Matlab Assignment 4: Conditional Probability

I've given you a datafile on Canvas called ECE2523_Project4_Data.mat. You are going to manipulate this dataset. This assignment is due **Friday, October 15, 2021, at 11:59 PM**. It will be turned in electronically using Canvas.

Project To-Do List Summary

To summarize the long description below, do these following steps:

1. Write a main script called "LastnameFirstname_Project_4.m" that
    a. Loads a data file with two different datasets
    b. Calls the function "CalcMKJointProb" to calculate 9 different probabilities corresponding to the intersection of different events
2. Write a function "CalcMKJointProb " that
    a. Takes a matrix of observations, where the row corresponds to an observation and each column corresponds to an **observed pair of events**
    b. Calculates the probability of each joint observation
        i. 3 unique possible observations in column 1 combined with 3 unique possible observations in column 2 means 9 joint probabilities to estimate
        ii. Each probability is found by counting the number of times each event is observed and divided by the total number of observations
    c. Returns the probability **matrix** as the output of the function

For further description, see the individual sections.

Part 1: Loading the data

Create a new script in Matlab. Name it "LastnameFirstname_Project_4.m". Save this file, and in the same directory download the file "ECE2523_Project4_Data.mat" from Canvas. At the very top of your "LastnameFirstname_Project_4.m" script, put your name and the date as comments. Below those comments, on a single line, use the command "*load ECE2523_Project4_Data;*".

Part 2: Manipulating the data

Introduction

The OU eSports team has been practicing the racing game *MK*, where three players race around a track. Periodically each player receives an item they may use to disrupt the other players. The possible items are: {blue shell, red shell, banana}. The probability of getting each item depends on the players place in the race (where 1st place is in the lead and 3rd place is last place). The OU eSports team has been trying to understand the probability of getting each item and has been logging the items they have received during races.

The team has collected the data into an $N \times 2$ matrix named "ouData", where the N *rows* of the matrix correspond to each observation made by the team. The *first column* of the matrix corresponds to the place the player was in, with possible values $\{1,2,3\}$. The *second column* of the matrix corresponds to the item the player obtained, with possible values $\{1,2,3\}$, where a value of 1 denotes the player received a blue shell, a value of 2 denotes the player received a red shell, and a 3 denotes the player received a banana. For example, consider this matrix:

$$exampleMatrix = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

The example matrix has 3 different observations (i.e. 3 rows). In the first observation, the player was in first place (i.e. exampleMatrix(1,1)=1) **and** they obtained a banana (i.e. exampleMatrix(1,2)=3). In the second observation, the player was in second place (i.e. exampleMatrix(2,1)=2) **and** they obtained a blue shell (i.e. exampleMatrix(2,2)=1). In the third observation, the player was in third place (i.e. exampleMatrix(3,1)=3) **and** they obtained a red shell (i.e. exampleMatrix(3,2)=2).

Task

In a *new file* named "CalcMKJointProb.m" make a function whose first line is "function [jointProb] = CalcMKJointProb (obsData)", where obsData is a data vector (with **any length**). The output jointProb will be a 3×3 matrix with entries that correspond to the probabilities of the following *intersecting events*:

	Blue Shell	Red Shell	Banana
1st Place	$P[1^{st} \cap BS]$	$P[1^{st} \cap RS]$	$P[1^{st} \cap B]$
2nd Place	$P[2^{nd} \cap BS]$	$P[2^{nd} \cap RS]$	$P[2^{nd} \cap B]$
3rd Place	$P[3^{rd} \cap BS]$	$P[3^{rd} \cap RS]$	$P[3^{rd} \cap B]$

For discussion on interpreting the probabilities of intersecting events, it might be helpful to review slides 46-59 in Lecture 1.3. Note, the table above is for illustration. You will fill out two tables below in Part 3.

There are MANY ways to program the function CalcMKJointProb – I have 2 different examples in my solution. Therefore, write the function so that it mimics how you would solve this problem if you had to solve it by hand. As long as you get the correct answer, you will get full points – I'm not looking for a particular solution. The "easiest" solution may involve nested loops (i.e. loops within loops), but it is possible to solve this problem with a single loop. In my solution I will show two different solutions, one with three nested loops and one with a single loop. **Take your time and think through the problem on paper before starting your code.**

Part 3: Answer Questions

Use your function to fill out the following table using the data stored in the vector "ouData". Note that each entry in the table is the probability of the intersection of the two events.

	Blue Shell	Red Shell	Banana
1st Place			
2nd Place			
3rd Place			

Use this table and your knowledge of probability to answer the following questions (**enter the answer below – don't use fprintf**):

1. What is the probability the OU team is in first place?

2. What is the probability the OU team obtained a blue shell?

3. *Given* the OU team obtained a blue shell, what is the probability they are in 2nd place?

Next, use the function CalcMKJointProb to calculate the joint probabilities of the dataset "txData" and fill in the table below. This dataset was collected by the University of Texas eSports team.

	Blue Shell	Red Shell	Banana
1st Place			
2nd Place			
3rd Place			

Use the answers from the Texas joint probability table to answer the following questions:

4. What is the probability the Texas team is in second place?

5. What is the probability the Texas team obtained a blue shell?

6. *Given* the Texas team obtained a blue shell, what is the probability they are in 3rd place?

7. From a *statistical perspective* do you think there is evidence that the OU team could beat the Texas team? Why or why not?

Comment all of your code. Comments will be a portion of your grade.

Turn in checklist:

3 separate files:

1. Main script called "LastnameFirstname_Project_4.m" (using your last name and first name)
2. A function file called "CalcMKJointProb.m"
3. A Word or PDF document called "LastnameFirstname_Project_4"

Please **zip** all three files together (i.e. compress the files into a single .zip file called LastnameFirstname_Project_4.zip). and submit **electronically** using Canvas. The submission will **close at 11:59**, so please turn this in on time! Note: please do not .rar the files – make sure to use zip files.