Chapters covered: Chapter 5

Show all of your work to receive full credit.

Textbook Problems : (Statistics, the art and science of learning from data, 4th)

- **Problem 1: Exercise 5.16**
- **Problem 2: Exercise 5.36**
- **Problem 3: Exercise 5.37**
- **Problem 4: Exercise 5.42**
- **Problem 5: Exercise 5.66**
- **Problem 6: Exercise 5.73** You do not need to draw tree diagram for part a.
- **Problem 7: A pair of balanced dice (numbered 1-6 on each side) are rolled with outcomes defined by (die 1, die 2), e.g. (4,2) means a 4 on die 1 and a 2 on die 2.**
  a. Let A be the event of getting the same outcome for both dice. List the outcome and find P(A).
  b. Let B be the event that the sum of the dice is 7. List the outcome and find P(B).
  c. Let C be the event that the sum of the dice is odd. Find P(C).
  d. Find P(B and C). What is the relationship between P(B and C) and P(B)? Explain why.
  e. Find P(B or C). What is the relationship between P(B or C) and P(C)? Explain why.

R Problem

Consider the Georgia student survey dataset from the textbook, Statistics: The Art and Science of Learning from Data by Agresti, Franklin, and Klingenberg, 4th edition. The csv file of the dataset can be found on canvas. Download the dataset and use the read.csv() command to load the dataset in R.

a. What is the sample size and number of variables for this dataset? Show your R code.

b. What is the average study time among female students (suppose Gender=1 represents a female student) in this survey? Show your R code.

For question c-d, we examine the distribution of amount of haircuts among the students.

c. What is the probability that a randomly chosen student has over 10 haircuts?

d. What is the probability that a randomly chosen student is female (suppose Gender=1 represents a female student) and has over 10 haircuts?

For question e-h, we examine the joint distribution gender variable (whether the student is male or female) and the job variable (whether the student has a job). Use the following code to obtain the cross table for the distributions of the gender and job variables (suppose Gender=1 represents a female student, and Job=1 represents the student has a job):

```
(joint_freq<-table(data$Gender,data$Job))
```

e. What is the probability that a randomly chosen student has a job?

f. What is the probability that a randomly chosen student has a job given that the student is male?

g. Are the event that the student has a job and the student is male independent? Explain.

h. What is the probability that a randomly selected student has a job or is male (or both)?

REFERENCE: Agresti, A. and Franklin, C. (2017). Statistics: the art and science of learning from data. Pearson, 4th ed.