MAST20004 Probability

Tutorial Set 1

1. A simple poker machine consists of three wheels. On each wheel there are pictures of six types of fruit; one each of a banana (B), a strawberry (S), a pear (P), an apple (A), a watermelon (W) and a kiwi fruit (K). When a player pulls the arm of the machine, each of the wheels spins individually and comes to rest with a picture of one of the pieces of fruit in the viewing window. The player who pulls the arm thus sees pictures of three pieces of fruit.
   (a) Describe the sample space for the random experiment that occurs when the arm is pulled.
   (b) How many possible outcomes are there?
   (c) Consider the events
      (i) $E = \{\text{S showing on at least two wheels}\}$
      (ii) $F = \{\text{Only fruits that start with a letter after H in the alphabet visible on the wheels}\}$
   List the outcomes in E, F, $E \cap F$, $(E \cap F)^c$ and say how many outcomes there are in E^c, F^c and $E^c \cap F^c$.

   Solution:
   (a) Ω is the set of all three-letter words made up of the symbols, B, S, P, A, W and K.
   (b) $6^3 = 216$.
   (c) $E = \{$(S,S,B),(S,S,S),(S,S,P),...,(K,S,S) $\}$
   $F = \{$(S,S,S),(S,S,P),(S,S,W),...,(K,K,K) $\}$
   $E \cap F = \{$(S,S,S),(S,S,P),(S,S,W),...,(W,S,S),(K,S,S) $\}$
   $(E \cap F)^c = \{$(B,B,B),(B,B,S),(B,B,P),...,(K,K,K)$\}$
   The number of outcomes in E^c is $216 - 16 = 200$.
   The number of outcomes in F^c is $216 - 64 = 152$.
   The number of outcomes in $E^c \cap F^c$ is $216 - 70 = 146$.

2. Of 25 laptop computers in a supply room, eight run on Windows 7, ten use Internet Explorer as a browser, and ten use neither program. Use A to denote the set of computers that run on Windows 7 and B to denote the set of computers that use Internet Explorer. Consider the random experiment of randomly choosing a box containing a computer and opening it. Symbolically denote the following events and give the number of outcomes in each:
   (a) the computer ran on Windows 7 and used Internet Explorer,
   (b) the computer either ran on Windows 7 or used Internet Explorer,
   (c) the computer ran on Windows 7 but did not use Internet Explorer, and
   (d) the computer used exactly one of the programs.

   Solution:
   (a) $A \cap B$, 3 outcomes.
   (b) $A \cup B$, 15 outcomes.
   (c) $A \cap B^c$, 5 outcomes,
   (d) $(A \cap B^c) \cup (A^c \cap B)$, 12 outcomes.

3. Suppose that, for events A and B, we know the following probabilities: $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.3$ and $\mathbb{P}(A \cup B) = 0.5$. Compute:
   (a) $\mathbb{P}(A^c)$,
   (b) $\mathbb{P}(A^c \cap B^c)$,
   (c) $\mathbb{P}(A \cap B)$,
   (d) $\mathbb{P}(A^c \,|\, B)$, and
   (e) $\mathbb{P}(A \,|\, B^c)$.

   Solution:
   (a) $\mathbb{P}(A^c) = 0.7$,
   (b) $\mathbb{P}(A^c \cap B^c) = 0.5$,
   (c) $\mathbb{P}(A \cap B) = 0.1$,
   (d) $\mathbb{P}(A^c \,|\, B) = 2/3$, and
   (e) $\mathbb{P}(A \,|\, B^c) = 2/7$.

4. Let $A, B \subset \Omega$ be events. Using the Axioms of Probability and the properties on Slide 36 (except Property (9)) prove that
   (a) if $B \subset A$, then $\mathbb{P}(A \backslash B) = \mathbb{P}(A) - \mathbb{P}(B)$ ($A \backslash B = A \cap B^c$, that is, $A \backslash B$ contains all elements of A that are not in B;
   (b) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ (Hint: You may use the result from Part (a)).
   (c) $\mathbb{P}(A \cap B) \geq \mathbb{P}(A) + \mathbb{P}(B) - 1$. (Hint: You may use the result from Part (b)).

   By considering the rolling of a fair six-sided die, write down an example of events A and B to illustrate the expressions in Parts (a), (b), and (c). You may use different events for each part.

   Solution:
   (a) We have $A = B \cup (A \cap B^c)$ and $B \cap (A \cap B^c) = \emptyset$. Therefore by Axiom 3,

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}\left(B \cup (A \cap B^c)\right) \\ &= \mathbb{P}(B) + \mathbb{P}\left(A \cap B^c\right), \end{aligned}$$

   which gives $\mathbb{P}(A \backslash B) = \mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(B)$.

   Let $A = \{1, 2, 5, 6\}$ and $B = \{2, 6\}$. $A \subset B$ and $\mathbb{P}(A \backslash B) = \frac{1}{3} = \frac{2}{3} - \frac{1}{3} = \mathbb{P}(A) - \mathbb{P}(B)$.

   (b) We have $A \cup B = A \cup (B \cap (A \cap B)^c)$ and $A \cap (B \cap (A \cap B)^c) = \emptyset$. Therefore by Axiom 3,

$$\begin{aligned} \mathbb{P}(A \cup B) &= \mathbb{P}\left(A \cup (B \cap (A \cap B)^c)\right) \\ &= \mathbb{P}\left(A\right) + \mathbb{P}\left(B \cap (A \cap B)^c\right) \\ &= \mathbb{P}\left(A\right) + \mathbb{P}\left(B \backslash \left(A \cap B\right)\right) \\ &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \end{aligned}$$

   The fourth equality uses Part (a) since $A \cap B \subset B$.

   Let $A = \{1, 2, 3, 4\}$ and $B = \{4, 5, 6\}$. $\mathbb{P}(A \cup B) = 1 = \frac{2}{3} + \frac{1}{2} - \frac{1}{6} = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

   (c) Using Part (b) we have

$$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) \\ &\geq \mathbb{P}(A) + \mathbb{P}(B) - 1, \end{aligned}$$

since $\mathbb{P}(A \cup B) \leq 1$ by Property (8) on Slide 36.

Let $A = \{1, 6\}$ and $B = \{3, 4, 5\}$. $\mathbb{P}(A \cap B) = 0 \geq \frac{1}{3} + \frac{1}{2} - 1 = -\frac{1}{6} = \mathbb{P}(A) + \mathbb{P}(B) - 1$.

MAST20004 Probability

Computer Lab 1

The aim of this lab is to

- recall some basic features of MATLAB;
- use MATLAB to simulate a simple die experiment;
- use MATLAB to investigate the bus stop paradox from lectures.

You are unlikely to have much time to look at the Bus Stop paradox (Exercise B) unless you are already familiar with MATLAB, so it is optional.

Introduction to MATLAB

The expectation is that you should all have a basic familiarity with MATLAB.

A document "Introduction to MATLAB" is available on the LMS (Canvas). Using this document, your tutors will give a presentation introducing you to the basics of MATLAB including:

- how to launch MATLAB;
- what the Command and Workspace windows are for and choices for arranging them on your desktop;
- how to define vectors and matrices in MATLAB and perform calculations with them;
- what relational and logical operators are and how to use them;
- how to use the comprehensive help system;
- how to copy the supplied lab programs (m-files) from the Maths and Stats lab materials server into the Student Data folder and open them in the MATLAB editor;
- how to use the MATLAB m-file editor to inspect, edit and run a MATLAB mfile (which is basically a file containing a computer program made up of MATLAB commands), and how to run an m-file from the Command Window.

You should then break up into groups of 2-3 to try to apply these skills to tackle at least Exercise A. Some students will have some (or quite a lot) of previous programming experience. If you are one of these please try to join a group with less experienced students so you can actively help them to start learning - and remember that some students will have no previous programming experience.

Simulation

To *simulate* means to reproduce the results of a real phenomenon by artificial means. In this subject we are interested in phenomena which have a random element. What we will simulate throughout the semester are the outcomes (elements of an outcome space) and/or the values of a random variable which summarises the outcome in some way. MATLAB (and computers in general) provide a fast and convenient way of carrying out these types of simulations.

There are several reasons to use simulation. Observing a random experiment repeatedly could give us intuition about its behaviour. In such cases, simulation may suggest how to approach modelling the experiment. Often the system under study is so complex that direct analysis is either impossible or very difficult. Then simulation may be the only means of studying the system in question.

Exercise A - Die throwing experiment

1. Consider the experiment of rolling a fair die twice and recording the scores. Let A denote the event that the first die score is 1 and B the event that the sum of the scores is 7.
   - (a) Define an outcome space Ω (mathematically).
   - (b) Express the events A and B as subsets of Ω.
   - (c) Express the events $A \cup B$ and $A \cap B$ as subsets of Ω. Are the events A and B independent? (ie does $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$).
   - (d) Compute $\mathbb{P}(A \cup B)$ and $\mathbb{P}(A \cap B)$.
2. Now we will test your theoretical values for $\mathbb{P}(A \cup B)$ and $\mathbb{P}(A \cap B)$ against empirical estimates from a MATLAB simulation. The relevant m-file **Lab1ExA.m** is available on the Maths and Stats lab materials server.

   Before running Lab1ExA.m you should study it in the editor and try to start to understand how it works and what result it produces (you will not understand everything in this first lab but hopefully more later), in particular:
   - (a) briefly look at the m-file **randunifd.m** which is a function called by Lab1ExA (this must also be copied to the Student data file for the program to work). Seeking help on the rand function might help you to understand how it works but don't worry too much about the details at this stage;
   - (b) remove the semicolon from the line *dicethrows=randunifd(1,6,2)*; and check the output in the command window to see what it produces on each repetition;
   - (c) use help for relational and logical operators to understand lines 13 and 16.
   - (d) use help for **for** and **if** statements to understand the **for** loop and the **if** statements.

   You can run the program by typing its name in the Command Window (Lab1ExA) and the output will also appear there. Compare the empirical results with your calculations. Run the program several times for different nrep values and note the impact on the variability of your empirical estimates.

Exercise B - Bus stop paradox (Optional)

1. Recall the Bus stop paradox discussed in the first lecture. This paradox is simulated in the m-file **Lab1ExB1.m** in the following way:
   (a) each day Rob arrives at the bus stop at a random instant (which in this simulation we choose to be uniformly distributed between 11am and 12noon) and waits for a bus.
   (b) buses arrive randomly at an average of one per hour (actually this means that the time between buses is exponentially distributed with mean 1 hour - we will learn more about this distribution later in the course). Each day buses start at time 0 (midnight) and keep running until the first bus after Rob's arrival time (we don't simulate any further as what happens after this is not of interest to us).

   Lab1ExB1 simulates a possible day of this experiment and plots the arrival times of the buses together with three waiting times:
   (a) W-1-2 - the wait between the first and second bus (provided there are two);
   (b) W-B-A - the wait from the last bus before Rob's arrival (or midnight if there is none) until the first bus after Rob's arrival;
   (c) W-Rob - how long Rob waits for a bus.

   The program works on an indefinite loop - so you can see another day simulated by hitting any key. To terminate the program you must hold down the Control key and hit C. Run several days and note the random arrival time for Rob between 11 and 12 noon.

   We claimed in lectures that the mean waiting time for Rob is still 1 hour. In fact the distribution of Rob's waiting time is exactly the same as the time between buses so half the time W-1-2 should exceed W-Rob and the other half the reverse should occur. As a rough check do 20 simulations and count the number of times W-1-2 is larger than W-Rob.

2. Of course we need to do more simulations to get an accurate idea of the average or mean waiting time. The m-file **Lab1ExB2.m** simulates 4000 days and prints out the estimated means for the three waiting times. You can run it directly from the Command window simply by typing Lab1ExB2, and the output is also printed in the Command window. Run this file to check the claim in lectures that W-1-2 and W-Rob have the same mean. What is the mean for W-B-A? Can you think (intuitively) of a way to explain this? If you have time run the program with uniformly distributed times between buses (change exponential to uniform in line 12) and think about the results.