# MAST20004 Probability – Assignment 2

# Name: Lucas Fern

# Student ID: 1080613

---

- ## New completion process

  Note this assignment is being handled using a similar process to that now planned for the final exam so you can start to become familiar with it.

  To complete this assignment, you need to write your solutions into the blank answer spaces following each question in this assignment pdf.

  If you have a printer (or can access one), then you must print out the assignment template and handwrite your solutions into the answer spaces.

  If you do not have a printer (**NB you must have one for the exam so get one asap**), but you can figure out how to annotate a PDF using an iPad/Android tablet/Graphics tablet or using Adobe Acrobat, then annotate your answers directly onto the assignment PDF and save a copy for submission.

  Failing both of these methods, you may handwrite your answers as normal on blank paper and then scan for submission (but note that you will thereby miss valuable practice for the exam process).

  Scan your assignment to a PDF file using your mobile phone then upload by going to the Assignments menu on Canvas and submit the PDF to the GradeScope tool by first selecting your PDF file and then clicking on 'Upload PDF'.

- The **strict** submission deadline is **3 pm Melbourne time on Friday 1 May**. You have two weeks instead of the normal one week to complete this assignment. Consequently late assignments will **NOT** be accepted. We recommend you submit at least a day before the due date to avoid any technical delays. If there are extenuating, eg medical, circumstances, contact the Tutorial Coordinator.

- There are 5 questions, of which 3 randomly chosen questions will be marked. Note you are expected to submit answers to **all** questions, otherwise **a mark penalty will apply**.

- Working and reasoning **must** be given to obtain full credit. Give clear and concise explanations. Clarity, neatness, and style count.

1. Let $X$ be a random variable with $S_X = \{0, 1, 2, \ldots\}$ and probability mass function of the form

$$p_X(k) \;=\; c\, 2^{-k}.$$

(a) Determine $c$.

$$c \cdot \sum_{k=0}^{\infty} 2^{-k} = 1$$

$$= \frac{c}{1-(2^{-1})}$$

$$1 = 2c$$

$$c = \frac{1}{2}$$

(b) Is $X$ more likely to take values that are divisible by 4, or values that are not divisible by 4? Justify your answer.

$$\sum_{n=0}^{\infty} p_X(4n) = \sum_{n=0}^{\infty} \frac{1}{2} 2^{-4n}$$

$$= \frac{1}{2} \sum_{n=0}^{\infty} (2^{-4})^n$$

$$= \frac{1}{2} \left( \frac{1}{1-2^{-4}} \right)$$

$$= \frac{8}{15}$$

$\frac{8}{15} > 1 - \frac{8}{15} \Rightarrow P(X=4n)$; $n = 0, 1, 2 \ldots$ is greater than $1 - P(X=4n)$.

more likely to be divisible by 4.

(c) If they exist, calculate $\mathbb{E}(X)$ and $\mathrm{sd}(X)$ from first principles.

$$E(X) = \sum_{x \in S_X} x \cdot p_X(x)$$

$$= \sum_{x=0}^{\infty} \frac{1}{2} x 2^{-x}$$

$$= \frac{1}{2} \sum_{x=0}^{\infty} \frac{x}{2^x}$$

$$= \frac{1}{2}\left(0 + \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{4}{16} \dots\right)$$

$$E(X) = 1$$

$$E(X^2) = \frac{1}{2} \sum_{x=0}^{\infty} x^2 \cdot 2^{-x}$$

$$= \frac{1}{2}\left(0 + \frac{1}{2} + \frac{2^2}{4} + \frac{3^2}{8} + \frac{4^2}{16} + \dots\right)$$

$$= \frac{1}{2}(6)$$

$$= 3$$

$$\mathrm{Var}(X) = E(X^2) - E(X)^2$$

$$= 3 - 1$$

$$= 2$$

$$\therefore \ \mathrm{sd}(X) = \sqrt{2}$$

2. Suppose $X$ is a continuous random variable. Consider, for $x \in \mathbb{R}$, the distribution function $F_X$ given by

$$F_X(x) = \begin{cases} 0, & x < 0 \\ ax, & 0 \leq x < 3 \\ \frac{x}{12} + \frac{1}{4}, & 3 \leq x < b \\ 1, & x \geq b \end{cases},$$

where $a$ and $b$ are real constants.

(a) Determine $a$ and $b$.

$F_X(3) = \frac{3}{12} + \frac{1}{4} = \frac{1}{2}$

$\lim_{x \to 3^-} F_X(x) = 3a$ (set equal because continuous)

$3a = \frac{1}{2} \Rightarrow a = \frac{1}{6}$

and

$\lim_{x \to b^-} F_X(x) = \frac{b}{12} + \frac{1}{4} = 1$

$b = 12 \cdot \frac{3}{4}$

$= 9$

$a = \frac{1}{6}, \; b = 9$

(b) Find the probability density function of $X$, $f_X$.

$f_X(x) = \frac{d}{dx}\left(F_X(x)\right)$

$$f_X(x) = \begin{cases} 1/6, & 0 \leq x < 3 \\ 1/12, & 3 \leq x < 9 \\ 0, & \text{elsewhere} \end{cases}$$

(c) Calculate $\mathbb{P}(4 < X^2 < 16)$.

$$P(4 < X^2 < 16) = P(2 < X < 4)$$
$$= F_X(4) - F_X(2)$$
$$= \frac{7}{12} - \frac{2}{6}$$
$$= \frac{1}{4}$$

(d) Calculate $\mathbb{E}(X)$ in *two different* ways.

$$E(X) = \int_0^\infty (1 - F_X(x))\,dx$$
$$= \int_0^3 \left(1 - \frac{1}{6}x\right)dx + \int_3^9 \left(1 - \left(\frac{x}{12} + \frac{1}{4}\right)\right)dx$$
$$= \left[x - \frac{1}{12}x^2\right]_0^3 + \left[\frac{3}{4}x - \frac{x^2}{24}\right]_3^9$$
$$= \frac{9}{4} + \frac{3}{2} = \frac{15}{4}$$

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$
$$= \int_0^3 \frac{1}{6}x\,dx + \int_3^9 \frac{1}{12}x\,dx$$
$$= \frac{1}{12}\left[x^2\right]_0^3 + \frac{1}{24}\left[x^2\right]_3^9$$
$$= \frac{3}{4} + 3 = \frac{15}{4}$$

3. Consider the experiment where a fair coin is tossed 100 times, and let $X$ be the length of the longest sequence of consecutive Tails.

   (a) In the Matlab file Assignment2Ex3a_2020.m you will need to add some code to simulate the experiment.

   (i) Replace the comment in Line 16 with the required condition and write the resulting Line 16 in the box below.

```
if ~coin_tosses(i)
```

   (ii) Replace the comment in Line 22 with the required condition and write the resulting Line 22 in the box below.

```
if n_Tails > max_Tails
```

   (iii) Replace the comment in Line 27 with the required code and write this code in the box below.

```
p(max_Tails+1) = p(max_Tails+1)+1;
```

   (iv) Replace the comment in Line 33 with the required code and write this code in the box below.

```
mean = sum(RelFreq.*(0:N))
```

   (v) Replace the comment in Line 35 with the required code and write this code in the box below.

```
var = sum(RelFreq.*((0:N).^2))
```

   (vi) What is the most likely length of the longest sequence of Tails, and what is its estimated probability (stated to 2 decimal places)?

   Most likely is 5 ; $P(X=5) \approx 0.26$

   (vii) Write down your estimates for $\mathbb{E}(X)$ and $V(X)$. Give your answers to 2 decimal places.

   $E(X) = 6.00$ ; $V(X) = 3.22$

(b) Now let $Y$ be the length of the longest sequence of tosses which consecutively alternate between Heads and Tails. Assume that if the sequence of tosses is all Heads or all Tails, then $Y = 1$. In the Matlab file Assignment2Ex3b_2020.m you will need to add some code to simulate the experiment.

(i) What is the most likely length of the longest sequence of alternating Heads and Tails, and what is its estimated probability (stated to 2 decimal places)?

Most likely is 6 ; $P(Y=6) \approx 0.26$

(ii) Write down your estimates for $\mathbb{E}(X)$ and $V(X)$. Give your answers to 2 decimal places.

$E(X) = 6.98$ , $V(X) = 3.22$

(iii) Write down the relationship between $X$ and $Y$.

$Y = X + 1$

4. Let $X \stackrel{d}{=} \mathrm{Bi}(n, p)$. Show that

$$\mathbb{E}\left(\frac{1}{1+X}\right) = \frac{1-(1-p)^{n+1}}{p(n+1)}.$$

$$X \sim \mathrm{Bi}(n,p) \Rightarrow p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\therefore E\left(\frac{1}{1+X}\right) = \sum_{x=0}^{n} \left(\frac{1}{1+x}\right) \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \left(\frac{1}{1+x}\right) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \frac{n!}{(x+1)!(n-x)!} p^x (1-p)^{n-x}$$

$$= \frac{1}{p(n+1)} \sum_{x=0}^{n} \frac{(n+1)!}{(x+1)!(n-x)!} p^{x+1} (1-p)^{n-x}$$

$$= \frac{1}{p(n+1)} \sum_{x=0}^{n} \binom{n+1}{x+1} p^{x+1} (1-p)^{n-x}$$

let $n+1 = a$ and $x+1 = b$ $\quad a-1 = n$, $b-1 = x$

$$= \frac{1}{p(n+1)} \sum_{b=1}^{a} \binom{a}{b} p^b (1-p)^{a-b}$$

$\sum_{b=1}^{a} \binom{a}{b} p^b (1-p)^{a-b}$ is the sum of all terms in a binomial distribution $\mathrm{Bi}(a,p)$ EXCEPT $\binom{a}{0} p^0 (1-p)^a$.

$$\therefore \frac{1}{p(n+1)} \sum_{b=1}^{a} \binom{a}{b} p^b (1-p)^{a-b} = \frac{1}{p(n+1)}\left(1-(1-p)^a\right)$$

$$= \frac{1-(1-p)^{n+1}}{p(n+1)}$$

5. A certain electronic component's lifetime $T$, is distributed according to an exponential distribution. On average, 10 components fail per year.

   (a) Write down the probability density function of $T$, and state its mean.

   $X \sim \exp(10)$

   $f_T(t) = 10e^{-10t}$ and $E(T) = \frac{1}{10}$

   (b) The *hazard* rate $h_T$, of the random variable $T$, measures the rate at which an electronic component will fail, depending on its age $t$. For $t \geq 0$, the hazard rate is given by

   $$h_T(t) = \frac{f_T(t)}{1 - F_T(t)},$$

   where $f_T$ and $F_T$ are the density and distribution functions of $T$, respectively.

   Calculate $h_T(t)$ and give an interpretation for your answer.

   $F_T(t) = 1 - e^{-10t} dt$

   $h_T(t) = \dfrac{10e^{-10t}}{1-(1-e^{-10t})}$

   $= 10$

   The constant hazard rate tells us that at any value of t we should expect the same amount 10 failures per year.

(c) Calculate the probability that an electronic component will fail within 3 months.

3 months $= \frac{1}{4}$ of a year

$$\therefore P(T \leqslant 1/4) = F_T(1/4)$$
$$= 1 - e^{-\frac{10}{4}}$$
$$\approx 0.918$$

(d) Suppose that a customer buys 20 electronic components. Calculate the probability that at least 3 components will still be working after 3 months. Assume that all electronic components fail independently of each other.

$P(T > 1/4)$ is the probability of any one component working longer than 3 months.

$$P(T > 1/4) = 1 - P(T \leqslant 1/4)$$
$$= e^{-\frac{5}{2}} \text{ (using 5c. answer)}$$

Let $Y$ be the amount of components working after 3 months.

$Y \sim Bi(20, e^{-\frac{5}{2}})$

$$P(Y \geqslant 3) = 1 - P(Y=0) - P(Y=1) - P(Y=2)$$
$$= 1 - (1-e^{-\frac{5}{2}})^{20} - 20(e^{-\frac{5}{2}})(1-e^{-\frac{5}{2}})^{19} - \binom{20}{2}(e^{-\frac{5}{2}})^2(1-e^{-\frac{5}{2}})^{18}$$
$$\approx 0.223$$