## 14.310x: Data Analysis for Social Scientists
### Describing Data, Joint and Conditional Distributions – Part II

Suppose two sisters, Caroline and Anna, sleep in adjoining rooms. Each has a speaker with which she plays music, and each speaker has a volume dial going from 0 to 1. The joint distribution of the volumes of the two speakers is $f_{XY}(x, y) = c(x + y^2)$ over the unit square, 0 otherwise. Caroline's volume is denoted by X, Anna's by Y.
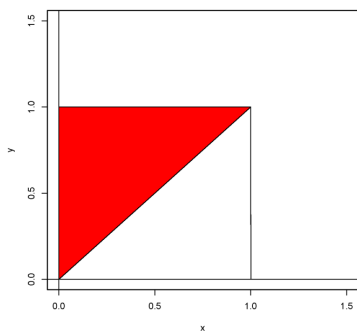
Question 1
Which of the following figures represent the domain (in red) in which the density function is defined as $f_{XY}(x, y) = c(x + y^2)$?
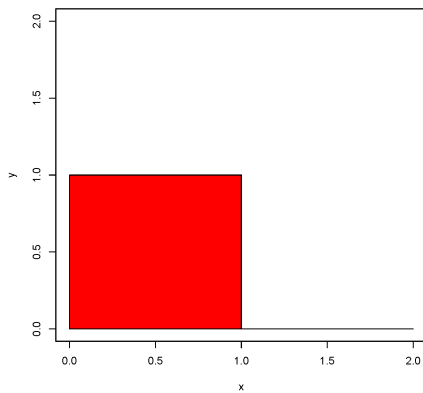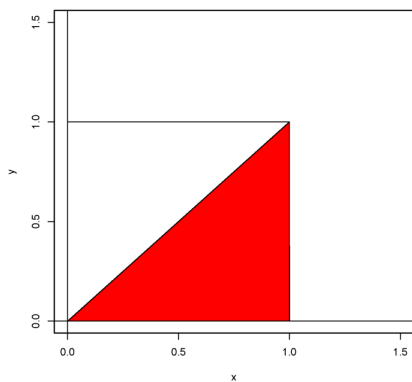
○

○
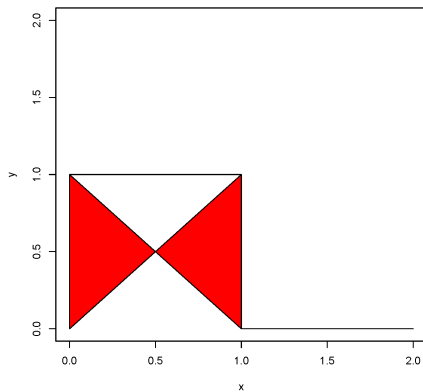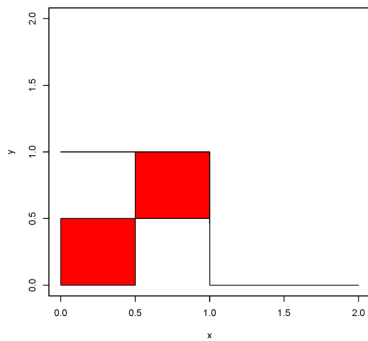
Question 2
What does the constant c represent? (Select all that apply)

    The constant c is a parameter whose value assures that the joint PDF integrates to 1.

    The constant c represents a parameter that changes both the joint PDF and the joint CDF of the random variables X and Y.

    The constant c is an irrelevant parameter in the shape of the joint CDF of the random variables X and Y.

    The constant c is a parameter that helps to infer whether the random variables X and Y are independent.
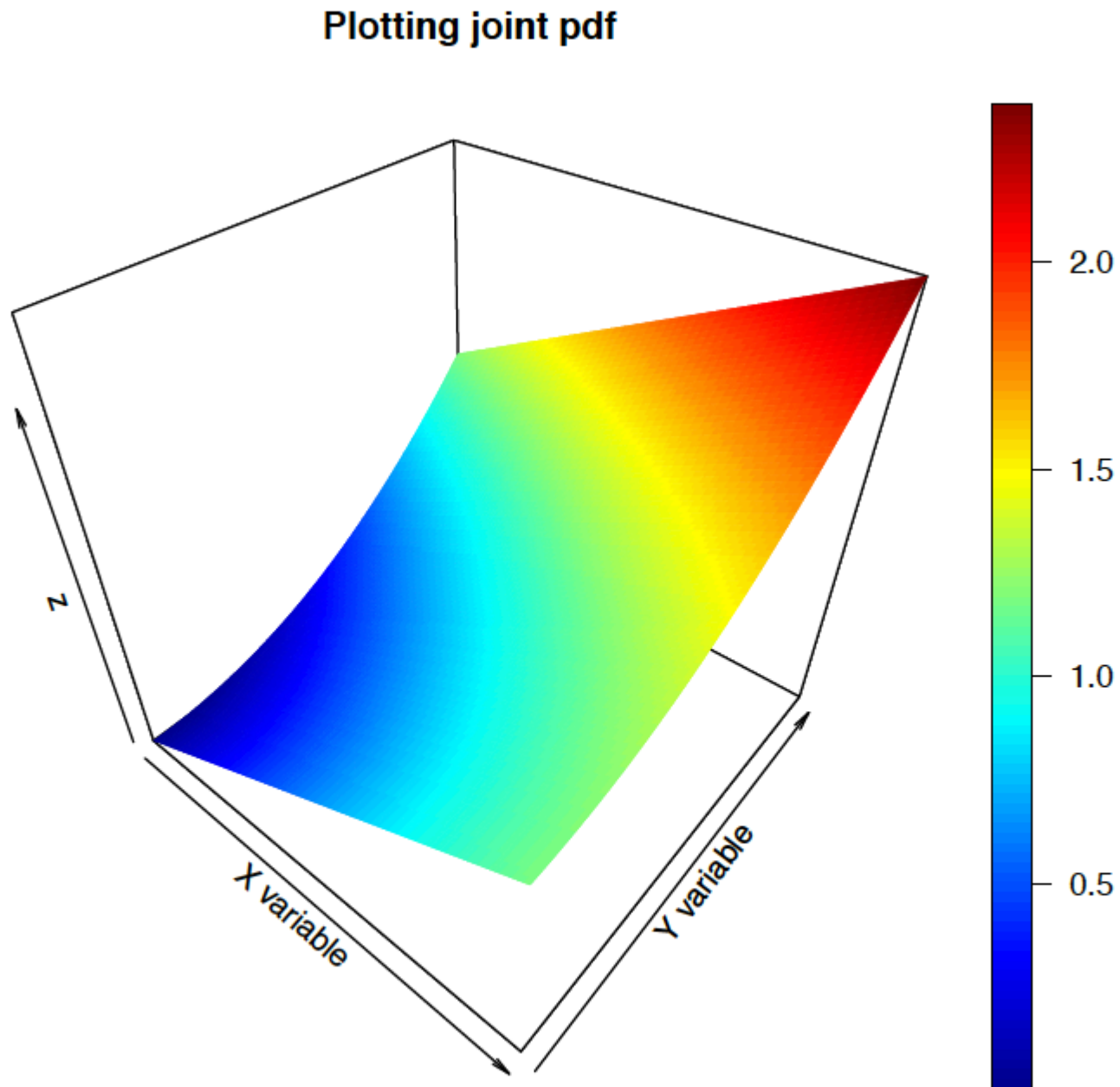
Question 3
What is the value of the constant c in this case?

Now we are going to work in R to plot the bivariate PDF. Download the code here and take a look at the following code in order to create a grid and a 3-dimensional plot of the PDF. Please note that you might need to install the package `plot3D`.

Question 4
The following plot was created by running the code. A student is claiming that this plot is wrong since there are certain regions in which the PDF shows values larger than 1. Is this student

correct that there is a mistake and therefore the plot does not correspond to the information given in the problem?

Plotting joint pdf

○ Yes
○ No

Question 5
Are the volumes of the two speakers independent random variables?
- o Yes
- o No

Question 6
Recall that Caroline's speaker volume is denoted by X and Anna's speaker volume is denoted by Y. What is the formula for the marginal distribution of Anna's speaker volume?

- o $\frac{5}{6}\left(\frac{1}{2} + y^2\right)$
- o $\frac{6}{5}\left(\frac{1}{2} + y^2\right)$
- o $\frac{6}{5}\left(\frac{1}{2} + \sqrt{y}\right)$
- o $\frac{5}{6}\left(\frac{1}{2} + \sqrt{y}\right)$

Question 7
Recall that Caroline's volume is denoted by X and Anna's volume is denoted by Y. What is the conditional distribution of Caroline's volume as a function of Anna's?

- o $\dfrac{x+y^2}{\frac{1}{2}+y^2}$
- o $\dfrac{\frac{5}{6}(x+y^2)}{\frac{6}{5}\left(\frac{1}{2}+y^2\right)}$
- o $\dfrac{x+\sqrt{y}}{\frac{1}{2}+y^2}$
- o $\dfrac{\frac{6}{5}(x+y^2)}{\frac{1}{2}+y^2}$

Question 8
From this conditional distribution can you infer whether Caroline likes Anna's music or not? Hint: Think about the probability that Caroline's volume is high when the volume of Anna's music increases.
- o Caroline does like Anna's music
- o Caroline does not like Anna's music

Question 9
What is the probability that Caroline's volume is less than ½ if Anna's volume is ½?

Question 10

Recall that Caroline's speaker volume is denoted by X and Anna's speaker volume is denoted by Y. What is the marginal distribution of Caroline's speaker volume?

- $\dfrac{5}{6}\left(x + \dfrac{2}{3}\right)$
- $\dfrac{5}{6}\left(x + \dfrac{1}{3}\right)$
- $\dfrac{6}{5}\left(x + \dfrac{2}{3}\right)$
- $\dfrac{6}{5}\left(x + \dfrac{1}{3}\right)$

Question 11

Is there a First Order Stochastic Dominance (FOSD) relationship between the random variables X and Y? (We suggest you compute the CDFs of both variables and plot them in R.)

- The distribution of X FOSD the distribution of Y
- The distribution of Y FOSD the distribution of X
- There is no clear relationship

Question 12

From this information, does Anna or Caroline prefer higher volumes? Hint: Think about the implications of the answer to Question 11)

- Anna
- Caroline
- We can't say