STATISTICS: Problem Set 2-Solutions

September 28, 2007

1. Let X_1, X_2, X_3 and X_4 be four mutually stochastically independent random variables, each with pdf $f_{X_i}(x_i) = 3(1-x_i)^2$, $0 < x_i < 1$. If Y is the minimum of these four variables, find the cdf and the pdf of Y.
(*Hint:* $F_Y(y) = P(Y \leq y) = 1 - P(Y > y)$ *and* $P(Y > y) = P(X_i > y,\ i = 1,2,3,4)$).

Thanks to the hint, we can write:

$$F_Y(y) = P(Y \leq y) = 1 - P(Y > y) = 1 - P(X_i > y,\ i = 1,2,3,4)$$

Now, we now from the text that the variables are independent between each other. Therefore, the probability that all of them are greater than y is equal to the product of the probabilities that each of them is greater than y. Thus, we can write:

$$F_Y(y) = P(Y \leq y) = 1 - P(X_1 > y)...P(X_4 > y)$$

Furthermore, we know that they are identically distributed, such that:

$$F_Y(y) = P(Y \leq y) = 1 - P(X_i > y)^4 = 1 - (1 - F_{X_i}(y))^4$$

Moreover, $F_{X_i}(y) = \int_0^y 3(1-u)^2 du = 1 - (1-y)^3$
Plugging the latter equation in the expression of $F_Y(y)$, we obtain:

$$F_Y(y) = 1 - ((1-y)^3)^4 = 1 - (1-y)^{12}$$

To get the pdf, we have to differentiate the cdf which we just obtained:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = 12(1-y)^{11} \mathbb{1}_{0<y<1}$$

2. Let X_1 and X_2 be jointly exponential with density function given by:

$$p_{X_1,X_2}(x_1,x_2) = \lambda^2 e^{-\lambda(x_1+x_2)},\ x_1 > 0,\ x_2 > 0.$$

(a) Find the marginal distribution of X_1
(b) Find the conditional distribution of $Y = X_1 + X_2$ given X_1
(c) Find $E(X_1)$, $Var(X_1)$ and $Var(Y|X_1)$

(a) We have:

$$f_{X_1}(x_1) := \int_0^\infty f_{X_1,X^2}(x_1,x_2)dx_2 = \int_0^\infty \lambda^2 e^{-\lambda(x_1+x_2)}dx_2 = \lambda e^{-\lambda x_1}\int_0^\infty \lambda e^{-\lambda x_2}dx_2$$

Calculate the obtained integral:

$$\int_0^\infty \lambda e^{-\lambda x_2}dx_2 = \int_0^\infty e^{-\lambda x_2}d\lambda x_2 = \int_0^\infty e^{-u}du = -e^{-u}|_0^\infty = 1$$

Therefore, we obtain:

$$f_{X_1}(x_1) = \lambda e^{-\lambda x_1} * 1 = \lambda e^{-\lambda x_1}$$

X_1 is distributed according to the exponential distribution, with the parameter λ. By analogy, we would find $f_{X_2}(x_2) = \lambda e^{-\lambda x_2}$. Note that

$$p_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$$

Hence, X_1 and X_2 are independent random variables.
(b) Let $Y = X_1 + X_2$. We know $f_{X_1,X_2}(x_1,x_2)$, but we need to determine $f_{Y|X_1}(y,x_1)$.

$$f_{Y|X_1}(y,x_1) = \frac{f_{Y,X_1}(y,x_1)}{f_{X_1}(x_1)},$$

so first, define the joint density of Y and X_1.
Let $Y_1 = X_1 + X_2$ and $Y_2 = X_1$. If we reverse the functions, we get:
$X_1 = Y_2$ and $X_2 = Y_1 - Y_2$. Let us define J as the determinant of the first-order partial derivatives matrix, we have:

$$J = \begin{vmatrix} \frac{\partial X_1}{\partial Y_1} & \frac{\partial X_1}{\partial Y_2} \\ \frac{\partial X_2}{\partial Y_1} & \frac{\partial X_2}{\partial Y_2} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Now, taking the absolute value and plugging it in the transformation formula, we obtain the joint density function of Y_1 and $X_1 = Y_2$:

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(y_2, y_1 - y_2) = \lambda^2 e^{-\lambda(y_2+y_1-y_2)} = \lambda^2 e^{-\lambda y_1}$$

Now we can apply the formula of conditional probability:

$$f_{Y_1|Y_2}(y_1, y_2) = \frac{f_{Y_1,Y_2}(y_1,x_2)}{f_{Y_2}(y_2)} = \frac{\lambda^2 e^{-\lambda y_1}}{\lambda e^{-\lambda y_2}} = \lambda e^{-\lambda y_1 - y_2}$$

Recalling that $Y = Y_1$ and $Y_2 = X_1$, we get:

$$f_{Y|X_1}(y, x_1) = \lambda e^{-\lambda(y-x_1)}$$

Intuition: Knowing one part of the sum, the rest $(Y|X_1)$ is distributed like the second part, i.e. the distribution of $Y|X_1$ is the one of X_2 shifted by x_1.
(c) Since X_1 is exponentially distributed, we know that $E(X_1) = \frac{1}{\lambda}$ and $V(X_1) = \frac{1}{\lambda^2}$
Moreover, in the calculation of the conditional variance, we can treat x_1 as a constant such that $V(X_1 + X_2|X_1 = x_1) = V(X_2|X_1 = x_1)$.
Finally, the independence between X_1 and X_2 allows us to write:

$$V(X_2|X_1 = x_1) = V(X_2) = \frac{1}{\lambda^2}$$

So, we obtain: $V(Y|X_1 = x_1) = \frac{1}{\lambda^2}$

3. Consider the joint density function of the random variables X and Y given by:

$$f_{X,Y}(x, y) = 8xy,\ 0 < x < y < 1;\ 0 \text{ elsewhere.}$$

Show that:
(a) The marginal density $f_X(x) = 4x(1 - x^2)$, $0 < x < 1$
(b) The marginal density $f_Y(y) = 4y^3$, $0 < y < 1$
(c) Find the conditional densities $f_{X|Y}(x, y)$ and $f_{Y|X}(y, x)$, paying attention to the interval over which the functions are defined.
(d) Show that $E(Y|X = x) = \frac{2}{3}\left(\frac{1-x^3}{1-x^2}\right)$
(e) Show that $E(X|Y = y) = \frac{2}{3}y$

(a)

$$f_X(x) := \int_x^1 8xydy = 8x(\frac{y^2}{2}|_x^1) = 8x(\frac{1-x^2}{2}) = 4x(1 - x^2), \qquad 0 < x < 1$$

(b)

$$f_Y(y) := \int_0^y 8xydx = 8y(\frac{x^2}{2}|_0^y) = 8y(\frac{y^2}{2}) = 4y^3, \qquad 0 < y < 1$$

(c)

$$f_{Y|X}(y,x) := \frac{f_{Y,X}(y,x)}{f_X(x)} = \frac{8xy}{4x(1-x^2)} = \frac{2y}{1-x^2}\mathbb{I}_{x<y<1}$$
$$f_{X|Y}(x,y) := \frac{f_{Y,X}(y,x)}{f_Y(y)} = \frac{8xy}{4y^3} = \frac{2x}{y^2}\mathbb{I}_{0<x<y}$$

(d)

$$E(Y|X=x) := \int_x^1 y f_{Y|X}(y,x)dy = \int_x^1 \frac{2y^2}{1-x^2}dy = \frac{2}{1-x^2}\int_x^1 y^2 dy = \frac{2}{1-x^2}(\frac{y^3}{3}|_x^1) = \frac{2(1-x^3)}{3(1-x^2)}$$

(e)

$$E(X|Y=y) := \int_0^y x f_{X|Y}(y,x)dx = \int_0^y \frac{2x^2}{y^2}dx = \frac{2}{y^2}\int_0^y x^2 dx = \frac{2}{y^2}(\frac{x^3}{3}|_0^y) = \frac{2y^3}{3(y^2)} = \frac{2}{3}y.$$

4. Using the result according to which if two random variables X and Y are independent, then the mgf of $Z = X + Y$ is the product of the mgf of both X and Y, show that the sum of two poisson is poisson distributed.

Recalling Problem set 1 (Q. 7), the moment generating function for a random variable which is Poisson distributed with parameter λ is equal to $M(t) = e^{\lambda(e^t-1)}$ Let us define Z as a random variable equal to the sum of 2 Poisson distributed random variables X_1 and X_2. We know that $M_Z(t) = M_{X_1}(t) * M_{X_2}(t)$. Thus, we have:

$$M_Z(t) = e^{\lambda_1(e^t-1)} * e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)} \tag{1}$$

We see that this is the moment generating function for the Poisson distribution with parameter $(\lambda_1 + \lambda_2)$. Therefore, Z is Poisson distributed with parameter $(\lambda_1 + \lambda_2)$.
To conclude, if X and Y are two **independent** random variables distributed according to the Poisson distribution with parameters λ_X and λ_Y, respectively, then $Z = X + Y$ is distributed according to the Poisson distribution with parameter $\lambda_Z = \lambda_X + \lambda_Y$.

5. Let X_1, X_2,..., X_T be iid with normal density $N(\mu, \sigma^2)$. Find the distribution of:

$$Y_n = \frac{\sum_{k=1}^n kX_k - \mu\sum_{k=1}^n k}{\sum_{k=1}^n k^2}$$

We have:

$$Y_n = \frac{\sum_{k=1}^n kX_k - \sum_{k=1}^n k\mu}{\sum_{k=1}^n k^2} = \frac{\sum_{k=1}^n k(X_k-\mu)}{\sum_{k=1}^n k^2} \tag{2}$$

$$\begin{gathered} X_k \sim N(\mu,\sigma^2) \\ \implies (X_k-\mu) \sim N(0,\sigma^2) \\ \implies k(X_k-\mu) \sim N(0,k^2\sigma^2) \end{gathered} \tag{3}$$

By independence between X_k, we can write:

$$\begin{array}{c} \sum_{k=1}^{n} k(X_k-\mu) \sim N(0, \sum_{k=1}^{n} k^2\sigma^2) \\ \Longrightarrow \frac{\sum_{k=1}^{n} k(X_k-\mu)}{\sum_{k=1}^{n} k^2} = Y_n \sim N(0, \frac{\sigma^2}{\sum_{k=1}^{n} k^2}) \end{array} \tag{4}$$

6. Let $Y_1 = \frac{1}{2}(X_1 - X_2)$, where X_1 and X_2 have the following joint pdf:

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{4}e^{-\frac{x_1+x_2}{2}} \mathbb{1}_{x_1>0}\mathbb{1}_{x_2>0}$$

Find the pdf of Y_1.

Let $Y_2 = X_2$. Since

$$f_{Y_1}(y_1) = \int_0^\infty f_{Y_1,Y_2}(y_1,y_2)dy_2,$$

let us first find the joint density function of Y_1 and Y_2. Taking the reverse functions, we have $x_1 = 2y_1 + y_2$ and $x_2 = y_2$

$$J = \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = 2$$

Using the transformation formula, we obtain:

$$f_{Y_1,Y_2}(y_1,y_2) = 2f_{X_1,X_2}(2y_1+y_2,y_2) = \frac{1}{2}e^{-y_1-y_2}\mathbb{1}_{-\infty<y_1<\infty}\mathbb{1}_{y_2>\max\{-2y_1,0\}}$$

To get the marginal density of Y_1, we just need to integrate over the support of Y_2. What we have to be careful about is the domain of the function f_{Y_1,Y_2}. Indeed, if y_1 is positive, y_2 takes its values from 0 to ∞:
$f_{Y_1}(y_1) = \int_0^\infty \frac{1}{2}e^{-y_1-y_2}dy_2 = \frac{1}{2}e^{-y_1}$. On the contrary, when y_1 is negative, y_2 takes its values from $-2y_1$ to ∞:
$f_{Y_1}(y_1) = \int_{-2y_1}^\infty \frac{1}{2}e^{-y_1-y_2}dy_2 = \frac{1}{2}e^{y_1}$.
In the end, we have:

$$f_{Y_1}(y_1) = \frac{1}{2}e^{-|y_1|}\mathbb{1}_{-\infty<y_1<\infty}$$

7. Suppose that the joint probability density function of the bivariate random variable (X,Y) is given by:

$$f_{X,Y}(x,y) = [1-\alpha(1-2x)(1-2y)]\mathbb{1}_{0\le x\le 1}\mathbb{1}_{0\le y\le 1}$$

(a) Work out $E(XY)$
(b) Work out the marginal density functions $f_X(x)$ and $f_Y(y)$ and hence $E(X)$ and $E(Y)$.
(c) Attempt to prove or disapprove the following statement:

"In this example, the variables X and Y are independent if and only if they are uncorrelated."

(a)

$$E(XY) := \int_0^1 \int_0^1 xy f_{X,Y}(x,y)dxdy = \int_0^1 y \int_0^1 x f_{X,Y}(x,y)dxdy$$
$$= \int_0^1 y \int_0^1 x[1-\alpha(1-2x)(1-2y)]dxdy = \int_0^1 y \int_0^1 x[1-\alpha+2\alpha y] + 2x^2\alpha(1-2y)dxdy$$
$$= \int_0^1 y[\tfrac{x^2(1-\alpha+2\alpha y)}{2} + \tfrac{2}{3}(\alpha-2\alpha y)x^3]|_0^1 dy = \int_0^1 y \int_0^1 \tfrac{3-3\alpha+6\alpha y+4\alpha-8\alpha y}{6} dy$$
$$= \int_0^1 y(\tfrac{1}{2}+\tfrac{1}{6}\alpha-\tfrac{1}{3}\alpha y)dy = \left[\tfrac{1}{4}+\tfrac{1}{12}\alpha)y^2 - \tfrac{1}{9}\alpha y^3\right]|_0^1$$
$$= \tfrac{1}{4}(1-\tfrac{\alpha}{9})$$

(b)

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)dy = \int_0^1 [1-\alpha(1-2x)+2\alpha(1-2x)y]dy$$
$$= 1-\alpha(1-2x)+\alpha(1-2x)y^2|_0^1 = 1$$
$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)dx = 1$$
$$E(X) := \int_0^1 x f_X(x)dx = \int_0^1 xdx = \tfrac{x^2}{2}|_0^1 = \tfrac{1}{2}$$
$$E(Y) = \tfrac{1}{2}$$

(c)

First let us compute the covariance between X and Y:

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{4}(1-\frac{\alpha}{9}) - \frac{1}{2} * \frac{1}{2} = -\frac{1}{36}\alpha$$

To be proved: Independence implies null covariance If both random variables are independent, this means that:

$$f_{X,Y}(x,y) = f_X(x) * f_Y(y) = 1$$

The only way for the joint density to be equal to 1 is to have $\alpha = 0$. But if $\alpha = 0$, then $Cov(X,Y) = 0$. Therefore, the independence implies that the covariance is equal to 0.

To be proved: Null covariance implies independence If the covariance is equal to 0, we must have $\alpha = 0$. But then if $\alpha = 0$, the joint density is equal to 1 , which means it is equal to the product of marginal densities. Thus, the two random variables are independent. Therefore, the statement is true.

8. Let X and Y be random variables with $E(X) = E(Y) = 0$. Assuming that $E(XY)$ exists and that $E(X|Y) = 0$, show that X and Y are uncorrelated.

We know that by definition $Cov(X,Y) = E(XY) - E(X)E(Y)$. Moreover, we know that $E(X) = 0$. Thus, in order to prove that X and Y are uncorrelated ($i.e$ $Cov(X,Y) = 0$), it is sufficient to show that $E(XY)$ is equal to 0.

$$E(XY) = \int\int xyf_{X,Y}(x,y)dxdy = \int y \int xf_{X|Y}(x,y)f_Y(y)dxdy =$$
$$\int y \int xf_{X|Y}(x,y)dxf_Y(y)dy = \int yE(X|Y=y)f_Y(y)dy = 0$$

We know that $E(X|Y=y) = 0$, so $E(XY) = 0$ as well. As a consequence, the covariance is also equal to 0. Therefore, X and Y are uncorrelated.

9. The joint density of 2 random variables X and Y is given by:

$$f_{X,Y}(x,y) = \frac{1}{y}e^{-\frac{x}{y}}e^{-y}\mathbb{1}_{x>0}\mathbb{1}_{y>0}$$

Compute E(X|Y=y).

Let us first compute $f_{X|Y}$. By definition, we have:

$$f_Y(y) := \int_0^\infty \frac{1}{y}e^{-\frac{x}{y}}e^{-y}\mathbb{1}_{y>0}dx = \frac{e^{-y}}{y}\mathbb{1}_{y>0}\int_0^\infty e^{-\frac{x}{y}}dx = e^{-y}\mathbb{1}_{y>0}$$

So,

$$f_{X|Y}(x,y) := \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{e^{-y}e^{-\frac{x}{y}}}{y}\mathbb{1}_{x>0}}{e^{-y}} = \frac{1}{y}e^{-\frac{x}{y}}\mathbb{1}_{x>0}$$

Now, we can write the expression for $E(X|Y=y)$:

$$E(X|Y=y) := \int_0^\infty \frac{x}{y}e^{-\frac{x}{y}}dx = -\int_0^\infty xde^{-\frac{x}{y}} = -xe^{-\frac{x}{y}}|_0^\infty + \int_0^\infty e^{-\frac{x}{y}}dx$$
$$= 0 - ye^{-\frac{x}{y}}|_0^\infty = y$$