Section 6.1 Joint Distribution Functions

We often care about more than one random variable at a time.

DEFINITION: For any two random variables X and Y the *joint cumulative probability distribution function* of X and Y is

$$\boxed{F(a,b) = P\{X \le a, Y \le b\}, \quad -\infty < a, b < \infty}$$

One can get the distribution of X alone from the joint distribution function

$$F_X(a) = P\{X \le a\} = P\{X \le a, Y < \infty\} = P\left\{\lim_{t\to\infty}\{X \le a, Y \le t\}\right\}$$

$$= \lim_{t\to\infty} F(a,t) \equiv F(a,\infty)$$

Similarly,

$$F_Y(b) = \lim_{t\to\infty} F(t,b) \equiv F(\infty,b)$$

We can also show that for $a_1 < a_2$ and $b_1 < b_2$ we have

$$\boxed{P\{a_1 < X \le a_2, b_1 < Y \le b_2\} = F(a_2,b_2) + F(a_1,b_1) - F(a_1,b_2) - F(a_2,b_1)}$$

For discrete random variables we define the joint probability mass function for X and Y by

$$p(x,y) = P\{X = x,\ Y = y\}$$

Note that

$$p_X(x) = P\{X = x\} = P\{X = x,\ Y \in \mathbb{R}\} = \sum_{y:p(x,y)>0} p(x,y)$$

and

$$p_Y(y) = \sum_{x:p(x,y)>0} p(x,\ y)$$

EXAMPLE: We flip a fair coin twice. Let X be 1 if head on first flip, 0 if tail on first. Let Y be number of heads. Find $p(x,y)$ and $p_X,\ p_Y$.

EXAMPLE: We flip a fair coin twice. Let X be 1 if head on first flip, 0 if tail on first. Let Y be number of heads. Find $p(x,y)$ and $p_X,\ p_Y$.

Solution: The ranges for X and Y are $\{0,1\},\ \{0,1,2\}$, respectively. We have

$$p(0,0) = P(X=0,\ Y=0) = P(X=0)P(Y=0 \mid X=0) = (1/2)(1/2) = 1/4$$
$$p(0,1) = P(X=0,\ Y=1) = P(X=0)P(Y=1 \mid X=0) = (1/2)(1/2) = 1/4$$
$$p(0,2) = P(X=0,\ Y=2) = 0$$
$$p(1,0) = P(X=1,\ Y=0) = 0$$
$$p(1,1) = P(X=1,\ Y=1) = P(X=1)P(Y=1 \mid X=1) = (1/2)(1/2) = 1/4$$
$$p(1,2) = P(X=1,\ Y=2) = P(X=1)P(Y=2 \mid X=1) = (1/2)(1/2) = 1/4$$

$x \ddots y$	0	1	2	Row Sum $= P\{X=x\}$
0	1/4	1/4	0	1/2
1	0	1/4	1/4	1/2
Column Sum $= P\{Y=y\}$	1/4	1/2	1/4	

Further,

$$p_X(0) = \sum_{y=0}^{2} p(0,y) = p(0,0) + p(0,1) + p(0,2) = 1/2$$
$$p_X(1) = \sum_{y=0}^{2} p(1,y) = p(1,0) + p(1,1) + p(1,2) = 1/2$$
$$p_Y(0) = \sum_{x=0}^{2} p(x,0) = p(0,0) + p(1,0) = 1/4$$
$$p_Y(1) = \sum_{x=0}^{2} p(x,1) = p(0,1) + p(1,1) = 1/2$$
$$p_Y(2) = \sum_{x=0}^{2} p(x,2) = p(0,2) + p(1,2) = 1/4$$

DEFINITION: We say that X and Y are *jointly continuous* if there exists a function $f(x,y)$ defined for all x,y such that for any $C \subset \mathbb{R}^2$ we have

$$P\{(X,\ Y) \in C\} = \iint_{(x,y)\in C} f(x,y)dxdy$$

where $f(x,\ y)$ is called the *joint probability density function* of X and Y.

Thus,

$$P\{X \in A,\ Y \in B\} = \int_B \int_A f(x,y)dxdy$$

Also,

$$\frac{\partial^2}{\partial a \partial b} F(a,b) = \frac{\partial^2}{\partial a \partial b} \int_{-\infty}^{b} \int_{-\infty}^{a} f(x,y) dx dy = f(a,b)$$

Similar to before, we view the joint density as a measure of the likelihood that the random vector (X, Y) will be in the vicinity of (a,b). As before,

$$P\{a < X < a + da,\ b < Y < b + db\} = \int_{b}^{b+db} \int_{a}^{a+da} f(x,y) dx dy \approx f(a,b) da db$$

where da and db should be thought of as very small.

We can find the marginal probability density functions as follows

$$P\{X \in A\} = P\{X \in A,\ Y \in (-\infty,\ \infty)\} = \int_{A} \int_{-\infty}^{\infty} f(x,y) dy dx = \int_{A} f_X(x) dx$$

where

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy$$

is therefore the density function for X. Similarly,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y) dx$$

EXAMPLE: Suppose the joint density is

$$f(x,y) = 2e^{-x}e^{-2y}$$

if $x, y > 0$ and zero otherwise. Find the marginal of X, $f_X(x)$. Compute $P\{X > 1, Y < 1\}$ and $P\{X < Y\}$.

EXAMPLE: Suppose the joint density is

$$f(x,y) = 2e^{-x}e^{-2y}$$

if $x, y > 0$ and zero otherwise. Find the marginal of X, $f_X(x)$. Compute $P\{X > 1, Y < 1\}$ and $P\{X < Y\}$.

Solution: For $x > 0$ we have

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{0}^{\infty} 2e^{-x}e^{-2y}dy = e^{-x}$$

Next,

$$P\{X > 1,\ Y < 1\} = \int_{0}^{1}\int_{1}^{\infty} 2e^{-x}e^{-2y}dxdy = \int_{0}^{1} 2e^{-2y}e^{-1}dy = e^{-1}(1 - e^{-2})$$

Finally,

$$P\{X < Y\} = \iint_{(x,y):x<y} f(x,y)dxdy = \int_{0}^{\infty}\int_{0}^{y} 2e^{-x}e^{-2y}dxdy = \int_{0}^{\infty} 2e^{-2y}(1 - e^{-y})dy$$

$$= \int_{0}^{\infty} 2e^{-2y}dy - \int_{0}^{\infty} 2e^{-3y}dy = 1 - \frac{2}{3} = \frac{1}{3}$$

Everything extends to more than two random variables in the obvious way. For example, for n random variables we have

$$F(a_1, a_2, \ldots, a_n) = P\{X_1 \le a_1, X_2 \le a_2, \ldots, X_n \le a_n\}$$

Also, joint probability mass functions and joint densities are defined in the obvious manner.

Section 6.2 Independent Random Variables

DEFINITION: Two random variables are said to be *independent* if for any two sets of real numbers A and B we have

$$\boxed{P\{X \in A,\ X \in B\} = P\{X \in A\}P\{X \in B\}} \qquad (1)$$

That is, the events $E_A = \{X \in A\}$ and $F_B = \{Y \in B\}$ are independent.

One can show that the above equation holds if and only if for all a and b we have

$$P\{X \le a,\ Y \le b\} = P\{X \le a\}P\{Y \le b\}$$

Therefore, two random variables X and Y are independent if and only if the joint distribution function is the product of the marginal distribution functions:

$$\boxed{F(a,b) = F_X(a)F_Y(b)}$$

It's easy to see that for *discrete* random variables the condition (1) is satisfied if and only if

$$\boxed{p(x,y) = p_X(x)p_Y(y)} \tag{2}$$

for all x, y where $p(x,y) = P\{X = x, Y = y\}$ is the joint probability density function. This is actually easy to show.

Proof: Note that (2) follows immediately from condition (1). However, for the other direction, for any A and B we have

$$P\{X \in A, Y \in B\} = \sum_{y \in B} \sum_{x \in A} p(x,y) = \sum_{y \in B} \sum_{x \in A} p_X(x)p_Y(y)$$

$$= \sum_{y \in B} p_Y(y) \sum_{x \in A} p_X(x) = P\{Y \in B\}P\{X \in A\} \blacksquare$$

Similarly, in the jointly continuous case independence is equivalent to

$$\boxed{f(x,y) = f_X(x)f_Y(y)}$$

It should be clear that independence can be defined in the obvious manner for more than two random variables. That is, $X_1, X_2, \ldots, X_n$ are *independent* if for all sets of real numbers $A_1, A_2, \ldots, A_n$ we have

$$\boxed{P\{X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n\} = \prod_{i=1}^{n} P\{X_i \in A_i\}}$$

The obvious conditions for the joint probability mass functions (in the discrete case) and joint densities (in the continuous case) hold. That is, in both cases independence is equivalent the join probability mass function or density being equal to the respective product of the marginals.

EXAMPLES:

1. Let the joint density of X and Y be given by

$$f(x,y) = 6e^{-2x}e^{-3y} \quad \text{for} \quad x, y > 0$$

and zero otherwise. Are these random variables independent?

1. Let the joint density of X and Y be given by

$$f(x,y) = 6e^{-2x}e^{-3y} \quad \text{for} \quad x,y > 0$$

and zero otherwise. Are these random variables independent?

Solution: We compute the marginals. For $x > 0$ we have

$$f_X(x) = \int_0^\infty 6e^{-2x}e^{-3y}dy = 2e^{-2x}$$

For Y we have

$$f_Y(y) = \int_0^\infty 6e^{-2x}e^{-3y}dx = 3e^{-3y}$$

Therefore, $f(x,y) = f_X(x)f_Y(y)$ for all $x, y \in \mathbb{R}$ (note that the relation also holds if one, or both, of x and y are negative). Also, X is Exp(2), and Y is Exp(3).

2. Let the joint density of X and Y be given by

$$f(x,y) = 24xy \quad \text{for} \quad 0 < x < 1, \ 0 < y < 1, \ 0 < x + y < 1$$

and zero otherwise. Are X and Y independent?

2. Let the joint density of X and Y be given by

$$f(x,y) = 24xy \quad \text{for} \quad 0 < x < 1,\ 0 < y < 1,\ 0 < x+y < 1$$

and zero otherwise. Are X and Y independent?

Solution (see the book for another way to solve): Note that the domain in which f is non-zero depends upon an interplay between x and y. Therefore, you expect that these are not independent. Let us compute the marginals. For $x \in (0,1)$ we have $f(x,y) = 0$ unless $0 < y < 1-x$. Therefore,

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{0}^{1-x} 24xydy = 12x(1-x)^2$$

Similarly,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{0}^{1-y} 24xydx = 12y(1-y)^2$$

Therefore, we do *not* have $f(x,y) = f_X(x)f_Y(y)$ that for all x,y and hence X and Y are not independent.

3. For $i = 1,\ldots,n$, let $X_i \sim \text{Exp}(\lambda_i)$ be *independent* exponential random variables with parameters $\lambda_i > 0$. What is the distribution of the *minimum* of the X_i?

Solution: Let $Z = \min\{X_i\}$. Clearly, the range of Z is $t > 0$. Therefore

$$P\{Z > t\} = P\{\min\{X_i\} > t\} = P\{X_1 > t,\ X_2 > t, \ldots, X_n > t\}$$

$$= P\{X_1 > t\}P\{X_2 > t\}\ldots P\{X_n > t\} \text{ (by independence)}$$

$$= e^{-\lambda_1 t}e^{-\lambda_2 t}\ldots e^{-\lambda_n t} = \exp\left\{-\left(\sum_{i=1}^{n}\lambda_i\right)t\right\}$$

It follows that for $t \geq 0$ we have

$$F_Z(t) = 1 - P\{Z > t\} = 1 - \exp\left\{-\left(\sum_{i=1}^{n}\lambda_i\right)t\right\}$$

therefore $Z \sim \text{Exp}(\lambda_1 + \ldots + \lambda_n)$.

One natural way to interpret this result is the following: if n alarm clocks are set to go off after an exponentially distributed amount of time (each with a potentially different rate λ_i), then the time that the *first* alarm rings is also exponentially distributed with the parameter which equals to the sum of the parameters of all the clocks, that is $\sum_i \lambda_i$.