ELEN90054 Probability and Random Models

MATLAB Workshop 1, 2018; week 3 = week of 12 March

Prepare this workshop individually, in particular **complete** Questions (I-a,b,c,d,e,f) as well as (II-d), and read through all questions before you come to your workshop session.

This workshop is worth 5% of the overall subject assessment and will be done in pairs. Allocation will be random by your demonstrator at the start of your workshop session, therefore you should not choose your project partner yourself. The next workshop will have a new allocation. 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 include workshop number, day, time and the assigned group number, e.g.
       Workshop_1_Mon09_Gp5.pdf for group 5 in the Monday 0900 hrs workshop slot.
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 following naming convention should **strictly** be followed:
       Workshop_1_Mon12_Gp5_Matlab.zip for group 5 in the Monday 1200 hrs workshop slot.

The workshop times are: Mon09, Mon12, Tue10, Tue17, Wed13, Wed17, Wed18, Wed19, Thu11, Thu19, Fri10, Fri13. The group numbers are randomly assigned by your demonstrator at the start of the workshop.

Both submissions should be made before the start of next workshop. **This is a strict deadline**.

In special circumstances you can email the pdf/code to your demonstrator. For demonstrator email address information, see “staff info” on the ELEN90054 LMS site.

Notes:

1) The MATLAB command *randi*(k) generates a random integer in $[1,\ldots,k]$ that’s 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, they are actually pseudo-random number generators, and cannot yield true unpredictability). For more information, type ‘help randi’ in the MATLAB command window.

2) In a large number n of *independent* trials of an experiment, the **empirical** number of times k that an event B occurs in a trial satisfies

$$\frac{k}{n} \approx \mathsf{P}[B].$$

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 recognize "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 Games with Dice and Goats (total = 36 marks + 4 on-time attendance marks)

Part I. Let's introduce the following game of chance: it is played by rolling a pair of dice and observing the total number of spots on their top faces. Assume that each of the dice has three sides with only 1 spot and 3 sides with only 2 spots (all sides are assumed to occur equally likely). As a result, for each roll the total number of spots on their top faces is an integer between 2 and 4. The rules of the game are as follows:

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

Part I Questions

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

I-b)(2 marks) Specify the sample space S for this game.

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

I-d) (2 marks) Show that the probabilities of the outcomes of your answer to the previous question add up to 1, by using the geometric series formula.

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

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

I-g)(5 marks) Write MATLAB code to simulate the game. Make sure that you generate an output that indicates whether the player lost or won.

I-h)(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.

Questions continue on the next page

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

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

I-k) (4 marks) Repeat (I-g) , (I-h) and (I-i) for the case that each dice is a normal dice, with 1,2,3,4,5 and 6 spots on its sides (so then, for each roll, the total number of spots is an integer between 2 and 12). From your simulation give an estimate of the probability that the player wins.

Part II.

<u>*The Monty Hall Game Show problem.*</u> 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.

Part II Questions

II-a) (3 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.

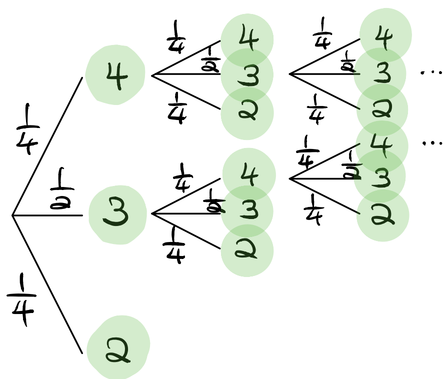
II-f) (2 marks) What should the contestant do if she wants the car? Switch or not switch? Explain your answer.

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….

End of Workshop 1 Questions

Dice Game

1a



$$\{1,1\},\ \{1,2\},\ \{2,1\},\ \{2,2\}$$
$$\Downarrow \quad \Downarrow \quad \Downarrow \quad \Downarrow$$
$$S = \{2,\ 3,\ 3,\ 4\}$$

1b

$$\left.\begin{array}{l} S = \{2\}, \\ \{3,2\}, \{4,2\}, \\ \{3,4,4,4,\dots,4,2\}, \{4,3,3,3,\dots,3,2\}, \end{array}\right\} \text{Lose}$$

$$\left.\begin{array}{l} \{3,3\}, \{4,4\}, \\ \{4,3,3,3,\dots,3,4\}, \{3,4,4,4,\dots,4,3\} \end{array}\right\} \text{Win}$$

1c Total number of rolls $= \{1, 2, 3, \dots, \infty\}$

1d $P(S)$

$= P(\text{starting with } 2) + P(\text{starting with } 3) + P(\text{starting with } 4)$

$= \frac{1}{4} + \left[\frac{1}{2}\frac{3}{4} + \frac{1}{2}\frac{1}{4}\frac{3}{4} + \frac{1}{2}\left(\frac{1}{4}\right)^2\frac{3}{4} + \dots\right] + \left[\frac{1}{4}\frac{1}{2} + \frac{1}{4}\left(\frac{1}{2}\right)^2 + \frac{1}{4}\left(\frac{1}{2}\right)^3 + \dots\right]$

$= \frac{1}{4} + \frac{3}{8}\left[1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \dots\right] + \frac{1}{8}\left[1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \dots\right]$

$= \frac{1}{4} + \frac{3}{8}\sum_{k=0}^{\infty}\left(\frac{1}{4}\right)^k + \frac{1}{8}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k$

$= \frac{1}{4} + \frac{3}{8}\frac{4}{3} + \frac{1}{8}2$

$= 1$

1e P(Game finishes)

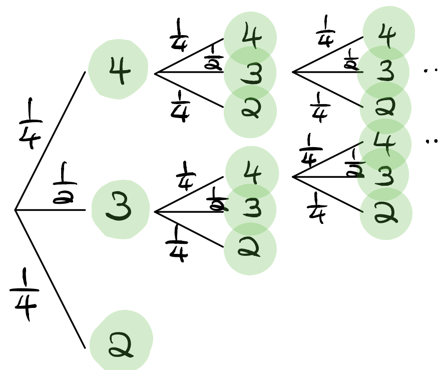
$$= \frac{1}{4} + \left[\frac{1}{2}\frac{3}{4} + \frac{1}{2}\frac{1}{4}\frac{3}{4} + \frac{1}{2}\left(\frac{1}{4}\right)^2\frac{3}{4} + \dots\right] + \left[\frac{1}{4}\frac{1}{2} + \frac{1}{4}\left(\frac{1}{2}\right)^2 + \frac{1}{4}\left(\frac{1}{2}\right)^3 + \dots\right]$$

$$= \frac{1}{4} + \frac{3}{8}\left[1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \dots\right] + \frac{1}{8}\left[1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \dots\right]$$

$$= \frac{1}{4} + \frac{3}{8}\sum_{k=0}^{\infty}\left(\frac{1}{4}\right)^k + \frac{1}{8}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k$$

$$= \frac{1}{4} + \frac{3}{8}\frac{4}{3} + \frac{1}{8}2$$

$$= 1$$



P(Game does not finish)

$$= 1 - P(\text{Game finishes})$$

$$= 0$$

1f P(player wins)

$$= P(\text{player wins starting with } 3) + P(\text{player wins starting with } 4)$$

$$= \left[\frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{4}\frac{1}{2} + \frac{1}{2}\frac{1}{4}\frac{1}{4}\frac{1}{2} + \dots\right] + \left[\frac{1}{4}\frac{1}{4} + \frac{1}{4}\frac{1}{2}\frac{1}{4} + \frac{1}{4}\frac{1}{2}\frac{1}{2}\frac{1}{4} + \dots\right]$$

$$= \frac{1}{4}\left[1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \dots\right] + \frac{1}{16}\left[1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \dots\right]$$

$$= \frac{1}{4}\sum_{k=0}^{\infty}\left(\frac{1}{4}\right)^k + \frac{1}{16}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k$$

$$= \frac{1}{4}\frac{1}{1-\frac{1}{4}} + \frac{1}{16}\frac{1}{1-\frac{1}{2}}$$

$$= \frac{1}{3} + \frac{1}{8}$$

$$= 0.458$$

1h Matlab Output $\Rightarrow$ 0.3

1i Matlab Output $\Rightarrow$ 0.456

1j For $n = 50{,}000$, the probability of winning converges to theoritical probability. $P(\text{winning}) \approx 0.456$

For $n = 10$, the probability of winning is not close to the theoritical probability at all, and is a poor estimate of the theoretical probability.

$P(\text{winning}) \not\approx 0.3$

1k Matlab Output $\Rightarrow 0.866$

$\therefore P(\text{winning}) \approx 0.866$

Monty Hall Problem

2b Matlab Output for staying with selection $\Rightarrow \frac{5}{10}$

Matlab Output for switching selection $\Rightarrow \frac{8}{10}$

2c Matlab Output for staying with selection $\Rightarrow 0.331$

Matlab Output for switching selection $\Rightarrow 0.665$

2d

Goat	Car	Goat

A is the event of selecting a car

B is the event of switching

$P(A) = \frac{1}{3}$

$P(B) = 1$ because player will definitely switch

$P(A \cap B) = \frac{2}{3}$ because to switch and win, player must select a goat and $P(\text{selecting goat}) = P(A^c)$

$= \frac{2}{3}$

$$\therefore P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{2}{3}$$

2e The result from 50 000 trials is close to the theoritical probabilities.

$P(A) \approx 0.331$ $\quad P(A|B) \approx 0.665$

The result from 10 trials is not close to the theoritcal probability.

$P(A) \not\approx \frac{5}{10}$ $\quad P(A|B) \not\approx \frac{8}{10}$

2f Contestant should always switch his/her selection because,

$P(A|B) = \frac{2}{3}$

$P(A) = \frac{1}{3}$

$\therefore P(A|B) > P(A)$

The chance of winning is higher if contestant switches his/her selection.