Sample: Statistics and Probability - Jointly Continuous Random Variables

**Question 1**
Let $X$ and $Y$ be jointly continuous random variables with joint density function

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

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:

$$\int_{-y}^{y} \int_{0}^{\infty} c(y^2 - x^2)e^{-y} dx \, dy = 1$$

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

$$\int_{-y}^{y} \int_{0}^{\infty} c(y^2 - x^2)e^{-y} dx \, dy = c \int_{-y}^{y} \left( y^2 - x^2 \right) e^{-y} \, dx \, dy = c \int_{-y}^{y} \left( y^2 - \frac{x^2}{3} \right) e^{-y} \, dy$$

$$= c \int_{-y}^{y} e^{-y} \left( y^2 - \frac{y^3}{3} \right) dy$$

$$= \frac{4}{3} c \int_{0}^{\infty} y^3 e^{-y} \, dy \quad \text{let } u = y^3, \, dv = e^{-y} \, dy$$

$$= \frac{4}{3} \left( \frac{y^3}{e^y} \right)_{-y}^{y} + 3 \int_{0}^{\infty} y^2 e^{-y} \, dy$$

$$= \frac{4}{3} c \left( \frac{y^3}{e^y} \right)_{0}^{\infty} + 3 \int_{0}^{\infty} y^2 e^{-y} \, dy$$

$$= \frac{4}{3} \left( \frac{y^3}{e^y} \right)_{0}^{\infty} + 3 \left( \frac{y^2}{e^y} \right)_{0}^{\infty} + 6 \left( -y e^{-y} \right)_{0}^{\infty}$$

$$= \frac{4}{3} \left( \frac{0}{e^0} - 3 \frac{1}{e^0} - 6 \frac{1}{e^0} - 6 \frac{1}{e^0} \right)$$

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

Thus, the solution is:
\[ c = \frac{1}{8} \]

(b) \[ f_X(x) = \int f(x, y) dy = \int_0^\infty c(y^2 - x^2)e^{-y} dy = \frac{1}{8} \int_0^\infty (y^2 - x^2)e^{-y} dy \]

\[ = \left| \text{let } u = y^2 - x^2, dv = e^{-y} dy \right| = \frac{1}{8} \left( -(y^2 - x^2)e^{-y} \right)_0^\infty + 2 \int_0^\infty ye^{-y} dy \]

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

\[ = \frac{1}{8} \left( -(y^2 - x^2)e^{-y} - 2ye^{-y} - 2e^{-y} \right)_0^\infty = \frac{1}{8} (0 + 0 + (0 + x^2) \cdot 1 + 0 + 2 \cdot 1) = \frac{x^2 + 2}{8} - y \leq x \leq y \]

\[ f_Y(y) = \int_x f(x, y) dx = \int_y c(y^2 - x^2)e^{-y} dx = \frac{1}{8} e^{-y} \int_y^{\infty} (y^2 - x^2) dx \]

\[ = \frac{1}{8} e^{-y} \cdot \left( y^2 x - \frac{x^3}{3} \right)_y^{\infty} = \frac{1}{8} e^{-y} \left( y^2 \cdot y - \frac{y^3}{3} - y^2 \cdot (-y) + \frac{(-y)^3}{3} \right) \]

\[ = \frac{1}{8} e^{-y} \cdot \frac{4}{3} y^3 = \frac{y^3 e^{-y}}{6}, y \geq 0 \]

(c) \[ E(X) = \int_x x f_X(x) dx = \int_y x \cdot \frac{y^2 + 2}{8} dx = \frac{1}{8} \int_y^{\infty} (x^3 + 2x) dx = \frac{1}{8} \left( \frac{x^4}{4} + x^2 \right)_y^{\infty} \]

\[ = \frac{1}{8} \left( \frac{y^4}{4} + y^2 - \frac{(-y)^4}{4} - (-y)^2 \right) = 0 \]
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: $c \leq a$ 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 \( (c-b) \):

\[
S_2 = \frac{c^2 - (c-b)^2}{2} = \frac{2bc - b^2}{2}
\]

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:

\[
S_3 = \frac{c^2 - (c-a)^2}{2} = \frac{2ac - a^2}{2}
\]

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 \( (c-a) \):

\[
S_4 = \frac{c^2 - (c-b)^2 - (c-a)^2}{2}
\]

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:
Case 6: \( c > a + b \)

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

\[
S_6 = ab
\]

Summarize the areas found to build one function:

\[
S = \begin{cases}
\frac{c^2}{2}, & \text{if } c \leq a \text{ and } c \leq b \\
\frac{c^2 - (c - b)^2}{2}, & \text{if } b < c \leq a \\
\frac{c^2 - (c - a)^2}{2}, & \text{if } a < c \leq b \\
\frac{c^2 - (c - b)^2 - (c - a)^2}{2}, & \text{if } a < b < c \leq a + b \text{ or } b < a < c \leq a + b \\
ab, & \text{if } c > a + b
\end{cases}
\]

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

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

Area of the rectangle:

\[
S_{\text{full}} = ab
\]

Use the formulas for areas to find the probability for each case:
\[ P(aX + bY \leq c) = \begin{cases} 
\frac{c^2}{2ab}, & \text{if } c \leq \min(a, b) \\
\frac{c^2 - (c - \min(a, b))^2}{2ab}, & \text{if } \min(a, b) < c \leq \max(a, b) \\
\frac{c^2 - (c - b)^2 - (c - a)^2}{2ab}, & \text{if } \max(a, b) < c \leq a + b \\
1, & \text{if } c > a + b
\end{cases} \]

**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

\[ P(X \leq s, Y \leq t) = \int_{x \leq s, y \leq t} f_{XY}(x, y) \, dx \, dy \]

Now, consider the random variable \( X + Y \). Consider the following probability.

\[ P(X + Y \leq a) = \int_{s \leq a} P(X \leq s, Y \leq a - s) \, ds = \int_{s \leq a} \int_{x \leq s, y \leq a - s} f_{XY}(x, y) \, dx \, dy \, ds \]

The function \( f_{XY}(x, y) \) is continuous in \( \mathbb{R}^2 \). 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

\[ P_j = P(X \leq s, Y \leq t, X + Y \leq a) = \int_{x \leq s, y \leq t, x + y \leq a} f_{XY,X+Y}(x, y, x + y) \, dx \, dy \]

When looking at the formula above we can understand that the conditions \( x \leq s, y \leq t, x + y \leq a \) 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 \( P_j \) will exist and will be non-zero in general case. But just above the line \( s + t = a + \epsilon \) we are sure to get \( P_j = 0 \), because if \( s + t > a \) the events \( x \leq s, y \leq t, x + y \leq a \) will never occur simultaneously.

As we can see, \( P_j \) 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.