to look more and more like a normal distribution.

Page 45

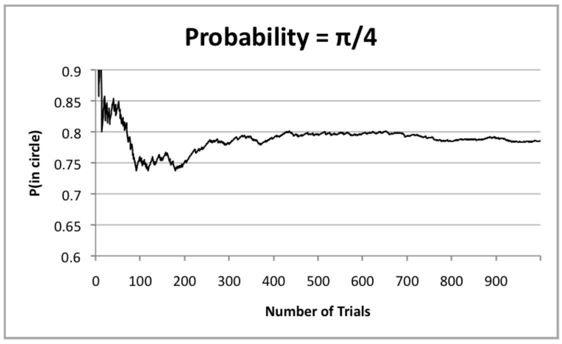
Simulation 10 Squaring the Circle

Description

The simulation was conducted in Excel on the computer. For each trial a random number from 0.000 000 000 to 1.999 999 999 was produced by multiplying the random number by 2. For the x ordinate one was subtracted so the number varied from $^-1$ to 1. This was also done for the y ordinate.

The distance this coordinate was from the centre was calculated with Pythagoras and if the distance was under 1 it was classified as in the circle otherwise out of the circle.

The number of trials was determined by graphing the probability that the shot was in the circle until the graph became stable.



The results were recorded as in the circle or out of the circle for each trial and the probability was the number in the circle divided by the number of trials.

This was repeated 1000 times and 785 were inside the circle as shown in this graph.

$$\frac{\pi}{4} = 0.785$$

$$\pi = 3.140 \qquad \text{(3 dp)}$$

Assumptions include ignoring the shot being exactly on the edge of the circle and assuming the shot was infinitely small.

The idea was good, but the probability did not quickly become stable so this is not recommended as a method of calculating pi.

Improvements include increasing the number of trials. With 1000 trials the most accuracy you could expect is 3 decimal places as the number of shots inside is going to be 3 digits. If you wanted 6 decimal places you would have to undertake the simulation one million times.

Pages 49 – 55 Practice Internal Assessment Task

Concepts	Achieved	Merit	Excellence
	To **Achieve** students show evidence of using each component of the simulation process.	To Achieve with **Merit** students link components of the simulation process to the context, explaining the relevant considerations made in the design of the simulation and supporting findings with statements that refer to evidence gained from the simulation.	To Achieve with Excellence, students show evidence of investigating the situation using each component of the simulation process, integrating statistical contextual knowledge throughout the process.
Design	Designs the simulation for the given situation.	Designs the simulation for the given situation.	Designs the simulation for the given situation.
Tools	Identifies the tools to be used in the simulation.	Describes **in detail** the tools to be used in the simulation.	Describes in detail the tools to be used in the simulation.
Trial	Defines what constitutes one trial and the number of trials in the simulation.	Defines what constitutes one trial and the number of trials in the simulation.	Defines what constitutes one trial and the number of trials in the simulation.
Data recording	Determines the data recording methods.	Determines the data recording methods.	Determines the data recording methods.
Simulation	Successfully carries out the simulation.	Successfully carries out the simulation **and repeats the simulation for different parameters.**	Successfully carries out the simulation and repeats the simulation for different parameters.
Assumptions		**Identifies at least one assumption.**	Identifies at least **two** assumptions in designing the simulation.
Record	Records the outcomes appropriately.	Records the outcomes appropriately.	Records the outcomes appropriately.
Displays and measures	Selects and uses appropriate displays and measures.	Selects and uses appropriate displays and measures.	Selects and uses appropriate displays and measures **and discusses the overall aspects of the distribution or simulation.**
Conclusion	Communicates findings in a conclusion.	Communicates findings clearly and links the simulation to the problem.	Communicates findings clearly and links the simulation to the problem. **Discusses in context how the simulation could be extended to improve our understanding of the problem.**

Design – The Lift to School
The simulation was set up in Excel.
Each trial consisted of randomly producing an arrival time for John. A random number 0 to 29 represented the time after 8 am that John arrived. We used Time = Integer(30 x Random #) as this makes a number from 0 to 29.99999 and when the integer value is taken it gives whole numbers 0 to 29.
In Excel this is = INT(30*RAND())
The departure time was 4 minutes later or 8:30 am if that was first.
Pete's arrival time was similarly generated to John's but his departure time was 8 minutes later or 8:30 am if that was first.

A decision needs to be made if the two times overlap. To do this we decide whether the start time of the shorter time interval (John's) is greater than when Pete arrives AND less than the time Pete departs OR if the departure time of John is between the interval he is stopped and waiting.
In Excel we use an IF statement so that IF (John > Pete AND John is less than Pete Departs) it records a 1. We could then do the same for John's departure time. If John Departs is less than Pete Departs AND John departs is greater than Pete arrives it records a 1 in a second column. If either of these if statements produce a 1 then there is an overlap.
This is easier if these decisions are made over several

Pages 49 – 55 Practice Internal Assessment Task cont...

columns and then a check made whether the start or departure time of John falls in the interval. It can be done in one statement in Excel.

Assumptions 1: That when the times are exactly the same then Pete and John see each other and Pete gets a lift. Assumption 2: That both do not vary the criteria even when they arrive close to 8:30.
The simulation will be run for at least 100 times but the probability Pete gets a lift will be watched to see if it is settling on a value or still moving.

The cumulative relative frequency (probability) that Pete gets a lift with John is about 0.345 but is still varying.
When the simulation was run repeatably the probability varied from 0.435 to 0.315 with a mean after 1000 of 0.377.

Sets 200	P(Lift)
Sim 1	0.345
Sim 2	0.360
Sim 3	0.430
Sim 4	0.315
Sim 5	0.435
Average	**0.377**

Taking the average probability (result of 1000 simulations) it means that in a 38 week school year of 5 days per week 72 days Pete will get a lift with John. It could be as high as 83 days or as low as 60 days.

If we keep the waiting time at 12 minutes in total but make John and Pete wait 6 minutes each it appears to have little or no effect on the probability of John getting a lift with Pete. To test this we increased the waiting time of Pete to 11 minutes and cut John's waiting time to 1 minute and again the probability was typically between 0.35 and 0.4.

If the total time was increased beyond 12 minutes then Pete's chances increase.

Keeping John at 4 minutes and changing Pete to 16 minutes resulted in a probability just over 0.5.

This is lower than you would expect because at 4 minutes and 16 minutes they are potentially covering $\frac{2}{3}$ of the 30 minute period, but because of the overlap or the interval going beyond 8:30 the probability never exceeded 0.6.

Trial	John arrive	John leaves	Pete arrive	Pete leaves	Lift	Prob lift
1	2	6	3	11	1	1
2	9	13	24	30	0	0.5
3	11	15	11	19	1	0.666 667
4	3	7	11	19	0	0.5
5	19	23	29	30	0	0.4
6	3	7	12	20	0	0.333 333
7	3	7	14	22	0	0.285 714
8	25	29	22	30	1	0.375
9	29	30	29	30	1	0.444 444
10	12	16	1	9	0	0.4