MATH 241 PSET 9

November 21, 2024

1.

a) We first check that $f_{X,Y}(x,y) \geq 0$. For $0 < x < 1$ and $0 < y < 1$, $f_{X,Y}(x,y) = x + y \geq 0$. Outside this region, $f_{X,Y}(x,y) = 0$. Thus, this condition is met.
The second condition to check that $f_{X,Y}$ is a valid joint PDF is $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$. We verify this below:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = \int_0^1\int_0^1 (x+y)dxdy = \int_0^1 (\frac{x^2}{2} + xy)\Big|_0^1 dy$$
$$= \int_0^1 (\frac{1}{2} + y)dy = \frac{y}{2} + \frac{y^2}{2}\Big|_0^1 = \frac{1}{2} + \frac{1^2}{2} = 1$$

b) Because we cannot factorize joint PDF $f_{X,Y}(x,y) = x+y$ into $g(x)h(y)$, where g and h are non-negative functions, we can conclude that X and Y are not independent.

c) We compute the marginal PDFs of $f_X(x)$ and $f_Y(y)$ below:

$$f_X(x) = \int_{-\infty}^{\infty} f_{x,y}(x,y)dy = \int_0^1 (x+y)dy = xy + \frac{y^2}{2}\Big|_0^1 = x + \frac{1}{2}$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{x,y}(x,y)dx = \int_0^1 (x+y)dx = \frac{x^2}{2} + xy\Big|_0^1 = y + \frac{1}{2}$$

d) The conditional PDF of Y given $X = x$ is given by: $f_{Y|X}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{x+y}{x+0.5} = \frac{2x+2y}{2x+1}$ for $0 < y < 1$.

2.

a) For $f_{X,Y}(x,y)$ to be a valid joint PDF, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy$ must equal 1. We solve this equation for c:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$$
$$\int_0^1\int_0^y (cxy)dxdy = \int_0^1 cy\frac{x^2}{2}\Big|_0^y dy = \int_0^1 c\frac{y^3}{2}dy = 1$$
$$c\frac{y^4}{8}\Big|_0^1 = 1$$
$$\frac{c}{8} = 1$$
$$c = 8$$

Thus, $c = 8$ for $f_{X,Y}$ to be a valid joint PDF.

b) Intuitively speaking, two random variables would be independent if knowing the value of one of the random variables does not give any information about the other random variables. This is not the case for the random variables X and Y: the given joint PDF $f_{X,Y}(x,y)$ is defined as non-zero across the region $0 < x < y < 1$. This means that knowing the given value of Y gives us information (i.e. an upper bound) on the value of X and thus we can see that random variables X and Y are not independent.[1]

c) We compute the marginal PDFs $f_X(x)$ and $f_Y(y)$ of X and Y for $0 < x < 1$ and $0 < y < 1$, respectively, below.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_x^1 (cxy)dy = cx\frac{y^2}{2}\Big|_x^1 = \frac{cx}{2} - \frac{cx^3}{2} = 4(x - x^3)$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_0^y (cxy)dx = cy\frac{x^2}{2}\Big|_0^y = \frac{cy^3}{2} = 4y^3$$

d) The conditional PDF of Y given $X = x$ is given by $f_{Y|X}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{8xy}{4(x-x^3)} = \frac{2y}{1-x^2}$.

3.

a) We compute the covariance of distributions $X + Y$ and $X - Y$ below. Note that $\mathbb{E}[X] = \mathbb{E}[Y] = 0.5$.

$$Cov(X+Y, X-Y) = \mathbb{E}[(X+Y)(X-Y)] - \mathbb{E}[X+Y]\mathbb{E}[X-Y]$$
$$= \mathbb{E}[X^2 - YX + YX - Y^2] - (\mathbb{E}[X] + \mathbb{E}[Y])(\mathbb{E}[X] - \mathbb{E}[Y]) =$$
$$\mathbb{E}[X^2 - Y^2] - (0.5 + 0.5)(0.5 - 0.5) = \mathbb{E}[X^2] - \mathbb{E}[Y^2]$$

Because r.v.s X and Y are identical and independent, $\mathbb{E}[X^2] = \mathbb{E}[Y^2]$ and so $Cov(X+Y, X-Y) = 0$.

[1] In more rigorous terms, this is because the support of the given joint distribution ($0 < x < y < 1$) is not defined over a rectangular region.

b) We demonstrate that $X-Y$ and $X+Y$ are not independent through a proof by contradiction. Let us define some tiny positive constant $\epsilon > 0$. If $X-Y$ and $X+Y$ were independent, then $P(X-Y > 1-\epsilon | X+Y > 2-\epsilon) = P(X-Y > 1-\epsilon)$. We know that $P(X-Y > 1-\epsilon) > 0$ as it is possible for $X-Y$ to be nearly equal to one. However, $P(X-Y > 1-\epsilon | X+Y > 2-\epsilon) = 0$ as if we know that $X+Y$ is approximately equal to two, both X and Y are equal to near one and so their difference cannot be greater than one. Thus, we have shown the following below:

$$P(X-Y > 1-\epsilon | X+Y > 2-\epsilon) = 0 \neq P(X-Y > 1-\epsilon)$$

and so we can conclude that $X-Y$ and $X+Y$ are not independent.

4.

a) Let us define constants $a, b, c \in \mathbb{R}$. A linear combination of $(X, Y, X+Y)$ can be given as $aX + bY + c(X+Y) = (a+c)X + (b+c)Y$, which is a summation of two independent normal distributions (i.e. $(a+c)X$ and $(b+c)Y$) and is thus a normal distribution. Thus, because all linear combinations of $(X, Y, X+Y)$ form a Normal distribution, $(X, Y, X+Y)$ is Multivariate Normal.

b) For $(X, Y, SX+SY)$ to be Multivariate Normal, all linear combinations of the set $\{X, Y, SX+SY\}$ must be normal distribution. Because with probability 50% r.v. $S = -1$, a possible linear combination of this set is $X + Y - X - Y = 0$ which does not follow a normal distribution, not all linear combinations of $\{X, Y, SX+SY\}$ follow a normal distribution. Thus, $(X, Y, SX+SY)$ is not Multivariate Normal.

c) To see if (SX, SY) is Multivariate Normal, we must test to see if all possible linear combinations of $\{SX, SY\}$ follow a normal distribution. All linear combinations of this set can be given as $aSX + bSY$, where $a, b \in \mathbb{R}$. In either case of whether $S = 1$ or $S = -1$, all linear combinations of $\{SX, SY\}$ can be given by the set $\{aX + bY : a, b \in \mathbb{R}\}$. Because each linear combination from this set is a sum of independent normal distributions, each linear combination from this set follows a normal distribution $\Rightarrow$ all linear combinations of (SX, SY) follow a normal distribution $\Rightarrow$ (SX, SY) is Multivariate Normal.

5. We first solve for c s.t. $Corr(Y - cX, X) = 0$.

$$\begin{aligned}
Corr(Y-cX, X) = \frac{Cov(Y-cX, X)}{\sqrt{Var(Y-cX)Var(X)}} &= 0 \\
Cov(Y-cX, X) &= 0 \\
Cov(Y, X) - cCov(X, X) &= 0 \\
Cov(X, Y) &= cVar(X) \\
Corr(X, Y)\sigma_1\sigma_2 &= c\sigma_1^2 \\
c &= \frac{\rho\sigma_2}{\sigma_1}
\end{aligned}$$

Because (X, Y) is Bivariate Normal, the distribution $(Y - cX, X)$ is also Bivariate Normal. This is because any linear combination of $\{Y - cX, X\}$ can be written as a linear combination of $\{X, Y\}$ (which we know follows a normal distribution as (X, Y) is Bivariate Normal), and thus any linear combination of $\{Y - cX, X\}$ follows a normal distribution $\Rightarrow (Y - cX, X)$ is Bivariate Normal. Because $(Y - cX, X)$ is Bivariate Normal and $Corr(Y - cX, X) = 0$, we can conclude that $Y - cX$ and X are independent.

6. Anish Lakkapragada. I worked independently.