ELEN90054 Probability and Random Models
MATLAB Workshop #1 (Week 3)
Basic Probability: Dice & Goat Games

Department of Electrical and Electronic Engineering
University of Melbourne

March 3, 2024

Prelab: Read through and even complete Questions (I-a,b,c,d,e,f,g) as well as (II-d), and read through all questions before you come to your workshop session. The Prelab will not be assessed, but it's very helpful if you get it done before your workshop.

Logistics: This workshop is worth 4% of the overall subject assessment and will be done in pairs. Be aware that seeking or providing detailed assistance from/to people other than your workshop partner is collusion - see `http://academichonesty.unimelb.edu.au/plagiarism.html`.

Each group is expected to upload two files:

1. A pdf (scanned/typed) file containing their worked solutions;
   (a) Only one member of the group needs to upload the pdf file.
   (b) The naming convention of the file should be **`Workshop_W1_GpX.pdf`**, where X is the Group Number that is assigned to you upon signing up as a group in Canvas (for both online-only and on-campus workshops). More details on this to be announced shortly.
2. A single zip file, containing all the required functions and a main script which calls these functions to generate the required outputs as outlined in the workshop questions.
   (a) Only one member of the group needs to upload the zip file.
   (b) The naming convention of the file should be **`Workshop_W1_GpX_Matlab.zip`**, where X is the Group Number that is assigned to you (again, more details on this coming soon).

Both submissions should be made **before 11:59pm on Sunday, 24th March**.

Helpful notes before starting:

1. The MATLAB command `randi(k)` generates a random integer in $[1, \ldots, k]$ that is uniformly distributed, i.e., every value in the integer range $[1, \ldots, k]$ is equally likely.

   Calling the `randi` function several times in a row (for example via a for-loop) yields mutually independent results (for practical purposes anyway; strictly speaking, the results are actually from pseudo-random number generators, and cannot yield true unpredictability).

   If you call `randi` with the command `randi(5,100,2)`, the result will be a 100×2 matrix where each element is a random integer that is generated independently from the range $[1, 2, 3, 4, 5]$.

   For more information, type `help randi` in the MATLAB command window.

2. In a large number n of independent trials of an experiment, the <u>empirical</u> number of times k that an event B occurs in a trial satisfies

   $$\frac{k}{n} \simeq P(B)$$

   This is an imprecise statement of the law of large numbers – you will get a first taste of this phenomenon in this workshop but the precise theory will come later in the subject.

3. Use the MATLAB help function to get support regarding working with MATLAB (googling what you are looking for helps too).

4. The main purpose of this workshop is to get you to recognise "sequential experiments" and practice analysing and structuring the information that is given to you in order to compute probabilities, particularly getting your head around the notion of "sample space". Also you'll get a first taste of the law of large numbers.

Simulation of a Dice Game and the Goat Problem

(total of 36 marks = 24 marks in Part 1 + 12 marks in Part 2)

Part I. Let us introduce the following game of chance: it is played by rolling a pair of six-sided dice and observing the total number of spots T on their top faces. Assume that each of the dice has three sides with only 1 spot and the remaining three sides sides with only 2 spots (all sides are assumed to occur equally likely).[1] As a result, for each roll j the total number of spots $T(j)$ is an integer between 2 and 4. The rules of the game are as follows:

- If $T(1)$ equals 4 then the player loses immediately.
- If $x = T(1)$ is any other number $x \neq 4$, then the player keeps rolling the two dice, yielding totals $T(2), T(3), T(4), T(5), \dots$ each time until either
  1. The player rolls a total of x again, in which case the player wins, or
  2. The player rolls a total of 4, in which case the player loses.

  Remark: Note that $x = T(1)$ is the value of the **first** roll. Also keep in mind that a "roll" here refers to a pair of dice.

PART I Questions (24 marks).

I-a) **(2 marks)** Plot the pmf (probability mass function) of the random variable $T(1)$ (you may choose to do this by hand or by using MATLAB).

I-b) **(1 mark)** Draw a tree diagram for this game, making sure that you label the branches with their corresponding probabilities.

I-c) **(2 marks)** Specify the sample space S for this game. If S is countably infinite, be as descriptive as possible to explain all the possible outcomes.

I-d) **(1 mark)** Observe the total number of rolls in the game. What are the possible outcomes?

I-e) **(2 marks)** Show that the probabilities of the outcomes of your answer to the previous question add up to 1. You may want to use the following hint: use the geometric series formula.

I-f) **(1 mark)** What is the probability that the game never finishes? Explain your answer.

[1] In other words, the result of each die is either a 1 or a 2.

I-g) **(2 marks)** Calculate the probability that the player wins, using the geometric series to get an exact value.

I-h) **(4 marks)** Write MATLAB code to simulate the game. Make sure that you generate an output that indicates whether the player lost or won.

I-i) **(2 marks)** Write a program that calls your MATLAB procedure from the previous question n=10 times and gives the fraction of times that the player wins. Include the output of your MATLAB simulation in your report.

I-j) **(1 mark)** Repeat the previous question with $n = 50,000$.

I-k) **(2 marks)** Comment on how well or poorly your results in (I-i) and (I-j) relate to your answer in (I-g). Show the evidence.

I-l) **(4 marks)** Repeat (I-h) , (I-i) and (I-j) for the case that each die[2] is a normal die (with $1, 2, 3, 4, 5$ and 6 spots on its sides), which means that for each roll, the total number of spots is an integer between 2 and 12. The rules of the game are modified as follows with the new dice:

- If $T(1)$ equals 12 then the player loses immediately.
- If $x = T(1)$ is any other number $x \neq 12$, then the player keeps rolling the two dice, yielding totals $T(2), T(3), T(4), T(5), \ldots$ each time until either
  (a) The player rolls a total of x again, in which case the player wins, or
  (b) The player rolls a total of 12, in which case the player loses.

From your simulation give an estimate of the probability that the player wins.

End of PART I

PART II. *The Monty Hall Game Show problem*. Consider a game show in which there are three closed doors, with a car behind one and goats behind the others. According to the rules of the game, the player first selects a door and the game show host (called Monty Hall) then opens one of the other doors, to reveal a goat. The player is then given a choice: either she retains the original selection, or she switches to another door.

[2] Note that 'die' is the singular form, while 'dice' is plural and is used to refer to two or more die.

PART II Questions (12 marks).

II-a) **(2 marks)** Write MATLAB code to simulate the game. More specifically, you need to write two MATLAB procedures:

(a) Procedure 1 assumes that the player sticks to her original selection;

(b) Procedure 2 assumes that the player switches.

Make sure that you generate an output that indicates whether the player lost or won.

II-b) **(2 marks)** For each of your MATLAB procedures from (II-a), write a program that calls the procedure $n = 10$ times, and gives the fraction of times that the player wins. Include the outputs of your MATLAB simulation in your report.

II-c) **(1 mark)** Repeat (II-b) with $n = 50,000$.

II-d) **(3 marks)** Calculate the probability of winning the car if the player switches to a door that is different from her original selection. Explain your answer.

II-e) **(2 marks)** Comment on how well or poorly your results in (II-b) and (II-c) relate to your answer in (II-d). Show the evidence.

II-f) **(2 marks)** What should the contestant do if she wants the car? Switch or not switch? Explain your answer.

Historical Note: Monty Hall's game show was called "Let's make a Deal" and it ran from 1963 to 1991 on American television. Marilyn vos Savant put this question to readers in her column in a 1992 American newspaper. Her answer was very controversial at the time, it even led to angry responses by reputable mathematicians. Apparently, they had not taken the trouble to run a simulation such as the one that you just did in this workshop. If you're curious, see the following links:

`https://ima.org.uk/4552/dont-switch-mathematicians-answer-monty-hall-problem-wrong/`

`https://en.wikipedia.org/wiki/Monty_Hall_problem#Confusion_and_criticism`

End of PART II