Sample: Statistics and Probability - Jointly Continuous Random Variables

Question 1

Let X and Y be jointly continuous random variables with joint density function

$$f(x,y) = c(y^2 - x^2)e^{-y}, \quad -y \le x \le y, \qquad 0 < y < \infty.$$

a) Find c so that f is a density function.

b) Find the marginal densities of X and Y.

c) Find the expected value of X.

Solution.

(a) If f is a true density function the following must be true:

$$\int_D f(x,y) = 1$$

Where D is domain of the function.

In our case the equality looks:

$$\int_0^\infty \int_{-y}^{y} c(y^2 - x^2)e^{-y} dx\, dy = 1$$

Solve the equation (use integration by parts) to get c:

$$\int_0^\infty \int_{-y}^{y} c(y^2 - x^2)e^{-y} dx\, dy = c\int_0^\infty e^{-y} \int_{-y}^{y} (y^2 - x^2) dx\, dy = c\int_0^\infty e^{-y} * \left(y^2 x - \frac{x^3}{3}\right)_{-y}^{y} dy$$

$$= c\int_0^\infty e^{-y}\left(y^2 * y - \frac{y^3}{3} - y^2 * (-y) + \frac{(-y)^3}{3}\right) dy = c\int_0^\infty e^{-y}\left(\frac{4}{3}y^3\right) dy$$

$$= \frac{4}{3}c\int_0^\infty y^3 e^{-y} dy = \left| \begin{matrix} let\ u = y^3, dv = e^{-y}dy \\ => du = 3y^2 dy, v = -e^{-y} \end{matrix} \right|$$

$$= \frac{4}{3}c\left((-y^3 e^{-y})_0^\infty + 3\int_0^\infty y^2 e^{-y} dy\right) = \left| \begin{matrix} let\ u = y^2, dv = e^{-y}dy \\ => du = 2y dy, v = -e^{-y} \end{matrix} \right|$$

$$= \frac{4}{3}c\left((-y^3 e^{-y})_0^\infty + 3(-y^2 e^{-y})_0^\infty + 3 * 2\int_0^\infty y e^{-y} dy\right)$$

$$= \left| \begin{matrix} let\ u = y, dv = e^{-y}dy \\ => du = dy, v = -e^{-y} \end{matrix} \right|$$

$$= \frac{4}{3}c\left((-y^3 e^{-y})_0^\infty + 3(-y^2 e^{-y})_0^\infty + 6(-y e^{-y})_0^\infty + 6\int_0^\infty e^{-y} dy\right) =$$

$$= \frac{4}{3}c(-y^3 e^{-y} - 3y^2 e^{-y} - 6y e^{-y} - 6e^{-y})_0^\infty$$

$$= \frac{4}{3}c(0 + 0 + 0 + 0 - 0 - 0 - 0 - (-6)) = 8c = 1$$

Thus, the solution is:

$$c = \frac{1}{8}$$

(b)

$$f_X(x) = \int_Y f(x,y)dy = \int_0^\infty c(y^2 - x^2)e^{-y}dy = \frac{1}{8}\int_0^\infty (y^2 - x^2)e^{-y}dy$$

$$= \left| \begin{matrix} let\ u = y^2 - x^2, dv = \ e^{-y}dy \\ => du = 2ydy, v = -e^{-y} \end{matrix} \right| = \frac{1}{8}\left((-(y^2 - x^2)e^{-y})_0^\infty + 2\int_0^\infty ye^{-y}dy \right)$$

$$= \left| \begin{matrix} let\ u = y, dv = \ e^{-y}dy \\ => du = dy, v = -e^{-y} \end{matrix} \right|$$

$$= \frac{1}{8}\left((-(y^2 - x^2)e^{-y})_0^\infty + 2(-ye^{-y})_0^\infty + 2\int_0^\infty e^{-y}dy \right)$$

$$= \frac{1}{8}(-(y^2 - x^2)e^{-y} - 2ye^{-y} - 2e^{-y})_0^\infty$$

$$= \frac{1}{8}(0 + 0 + 0 + (0 + x^2) * 1 + 0 + 2 * 1) = \frac{x^2 + 2}{8}, -y \le x \le y$$

$$f_Y(y) = \int_X f(x,y)dx = \int_{-y}^{y} c(y^2 - x^2)e^{-y}dx = \frac{1}{8}e^{-y}\int_{-y}^{y}(y^2 - x^2)dx$$

$$= \frac{1}{8}e^{-y} * \left(y^2x - \frac{x^3}{3} \right)_{-y}^{y} = \frac{1}{8}e^{-y}\left(y^2 * y - \frac{y^3}{3} - y^2 * (-y) + \frac{(-y)^3}{3} \right)$$

$$= \frac{1}{8}e^{-y} * \frac{4}{3}y^3 = \frac{y^3e^{-y}}{6}, y \ge 0$$

(c)

$$E(X) = \int_X xf_X(x)dx = \int_{-y}^{y} x * \frac{x^2 + 2}{8}dx = \frac{1}{8}\int_{-y}^{y}(x^3 + 2x)dx = \frac{1}{8}\left(\frac{x^4}{4} + x^2 \right)_{-y}^{y}$$

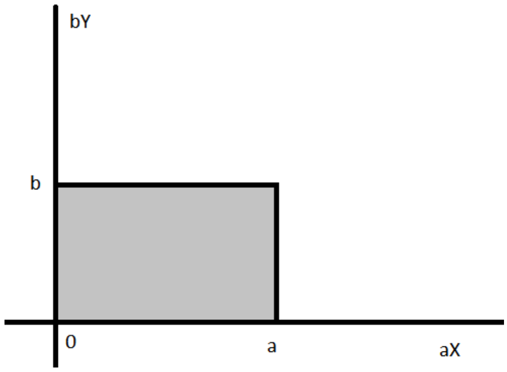
$$= \frac{1}{8}\left(\frac{y^4}{4} + y^2 - \frac{(-y)^4}{4} - (-y)^2 \right) = 0$$

Question 2

Let X and Y be independent standard uniform random variables and let a, b and c be positive real numbers. Find the probability that $aX + bY \leq c$.

Solution.

X and Y are uniformly distributed in the interval [0, 1]. Thus, aX and bY are uniformly distributed in the intervals [0, a] and [0, b] correspondently. Thus, the variable (X, Y) is uniformly distributed in the following rectangle:



The condition $aX + bY \leq c$ corresponds to the following one:

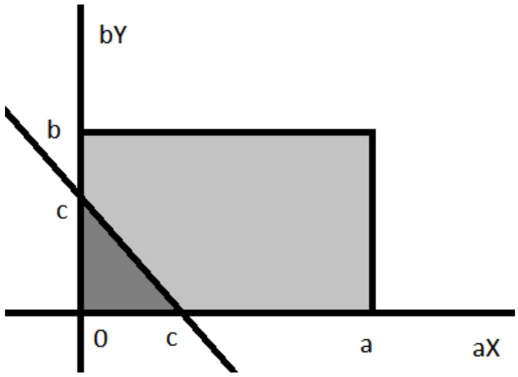
$$bY \leq -aX + c$$

Or, graphically, bY locates under the line bY = -aX + c.

The corresponding probability equals to percentage of rectangle that locates under the line bY = -aX + c.

Consider the possible cases of relations between a, b and c and find the area in each case.

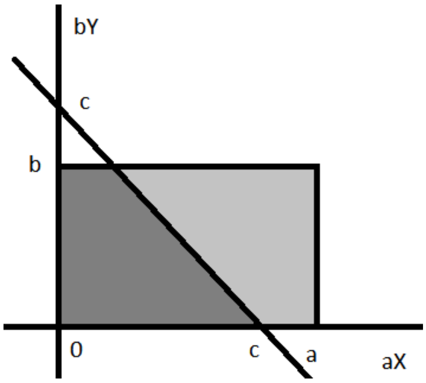
Case 1: $c \leq a \text{ and } c \leq b$



The area under the line equals to area of a right triangle with cathetus of length c:

$$S1 = \frac{c^2}{2}$$

Case 2: $b < c < a$



The area under the line equals to area of a right triangle with cathetus of length c minus area of a right triangle with cathetus of length (c-b):

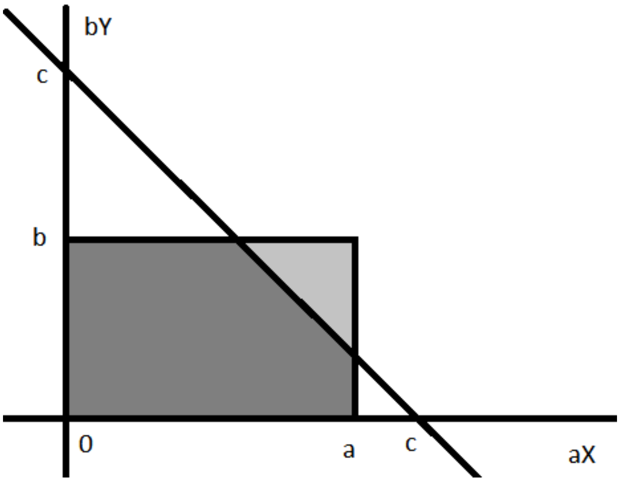
$$S2 = \frac{c^2 - (c-b)^2}{2} = \frac{2bc - b^2}{2}$$

Case 3:$a < c < b$

The figure for this case will be symmetrical to case 2 figure. The corresponding formulas for area are the same too, just switch a and b:

$$S3 = \frac{c^2 - (c-a)^2}{2} = \frac{2ac - a^2}{2}$$

Case 4:$b < a < c \leq a + b$



The area under the line equals to area of a right triangle with cathetus of length c minus area of a right triangle with cathetus of length (c-b) and minus area of a right triangle with cathetus of length (c-a):

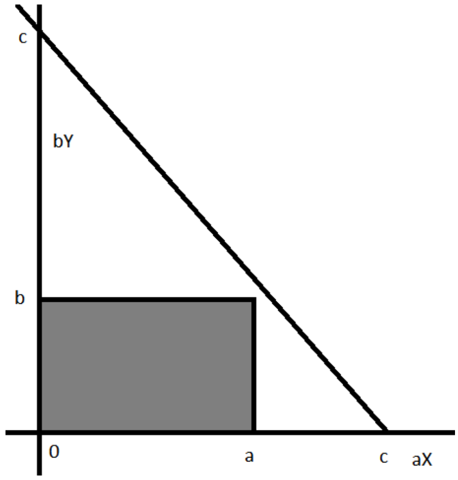
$$S4 = \frac{c^2 - (c-b)^2 - (c-a)^2}{2}$$

Case 5:$a < b < c \leq a + b$

The figure for this case will be symmetrical to case 4 figure. The corresponding formulas for area are the same too, just switch a and b:

$$S5 = \frac{c^2 - (c-a)^2 - (c-b)^2}{2}$$

Case 6: $c > a + b$



In this case the whole rectangle will locate under the line:

$$S6 = ab$$

Summarize the areas found to build one function:

$$S = \begin{cases} \frac{c^2}{2}, if\ c \le a\ and\ c \le b \\ \frac{c^2 - (c-b)^2}{2}, if\ b < c \le a \\ \frac{c^2 - (c-a)^2}{2}, if\ a < c \le b \\ \frac{c^2 - (c-b)^2 - (c-a)^2}{2}, if\ a < b < c \le a+b\ or\ b < a < c \le a+b \\ ab, if\ c > a+b \end{cases}$$

Combine and transform some of the cases to get more compact form:

$$S = \begin{cases} \frac{c^2}{2}, & if\ c \le \min(a,b) \\ \frac{c^2 - (c - min(a,b))^2}{2}, & if\ \min(a,b) < c \le \max(a,b) \\ \frac{c^2 - (c-b)^2 - (c-a)^2}{2}, & if\ \max(a,b) < c \le a+b \\ ab, & if\ c > a+b \end{cases}$$

Area of the rectangle:

$$S_{full} = ab$$

Use the formulas for areas to find the probability for each case:

$$P(aX+bY\le c)=\begin{cases}\dfrac{c^2}{2ab}, & if\ c\le\min(a,b)\\ \dfrac{c^2-(c-min(a,b))^{\,2}}{2ab}, & if\ \min(a,b)<c\le\max(a,b)\\ \dfrac{c^2-(c-b)^2-(c-a)^2}{2ab}, & if\ \max(a,b)<c\le a+b\\ 1, & if\ c>a+b\end{cases}$$

Question 3

Show that if X and Y are jointly continuous, then $X+Y$ is a continuous random variable while X, Y and $X+Y$ are not jointly continuous.

Solution.

If X and Y are jointly continuous random variables, there exists a continuous density function $f_{XY}(x,y)$ such that

$$P(X\le s, Y\le t)=\int\limits_{x\le s,\ y\le t} f_{XY}(x,y)dxdy$$

Now, consider the random variable $X+Y$. Consider the following probability.

$$P(X+Y\le a)=\int\limits_{s\le a} P(X\le s, Y\le a-s)ds=\int\limits_{s\le a}\ \int\limits_{x\le s,\ y\le a-s} f_{XY}(x,y)dxdy\,ds$$

The function $f_{XY}(x,y)$ is continuous in R^2. Thus, the integral above has a clear geometrical sense – volume of the curvilinear cone. Thus, the probability considered exists and is continuous for such X and Y. So, $X+Y$ is a continuous variable.

Now assume that X, Y and $X+Y$ are jointly continuous. In this case there must exist a function $f_{XY,X+Y}(x,y,x+y)$ such that

$$Pj=P(X\le s, Y\le t, X+Y\le a)=\int\limits_{x\le s,\ y\le t,\ x+y\le a} f_{XY,X+Y}(x,y,x+y)dxdy$$

When looking at the formula above we can understand that the conditions $x\le s,\ \ y\le t,\ \ \ x+y\le a$ are not independent. There are “border” points where the final equation will change its shape.

For example, assume s and t increase from some point and tend to the line $s+t\ =\ a$. Below this line $(s+t=a-\epsilon)$ the probability Pj will exist and will be non-zero in general case. But just above the line $(s+t=a+\epsilon)$ we are sure to get $Pj=0$, because if $s\ +\ t\ >\ a$ the events $x\le s,\ \ y\le t,\ \ x+y\le a$ will never occur simultaneously.

As we can see, Pj will have a “jump” in the set of points $s+t\ =\ a$. Thus, the probability is not continuous and so, X, Y and $X+Y$ are not jointly continuous.