Question 1 – Web Page Counter

Consider an application that counts the number of web pages visited on a server on a specific day.

Let A denote the event that at least 20 web pages were visited. Let B denote the event that the number of visited web pages is less than or equal to 37, and let C denote the event that at least 45 web pages were visited that day.

(a) What is the sample space for this experiment?

Solution:
The positive integers, $\mathbb{Z}^*$

(b) Are the events A, B, C mutually exclusive?

Solution:
No, as the events overlap

Describe the events:

(c) $A \cap B$

Solution:
This is the intersection of the two events so only contains the numbers which are in both events. This includes all the times greater than 20 pages were visited along with those less than or equal to 37. So that is that the number of web pages is between 20 and 37.

(d) $\overline{A}$

Solution:
This is the compliment to even A which is everything that is not in event A. Therefore This is the event that strictly less than 20 pages were visited.

(e) $A \cup C$

Solution:
This is all the counts in either A or C. As C is that at least 45 pages were visited it is completely contained in A which is that at least 20 pages were visited, so the set is simply, $A \cup C = A$.

(f) $\overline{B} \cap C$

Solution:
Here the count has to be in the compliment of B which is that the number of web pages visited is over 37 and in C which is that at least 45 pages were visited. Therefore, $\overline{B} \cap C = C$

(g) $A \cup \left(\overline{B} \cap C\right)$

Solution:
From above we have seen that $\overline{B} \cap C = C$ and $A \cup C = A$. Using these two results we get that

$$\begin{aligned} A \cup \left(\overline{B} \cap C\right) &= A \cup C \\ &= A. \end{aligned}$$

(h) $\overline{(A \cup C)}$

Solution:
As we have seen above $A \cup C = A$ so $\overline{A \cup C} = \overline{A}$. Or that less than 20 pages were visited.

(i) $A \cap B \cap \overline{C}$

Solution:
Here the values have to satisfy the events A, that as least 20 pages were visited, B, that less than or equal to 37 pages were visited, and $\overline{C}$ that less than 45 pages were visited. As can be seen $B \cap \overline{C} = B$ so $A \cap B \cap \overline{C} = A \cap B$. Therefore between 20 and 37 pagers were visited.

Question 2 – Transmitting Bits

Four bits are transmitted over a digital communications channel. Each bit is either distorted or received without distortion. Let A_i denote the event that the ith bit is distorted, $i = 1, \ldots, 4$.

(a) Describe the sample space for this experiment.

Solution:
For this question we will identify distorted bits by 1 and undistorted bits as 0. Here the sample space is all the combinations of four bits either distorted or not distorted, i.e.

$$\begin{aligned}\Omega &= \{0, 1\}^4 \\ &= \{0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111\}\end{aligned}$$

(b) Are the A_i's mutually exclusive?

Solution:
As can be seen above more than one bit can be distorted and they can happen in any order, therefore they are not mutually exclusive.

Describe the outcomes in each of the following events:

(c) A_3

Solution:
The third bit is distorted $\{0010, 0011, 0110, 0111, 1010, 1011, 1110, 1111\}$

(d) $\overline{A_2}$

Solution:
This event is the event that the second bit is not distorted $\{0000, 0001, 0010, 0011, 1000, 1001, 1010, 1011\}$

(e) $A_1 \cap A_2 \cap \overline{A_3} \cap A_4$

Solution:
This is the event that bits 1,2 and 4 are distorted while bit 3 is not distorted $\{1101\}$.

(f) $(A_1 \cup A_2) \cap (A_3 \cap A_4)$

Solution:
This is the event that either the first or second bit is distorted and the final two bits are distorted $\{1011, 0111, 1111\}$

(g) $(\overline{A_1 \cap A_2}) \cup (\overline{A_3} \cap A_4)$

Solution:
This is the event that the first two bits are not distorted or the third bit is not distorted while the fourth bit is distorted $\{0000, 0001, 0011, 0010, 1001, 1101, 0101\}$.

Question 3 – Basic Simulation with R

The sample space of a random experiment is $\{a, b, c, d, e\}$ with probabilities 0.1, 0.05, 0.4, 0.2, 0.25 respectively. Let A denote the event $\{a, b, c\}$, and let B denote the event $\{c, d, e\}$. Determine the following:

(a) $P(A)$

Solution:
To calculate this we add the probabilities of getting a, b and c, $P(A) = 0.1 + 0.05 + 0.4 = 0.55$

(b) $P(B)$

Solution:
To calculate this we add the probabilities of getting c, d and e, $P(B) = 0.4 + 0.2 + 0.25 = 0.85$. Note the events A and B are not disjoint as they both contain c

(c) $P(\overline{A})$

Solution:
This is the probability of event A so subtract the answer from part (a) from one, $P(\overline{A}) = 1 - 0.55 = 0.45$.

(d) $P(A \cup B)$

Solution:
This would be an event that contains all letters, so is the probability of getting a value in the sample space which is $P(A \cup B) = P(\Omega) = 1$.

(e) $P(A \cap B)$

Solution:
As discussed in part (b) c is the only letter in both events so this would be the probability of getting c which is $P(A \cap B) = P(c) = 0.4$.

The following R Code obtains a sample with values "1","2","3","4" chosen with weights of 0.1, 0.6,0.2,0.1 respectively:

```
> val <- 1:4
> weig <- c(0.1, 0.6,0.2,0.1)
> Wsam <- sample(val,1,prob = weig)
```

(f) Run the following R block:

```
> val <- 1:4
> weig <- c(0.1, 0.6,0.2,0.1)
> prop <- mean(replicate(10^6,any(c(1,3)==sample(val,1,prob = weig))))
> print(prop)
```

Explain what this R block is doing, i.e., what does the value of "prop" represent?

Solution:
In the code block the function checks a single randomly chosen value against the condition of equally 1 or 3, then calculating the proportion of the iterations this occurs. As these are disjoint events (a number can not be 1 and 3 at the same time) the proportion here can be calculated as

$$P(W = 1 \cup W = 3) = P(W = 1) + P(W = 3) = 0.1 + 0.2 = 0.3$$

(g) Modify the code in f) to simulate the experiment of the question (with $\Omega = \{a, b, c, d, e\}$) using 10^6 replications. Based on the simulation runs, present your estimates for the probabilities in (a)-(e) and compare them to your exact asnwers for (a)-(e).

Solution:

Below are the simulation results for each part

```
> # (a)
>
> val <- c('a','b','c','d','e')
> weig <- c(0.1, 0.05, 0.4, 0.2, 0.25)
> prop <- mean(replicate(10^6,any(c('a','b','c')==sample(val,1,prob = weig))))
> prop

[1] 0.549821

> # (b)
>
> val <- c('a','b','c','d','e')
> weig <- c(0.1, 0.05, 0.4, 0.2, 0.25)
> prop <- mean(replicate(10^6,any(c('c','d','e')==sample(val,1,prob = weig))))
> prop

[1] 0.850604

> # (c)
>
> val <- c('a','b','c','d','e')
> weig <- c(0.1, 0.05, 0.4, 0.2, 0.25)
> prop <- mean(replicate(10^6,any(c('d','e')==sample(val,1,prob = weig))))
> prop

[1] 0.449566

> # (d)
>
> val <- c('a','b','c','d','e')
> weig <- c(0.1, 0.05, 0.4, 0.2, 0.25)
> prop <- mean(replicate(10^6,any(c('a','b','c','d','e')==sample(val,1,
+ prob = weig))))
> prop

[1] 1

> # (e)
>
> val <- c('a','b','c','d','e')
> weig <- c(0.1, 0.05, 0.4, 0.2, 0.25)
> prop <- mean(replicate(10^6,any(c('c')==sample(val,1,prob = weig))))
> prop

[1] 0.399757
```

Question 4 – Car Features

BMW offers their cars as base models where clients can add different sets of features (e.g. screens on backseats, integrated GPS). Suppose, the following table lists the proportion of BMW cars bought in Australia in 2018 dependent on their features.

Features	Proportions bought in 2018
basic model	0.32
basic model with one feature	0.30
basic model with two features	0.22
basic model with three features	0.16

When answering the following questions, also formulate the events considered as sets.

(a) What is the probability that a BMW bought in 2018 has at least two features?

Solution:
For a BMW to have at least two features it has to be in the set of BMWs with two features or the set of BMW with three features. To work out the probability of this event we simply add together the proportions found in each group as the events are mutually exclusive (disjoint).

$$\begin{aligned} P(\text{At least two features}) &= P(\text{two features}) \cup P(\text{three features}) \\ &= 0.22 + 0.16 = 0.38 \end{aligned}$$

(b) What is the probability that a BMW does not contain more than one feature?

Solution:
As this is the compliment of the event above, we simply take the probability from above away from 1.

$$\begin{aligned} P(\text{not more than one feature}) &= P(\overline{\text{at least two features}}) \\ &= 1 - P(\text{at least two features}) \quad = 1 - 0.38 = 0.62. \end{aligned}$$

Question 5 – Hacking the NSA

A computer system uses passwords that contain exactly eight characters, and each character is one of 26 lower-case letters (a-z) or 26 upper-case letters (A-Z) or 10 integers (0-9). Let Ω denote the set of all possible passwords. Suppose that all passwords Ω are equally likely. Determine the probability of each of the following:

(a) The password contains only letters.

Solution:
First we need to know the number of possible passwords that can be generated with these characters. There are 62 possible characters (26+26+10=62) and each one is equally likely to be selected as a character of the password. This means there are $62^8 = 218340105584896$ possible passwords. Event A is the event of a password of only letters. The probability of any character in the password being a letter is $P(L) = \frac{52}{62}$. Each character can be considered an independent event so we raise this to the power of eight to get the probability. so

$$P(A) = \left(\frac{52}{62}\right)^8 = 0.2448.$$

(b) The password contains at least one integer

Solution:
This event is $P(\overline{A}) = 1 - P(A)$ so taking the value from part (a) we get a probability of

$$P(\overline{A}) = 0.7552$$

(c) A password contains exactly two integers and one lower-case letter.

Solution:
This is a little more involved as there are many combinations of letters and numbers that could forma password with 2 integers and one lower case letter. To work out this probability we need to first work out the probability of a password with exactly 2 integers and the probability of a password with 1 lower case letter. Consider the distribution of integers in the password as a series of Bernoulli trials. This means that we can use a Binomial Distribution with the probability of success as 0.1613. To calculate the probability of a password with exact two integers we perform the following calculation,

$$\begin{aligned} P(2 \text{ integers}) &= \binom{8}{6} P(N)^2 P(\overline{N})^6 \\ &= \frac{8!}{6!2!} 0.1613^2 \times 0.8387^6 \\ &= 28 \times 0.1613^2 \times 0.8387^6 \\ &= 0.2535 \end{aligned}$$

Now in the remaining 6 spaces we want to calculate the probability of obtaining one lower case letter. Again we use a Binomial distribution to calculate the probability, with the probability of success now being 0.5.

$$\begin{aligned} P(1 \text{ lower case}) &= \binom{6}{1} P(l)^1 P(\bar{l})^5 \\ &= 6 \times 0.5 \times 0.5^5 \\ &= 0.0938 \end{aligned}$$

As these events are independent we simply multiply the probabilities together to get the probability of both happening.

$$\begin{aligned} P(2 \text{ integers and } 1 \text{ lower case}) &= P(2 \text{ integers}) \cap P(1 \text{ lower case}) \\ &= P(2 \text{ integers}) \times P(1 \text{ lower case}) \\ &= 0.2535 \times 0.0938 \\ &= 0.0238 \end{aligned}$$

The following R code generates 100 random passwords and counts how many of them contain 1 or less lower case letters.

```
> poss_values <- c(0:9,letters[1:26],LETTERS[1:26])
>
> # function to count lower case values
> count_prop <- 0
>
```

```
> # 100 samples
> it <- 100
>
> for (i in 1:it) {

+ Sample_password <- sample(poss_values,8,replace=TRUE)

+ count_lower <- sum(Sample_password %in% letters)

+ count_prop <- count_prop + (count_lower <= 1)
+ }
>
> print(count_prop/it)


> poss_values <- c(0:9,letters[1:26],LETTERS[1:26])
>
> # function to count lower case values
> count_prop <- 0
>
> # 100 samples
> it <- 100
>
> for (i in 1:it) {

+         Sample_password <- sample(poss_values,8,replace=TRUE)

+         count_lower <- sum(Sample_password %in% letters)

+         count_prop <- count_prop + (count_lower <= 1)
+ }
>
> print(count_prop/it)

[1] 0.09
```

(d) In your view, are 100 passwords sufficient for obtaining a sensible estimate for the event of having 1 or less lower case characters? Modify the code to obtain a more accurate estimate.

Solution:
The probability of a character being a lower case character is $\frac{26}{62}$ and so the probability of obtaining a password with 1 or less lower case characters

$$\begin{aligned} P(\leq 1 \text{ lower case letter}) &= P(0 \text{ ower case letters}) + P(1 \text{ lower case letter}) \\ &= \left(1 - \frac{26}{62}\right)^8 + 8 \times \frac{26}{62} \times \left(1 - \frac{26}{62}\right)^7 \\ &= 0.0876. \end{aligned}$$

Given the size of the sample space, of the order of 10^{14}, it is not sensible to obtain an estimate from only 100 passwords. A reasonable sample would be 10^6 passwords as shown below:

```
> poss_values <- c(0:9,letters[1:26],LETTERS[1:26])
>
> # function to count lower case values
> count_prop <- 0
>
> # 10^6 samples
> it <- 10^6
>
> for (i in 1:it) {

+         Sample_password <- sample(poss_values,8,replace=TRUE)

+         count_lower <- sum(Sample_password %in% letters)

+         count_prop <- count_prop + (count_lower <= 1)
+ }
>
> print(count_prop/it)

[1] 0.087365
```

(e) Modify the code to obtain estimates for the probabilities of the events in (a)-(c). Compare it with your theoretical results.

Solution:

```
> # (a)
> poss_values <- c(0:9,letters[1:26],LETTERS[1:26])
>
> # function to count lower case values
> count_prop <- 0
>
> # 10^6 samples
> it <- 10^6
>
> for (i in 1:it) {

+         Sample_password <- sample(poss_values,8,replace=TRUE)

+         count_integer <- sum(Sample_password %in% 0:9)

+         count_prop <- count_prop + (count_integer == 0)
+ }
>
> print(count_prop/it)

[1] 0.244357

> # (b)
> poss_values <- c(0:9,letters[1:26],LETTERS[1:26])
>
```

```
> # function to count lower case values
> count_prop <- 0
>
> # 10^6 samples
> it <- 10^6
>
> for (i in 1:it) {

+         Sample_password <- sample(poss_values,8,replace=TRUE)

+         count_integer <- sum(Sample_password %in% 0:9)

+         count_prop <- count_prop + (count_integer >= 1)
+ }
>
> print(count_prop/it)

[1] 0.755766

> #(c)
> poss_values <- c(0:9,letters[1:26],LETTERS[1:26])
>
> # function to count lower case values
> count_prop <- 0
>
> # 10^6 samples
> it <- 10^6
>
> for (i in 1:it) {

+         Sample_password <- sample(poss_values,8,replace=TRUE)

+         count_integer <- sum(Sample_password %in% 0:9)
+         count_lower <- sum(Sample_password %in% letters)

+         count_prop <- count_prop + (count_integer == 2) * (count_lower == 1)
+ }
>
> print(count_prop/it)

[1] 0.023721
```

When comparing these result to those calculated analytically most results look similar. The result for part (b) however is slightly different but this is due to the level of floating point numbers in R.

Question 6 – Air Conditioning Systems

A maintenance firm has gathered the following information regarding the failure mechanisms for air conditioning systems:

		Gas Leaks	
		Yes	No
Electric Leaks	Yes	55	17
	No	32	3

Let A denote the event that an air condition has a gas leak, and let B denote the event that an air condition has an electric leak. Determine:

(a) $P(A)$

Solution: This is the probability of the air conditioner having a gas leak, so the number of air conditioners with a gas leak $(55 + 32)$ is divided by the total number of parts (107). That is

$$P(A) = \frac{55 + 32}{107} = 0.813$$

(b) $P(\overline{B})$

Solution: This is the probability of the air conditioner not having an electrical leak. 35 air conditioners do not have an electrical leak so the probability is

$$P(\overline{B}) = \frac{35}{107} = 0.327$$

(c) $P(A \mid B)$

Solution: This is the probability of an air conditioner with an electrical fault having a gas leak. Using the following formula (Bayes Rule)

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{55}{72} \approx 0.764$$

(d) $P(B \cap \overline{A})$

Solution: This is the probability of an air conditioner that does not have a gas leak and has an electrical fault.

$$P(B \cap \overline{A}) = \frac{17}{107} \approx 0.159$$

(e) If the selected air condition doesn't have a gas leak, what is the probability that it has an electric leak?

Solution: This is the probability

$$P(B \mid \overline{A}) = \frac{P(B \cap \overline{A})}{P(\overline{A})} = \frac{17}{20} \approx 0.85.$$

(f) If the selected air condition has a gas leak, what is the probability that it does not have an electric leak?

Solution: This is the probability $P(\overline{B} \mid A)$, so following the same procedure as above

$$P(\overline{B} \mid A) = \frac{P(A \cap \overline{B})}{P(A)} = \frac{32}{87} \approx 0.368.$$

Question 7 – Flaws in Research Journals

Suppose 1.8% of articles in physics journals and 3% of articles in engineering journals contain flaws. Of the articles read by a PhD student in Engineering, 38% are from physics journals and 62% are engineering articles. What is the probability that a randomly selected article by an engineering PhD student contains flaws?
Solution: Let the event that these is a flaw be F, the event that a Physics journal is selected be P and the event that an Engineering journal selected is E. The probability of a flaw can now be calculated by taking the union of the event that a flaw is present and a Physics journal is selected and the event that a flaw is present and an Engineering journal is selected. Hence:

$$\begin{aligned} P(F) &= P((F \cap P) \cup (F \cap E)) \\ &= P(F \cap P) + P(F \cap E) \quad \text{As disjoint events} \\ &= P(F|P)P(P) + P(F|E)P(E) \\ &= 0.018 \times 0.38 + 0.03 \times 0.62 \\ &= 0.025 \end{aligned}$$

Question 8 – Computer Keyboard Failure

Computer keyboard failures are due to faulty electrical connects (13%) or mechanical defects (87%). Mechanical defects are related to loose keys (22%) or improper assembly (78%), Electrical connect defects are caused by defective wires (27%), improper connections (26%), or poorly welded wires (47%).

(a) Find the probability that a failure is due to loose keys.

Solution:
As 22% of Mechanical defects are caused be loose keys and 87% of defects are mechanical to work out the probability of a failure due to loose keys we have to multiply these proportions together.

$$P(LK) = P(LK|MD)P(MD) = 0.22 \times 0.87 = 0.1914$$

(b) Find the probability that a failure is due to improperly connected or poorly welded wires

Solution:
Here we need to add the probability of a failure from a faulty electrical connect due to improperly connections and a faulty electrical connect due to poorly welded wires.

$$\begin{aligned} P(IC \cup PW) &= P(IC|FE)P(FE) + P(PW|FE)P(FE) \\ &= (P(IC|FE) + P(PW|FE))P(FE) \\ &= (0.26 + 0.47) \times 0.13 = 0.0949 \end{aligned}$$