Probability & Simulating Probabilities in R

In this assignment, you'll apply concepts of probability to real data. You'll also learn how to perform simple simulations in R. If you haven't already done so, work through the R tutorial provided on the Data Analysis 2 Canvas page. Once you've worked through the tutorial, write up your responses to the questions listed throughout the tutorial. The same questions are included below to help you format your submissions.

Submit a PDF copy of your responses to Gradescope by the deadline stated on Canvas.

Question 1 (2 points) In your simulation of 100 coin tosses, how many landed on heads? How many on tails?

Heads: 43 Tails: 57

Question 2 (3 points) In your simulation 100 tosses of the *unfair* coin, how many landed on heads? How many on tails? Do the simulated outcomes align with the expected probabilities (i.e. 20% landing on heads and 80% on tails)?

Heads: 24 Tails: 76. The simulated outcomes do not align with the expected probabilities because 20% of 100-time simulations should be 20 times heads and 80% should be 80 times tails.

Question 3 (3 points) Like coins, dice are commonly used in probability examples. Suppose we wanted to simulate the outcome of a single die roll. First, we'll consider a fair, six-sided die with sides numbered 1, 2, 3, 4, 5, and 6. A "fair" die refers to a die such that when rolled, each of the six sides has an equal likelihood of being chosen. Write a chunk of code that simulates ten rolls of a fair die. Include your code and your results of the ten simulated rolls.

Code:

```
D <- c (1, 2, 3, 4, 5, 6)
sim_fair_die<- sample (D, size = 10, replace = TRUE)
table(sim_fair_die)
```

Result:

```
1 2 3 4 6
2 3 2 1 2
```

Question 4 (3 points) Now adjust your code to simulate ten rolls of an **unfair** die. You can decide what the probabilities of rolling each face of the die are. Remember, that the total probabilities must sum to 1. In your response, include the probabilities you assigned to each outcome on the die, the code you wrote to simulate the unfair die rolls, and the results of the 10 simulated rolls.

Result:

```
1 2 3 4 5 6
2 1 4 1 1 1
```

Code:

```
D <- c(1, 2, 3, 4, 5, 6)
unfairDie <- sample (D, size = 10, replace = TRUE, prob = c(0.1, 0.1, 0.3, 0.1, 0.2, 0.2))
table(unfairDie)
```

Question 5 (7 points)

a. (2 points) Using the Law of Total probability, determine the probability of measuring wavelengths matching Ordering 1 (R1<R2<R3). That is, calculate P (Ordering 1). Show your work or include R code calculations to support your answer.

P(O1) = P(O1|G) *P(G) + P(O1|B) *P(B)
P(O1) = 0.6*0.25 + 0.1*0.75 = 0.225

b. (2 points) Suppose a rock selected from this region yielded wavelength measurements matching Ordering 1 (R1<R2<R3). Knowing this information, what is the probability the rock is granite? That is, calculate P (G|Ordering 1). Show your work or include R code calculations to support your answer.

P (Ordering 1|G) = P (G| Ordering1) *P(Ordering1)/P(G)
0.6 = P (G| Ordering1) *0.225/0.25
P (G| Ordering1) *0.225 = 0.15
P (G| Ordering1) = 0.6667

c. (3 points) Suppose a rock selected from this region yielded wavelength measurements matching Ordering 3 (R3<R1<R2). Knowing this information, should you classify the rock as granite or basalt? Show your work or include R code calculations to support your answer. *Hint: compare* P (G|Ordering 3) *and* P (B|Ordering 3).

P (G| Ordering 3) = P (Ordering 3|G) *P(G)/P (Ordering 3)
P (B| Ordering 3) = P(Ordering3|B) *P(B)/P(Ordering3)
P (Ordering 3) = P (Ordering 3|G) *P(G) + P (Ordering 3|B) *P(B)
P (Ordering 3) = 0.15*0.25 + 0.70*0.75 = 0.5625
P (G| Ordering 3) = 0.15*0.25/0.5625 = 0.0667
P (B| Ordering 3) = 0.70*0.75/0.5625 = 0.9333
According to my results that P (B| Ordering 3) is greater than P (G| Ordering 3), thus the rock should be basalt.

Question 6 (5 points)

a. The next purchaser will request both of the two options.
P (A and B) = P(A)+P(B)-P (A or B) =0.55+0.4-0.63 = 0.32
b. The next purchaser will request neither of the two options.
P(neither) = p(A)-(P A or B) = 1-0.63 = 0.37
c. The next purchaser will request only Apple Carplay and not a backup camera.
P (A not B) = P(A) – P (A and B) = 0.55 – 0.32 = 0.23
d. The next purchaser will request only a backup camera and not Apple Carplay.
P (B not A) = P(B) – P (A and B) = 0.40 – 0.32= 0.08
e. The next purchaser will request exactly one of these two options.
P (exactly one) = (P(A) + P(B))- 2P (A and B) = (0.55 + 0.40) – (2*0.32) = 0.31

Gradescope Page Matching (2 points)

When you upload your PDF file to Gradescope, you will need to match each question on this assignment to the correct pages. Video instructions for doing this are available in the Start Here module on Canvas on the page “Submitting Assignments in Gradescope”. Failure to follow these instructions will result in a 2-point deduction on your assignment grade. Match this page to outline item “Gradescope Page Matching”.