Math 151 - Probability Theory - Homework 7

your name here

Due: Friday, October 9, 2020, midnight PDT

Important Note:

You should work to turn in assignments that are clear, communicative, and concise. Part of what you need to do is not print pages and pages of output. Additionally, you should remove these exact sentences and the information about HW scoring below.

Click on the *Knit to PDF* icon at the top of R Studio to run the R code and create a PDF document simultaneously. [PDF will only work if either (1) you are using R on the network, or (2) you have LaTeX installed on your computer. Lightweight LaTeX installation here: https://yihui.name/tinytex/]

> Either use the college's RStudio server (https://rstudio.pomona.edu/) or install R and R Studio on to your personal computer. See: https://research.pomona.edu/johardin/math151f20/ for resources.

Assignment

1: PodQ

Describe one thing you learned from someone in your pod this week (it could be: content, logistical help, background material, R information, etc.) 1-3 sentences.

2: 3.11.13

Suppose that the joint distribution of X and Y is uniform over a set A in the xy-plane. For which of the following sets A are X and Y independent?

a. A circle with a radius of 1 and with its center at the origin.
b. A circle with a radius of 1 and with its center at the point $(3, 5)$.
c. A square with vertices at the four points $(1, 1)$, $(1, -1)$, $(-1, -1)$, $(-1, 1)$.
d. A rectangle with vertices at the four points $(0, 0), (0, 3), (1, 3)$, and $(1, 0)$.
e. A square with vertices at the four points $(0, 0), (1, 1), (0, 2)$, and $(-1, 1)$.

3: 3.6.2

Each student in a certain high school was classified according to their year in school (freshman, sophomore, junior, or senior) and according to the number of times that they had visited a certain museum (never, once, or more than once). The proportions of students in the various classifications are given in the following table. (Note: the solution to this problem is a calculation of parts of the discrete conditional distribution. You'd need to calculate all the possible options to arrive at the complete discrete conditional pf.)

	Never	Once	$>$ Once
Freshmen	0.08	0.10	0.04
Sophomores	0.04	0.10	0.04
Juniors	0.04	0.20	0.09
Seniors	0.02	0.15	0.10

a. If a student selected at random from the high school is a junior, what is the probability that they have never visited the museum?

b. If a student selected at random from the high school has visited the museum three times, what is the probability that they are a senior?

4: 3.6.8

Suppose that a person's score X on a mathematics aptitude test is a number between 0 and 1, and that the score Y on a music aptitude test is also a number between 0 and 1. Suppose further that in the population of all college students in the United States, the scores X and Y are distributed according to the following joint p.d.f.:

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y) & \text{for } 0 \le x \le 1 \text{ and } 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

a. What proportion of college students obtain a score greater than 0.8 on the mathematics test?

b. If a student's score on the music test is 0.3, what is the probability that his score on the mathematics test will be greater than 0.8?

c. If a student's score on the mathematics test is 0.3, what is the probability that his score on the music test will be greater than 0.8?

5: 3.7.2

Suppose that the three random variables X_1, X_2 and X_3 have a mixed joint distribution with p.f./p.d.f.:

$$f(x_1,x_2,x_3) = \begin{cases} cx_1^{1+x_2+x_3}(1-x_1)^{3-x_2-x_3} & \text{if } 0 < x_1 < 1 \text{ and } x_2, x_3 \in \{0,1\} \\ 0 & \text{otherwise} \end{cases}$$

(Notice that X_1 has a continuous distribution and X_2 and X_3 have discrete distributions.) Determine

a. the value of the constant, c;
b. the marginal joint p.f. of X_2 and X_3; and
c. the conditional p.d.f. of X_1 given $X_2 = 1$ and $X_3 = 1$.

6: R - breaking a stick

(Note: in the interst of this being a long week, I've done the majority of the R coding for you. There are three places where you need to replace the word `TRUE` with the appropriate condition. Additionally, you should make sure that you can follow the code I've written.)

Consider the problem where a stick of length 1 is broken in two places and the task is to determine the probability of the three pieces making a triangle. In this problem the stick is broken in two different ways:

Scenario 1. By choosing the breaks randomly (i.e., choosing two real numbers independently and uniformly from [0,1].)

Scenario 2. By first breaking the stick at random, and then breaking the longer piece at random.

Show that the two probabilities are quite different as follows:

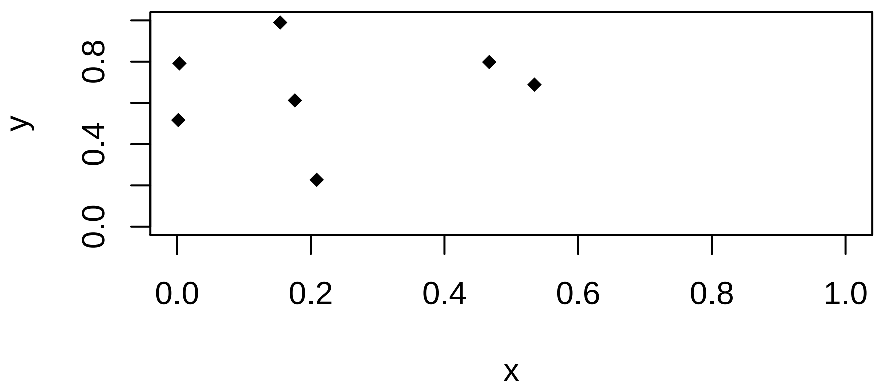
a. Write a program which simulates both cases for a run of 1000 trials, prints out the proportion of successes for each run. (Call a trial a success if the three pieces do form a triangle.) Have your program pick (x, y) at random in the unit square, and in each case use x and y to find the two breaks. For each experiment, have it plot (x, y) if (x, y) gives a success.

b. Show that in the first experiment the probability is $1/4$ and in the second experiment the theoretical probability of success is $2\ln(2) - 1$.

Hint: the joint pdf for b. is no longer the product of the marginals. Why not?

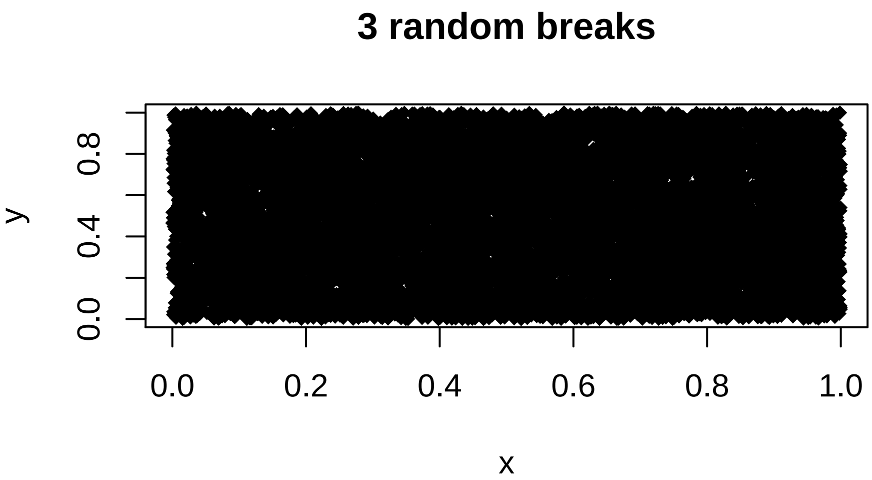
Here is some sample code to count some things. The first R chunck isn't really related to the problem.

```
count=0
nreps=10

plot(c(0,1), c(0,1), type = "n", xlab = "x", ylab = "y")

for(i in 1:nreps){
  xval=runif(1,0,1)
  yval=runif(1,0,1)

  if (xval < yval){
    count = count+1
    points(xval,yval, pch=18)}

}
```




```
#scenario 1
count = 0
nreps = 10000
plot(c(0,1), c(0,1), type = "n", xlab = "x", ylab = "y", main="3 random breaks")

xval = runif(nreps,0,1)
yval = runif(nreps,0,1)

for (i in 1: nreps){
  yval = runif(1,0,1)
  xval = runif(1,0,1)

  aval = min(xval,yval)
  bval = max(xval,yval) - min(xval,yval)
  cval = 1 - max(xval,yval)

  vals = sort(c(aval,bval,cval))  # what does line do?  What is `vals`?


  ##  vals[1] and vals[2] and vals[3] are three different objects (or numbers).
  ##  use them inside the `if()` statement below
  ##  delete the word TRUE and write the appropriate condition
```

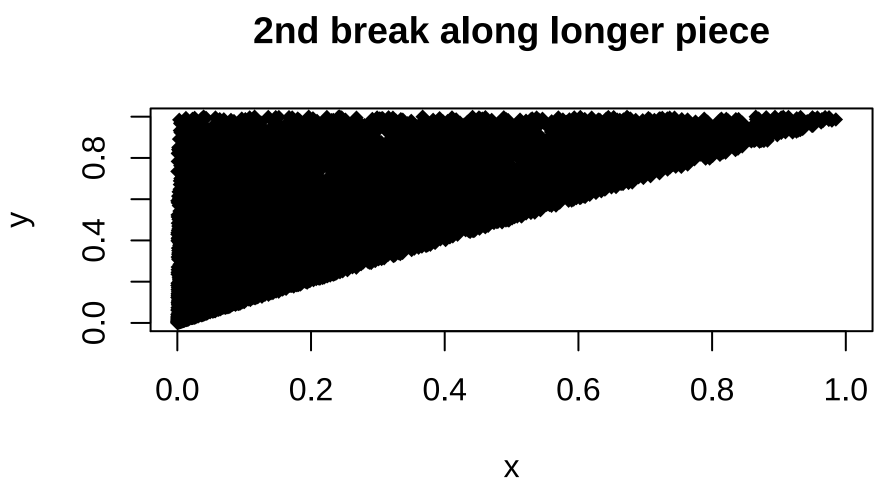
```
  if(TRUE){
    count = count+1
    points(xval,yval, pch=18)
  }
  }
```




```
count/nreps
```

```
## [1] 1
```

```
# scenario 2
count=0
plot(c(0,1), c(0,1), type = "n", xlab = "x", ylab = "y", main="2nd break along longer piece")

for (i in 1: nreps){
  yval = runif(1,0,1)

  ##  delete the word TRUE and write the appropriate condition
  ##  which fulfills the scenario 2 setting.

  if (TRUE){
    xval = runif(1,0, yval) }else{
    xval = runif(1,0,1-yval) +yval
    }

  aval = min(xval,yval)
  bval = max(xval,yval) - min(xval,yval)
  cval = 1 - max(xval,yval)

  vals = sort(c(aval,bval,cval))

  ## use the same condition from scenario 1, replace the word TRUE
  if(TRUE){
    count = count+1
    points(xval,yval, pch=18)}}
```




```
count/nreps
```

```
## [1] 1
```

```
2*log(2) - 1
```

```
## [1] 0.3862944
```