Chapter Two

Joint Marginal, and Conditional Distributions

The definition of a random vector is a straightforward generalization of the concept of random variables.

Def: Given the probability space $(\Omega, \mathcal{A}, P)$, an n- dimensional random vector $\boldsymbol{X}=(X_1, X_2, \ldots, X_n)$ is a function X: $\Omega \rightarrow \boldsymbol{R}^n$ such that $\forall\, \omega \in \Omega$, we have $\boldsymbol{X}(\omega)=(X_1, X_2, \ldots, X_n)(\omega)=(X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$.

The vector valued function $\boldsymbol{X}$ is a random vector if and only if its components X_i is a random variable.

In other words, a random vector is a measurable function from Ω to $\boldsymbol{R}^n$ just as a random variable is a measurable function from Ω to the real line $\boldsymbol{R}$.

Discrete Joint Distributions:

Def: If X and Y are discrete random variables, the function given by $f_{X,Y}(x, y) = p(X=x, Y=y)$ for each pair of values within the range of X and Y is called the joint probability function of X and Y.

A bivariate function can serve as the joint probability distribution of a pair of discrete random variables X and Y if and only if its values, $f_{X,Y}(x, y)$, satisfy the conditions:

1. $f_{X,Y}(x, y) \geq 0$ for each pair of values (x, y) within its domain.
2. $\sum\sum f_{X,Y}(x, y) = 1$ where the double summation extends over all possible pairs (x, y) within its domain.

Example 1:

Let X and Y be random variables with the joint distribution:

$$f_{X,Y}(x,y)=\frac{x+y}{21} \qquad x=1,2,3 \quad y=1,2$$

$$f(1,1)=p(X=1,Y=1)=\frac{2}{21}$$

$$f(1,2)=p(X=1,Y=2)=\frac{3}{21}$$

$$f(2,1)=p(X=2,Y=1)=\frac{3}{21}$$

$$f(2,2)=p(X=2,Y=2)=\frac{4}{21}$$

$$f(3,1)=p(X=3,Y=1)=\frac{4}{21}$$

$$f(3,2)=p(X=3,Y=2)=\frac{5}{21}$$

X \ Y	1	2	sum
1	2/21	3/21	5/21
2	3/21	4/21	7/21
3	4/21	5/21	9/21
sum	9/21	12/21	1

$$\sum_{x=1}^{3}\sum_{y=1}^{2} f(x,y)=\sum_{x=1}^{3}\sum_{y=1}^{2}\frac{x+y}{21}=\frac{2}{21}+\frac{3}{21}+\ldots+\frac{5}{21}=1$$

Example 2:

Determine the value of *k* for which the function given by:

f(x,y)= kxy for *x = 1,2,3 ; y = 1,2,3* , can serve as a joint probability distribution.

f(1,1) =k *f(2,1) = 2k* *f(3,1) = 3k*

f(1,2) = 2k *f(2,2) = 4k* *f(3,2) = 6k*

f(1,3) = 3k *f(2,3) = 6k* *f(3,3) = 9k*

K + 2k + 3k + 2k + 4k + 6k + 3k + 6k + 9k = 1

36k = 1 $\therefore$ *k = 1/36*

<u>Joint Distribution Function:</u>

Def: If *X* and *Y* are discrete random variables, the function given by:

$$F(x,y)=p(X\le x, Y\le y)=\sum_{s\le x}\sum_{t\le y} f(s,t) \qquad for \quad -\infty<x<\infty \quad -\infty<y<\infty$$

where *f(s,t)* is the value of the joint probability distribution of *X* and *Y* at *(s,t),* is called the joint distribution function, or the joint cumulative distribution, of X and Y.

Properties of bivariate cumulative distribution function

1. $F\left(-\infty, y\right) = \lim_{x\to-\infty} F(x,y) = 0$ for all y,

$F\left(x, -\infty\right) = \lim_{y\to-\infty} F(x,y) = 0$ for all x, and

$\lim_{\substack{x\to\infty \\ y\to\infty}} F(x,y) = F(\infty,\infty) = 1$

2. $If\ x_1 < x_2 and\ y_1 < y_2, then\ P[x_1 < X \le x_2, y_1 < Y \le y_2] =$
$F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1) \ge 0$

3. $F(x,y)$ is right continuous in each argument. That is,

$\lim_{0<h\to0} F(x+h,y) = \lim_{0<h\to0} F(x,y+h) = F(x,y)$

Example: With reference to example 1

$$f(x,y) = \frac{x+y}{21} \qquad x = 1,2,3 \quad y = 1,2$$

$$F(1,1) = p(X \le 1, Y \le 1) = f(1,1) = \frac{2}{21}$$

$$F(1,2) = p(X \le 1, Y \le 2) = f(1,1) + f(1,2) = \frac{2}{21} + \frac{3}{21} = \frac{5}{21}$$

$$F(2,2) = p(X \le 2, Y \le 2) = f(1,1) + f(1,2) + f(2,1) + f(2,2) = \frac{2+3+3+4}{21} = \frac{12}{21}$$

$$F(-2,1) = p(X \le -2, Y \le 1) = 0$$

$$F(3.7,4.5) = p(X \le 3.7, Y \le 4.5) = 1$$

<u>Joint Probability Density Function:</u>

Def: A bivariate function with values $f_{X,Y}(x,y)$, defined over the *xy*-plane is called the joint pdf of the continuous random variables X and Y if and only if:

$$p[(X,Y) \in A] = \iint_A f_{X,Y}(x,y)\,dx \quad dy$$

for any region A in the xy-plane.

A bivariate function can serve as a joint *pdf* of a pair of continuous random variables X and Y if its values, $f(x, y)$, satisfy the conditions

1. $\forall\ (x,y) \in \mathbb{R}^2,\ f_{X,Y}(x, y) \geq 0$ 2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = 1$

Example:

Show that the following function is a joint *pdf*.

$$f(x,y) = \begin{cases} \frac{3}{5}x(y+x) & for \quad 0 < x < 1, 0 < y < 2 \\ 0 & elsewhere \end{cases}$$

And find $p(0 < x < 1/2,\ 1 < y < 2)$.

$f(x, y) \geq 0 \qquad for\ -\infty < x < \infty,\ -\infty < y < \infty$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = \int_{0}^{2}\int_{0}^{1} \frac{3}{5}x(y+x) \quad dxdy$$

$$= \frac{3}{5}\int_{0}^{2}\int_{0}^{1}(xy + x^2)dxdy \qquad = \frac{3}{5}\int_{0}^{2} y\frac{x^2}{2} + \frac{x^3}{3} \ ]_0^1 dy$$

$$= \frac{3}{5}\int_{0}^{2}(\frac{y}{2} + \frac{1}{3})dy \qquad = \frac{3}{5}[\frac{y^2}{4} + \frac{y}{3}]_0^2 \qquad = \frac{3}{5}[\frac{4}{4} + \frac{2}{3}]$$

$$= \frac{3}{5}[1 + \frac{2}{3}] = \frac{3}{5}[\frac{5}{3}] = 1$$

$\therefore\ f(x, y)$ is a joint *pdf*

$$p(0 < x < \frac{1}{2}, 1 < Y < 2) = \int_1^2 \int_0^{\frac{1}{2}} \frac{3}{5} x(y+x)dxdy = \frac{3}{5} \int_1^2 \int_0^{\frac{1}{2}} (xy + x^2)dxdy$$

$$= \frac{3}{5} \int_1^2 y \frac{x^2}{2} + \frac{x^3}{3}]_0^{\frac{1}{2}} dy \quad = \frac{3}{5} \int_1^2 (\frac{y}{8} + \frac{1}{24})dy$$

$$= \frac{3}{5}[\frac{y^2}{16} + \frac{y}{24}]_1^2 \qquad = \frac{3}{5}[\frac{4}{16} + \frac{2}{24} - \frac{1}{16} - \frac{1}{24}]$$

$$= \frac{3}{5}[\frac{3}{16} + \frac{1}{24}] \qquad = \frac{3}{5}[\frac{9+2}{48}] = \frac{11}{80}$$

Joint Distribution Function:

Def: If *X* and *Y* are continuous random variables, $\forall\ (x, y) \in \boldsymbol{R}^2$ the function given by:

$$F_{X,Y}(x,y) = p(X \le x, Y \le y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(s,t)dsdt$$

where, $f(s, t)$ is the value of the joint *pdf* of *X* and *Y* at (s, t), is called the joint distribution function of *X* and *Y*.

Analogous to the relationship

$$f(x) = \frac{dF(x)}{dx}$$

Partial differentiation of $F_{X,Y}(x,y)$ leads to:

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

wherever these partial derivatives exist.

Example:

If the joint pdf of X and Y is given by:

$$f(x,y)=\begin{cases} x+y & 0<x<1, 0<y<1 \\ 0 & elsewhere \end{cases}$$

1. Find the distribution function of X and Y.

$$F(x,y)=\int_0^y\int_0^x (s+t) \quad dsdt = \int_0^y (\frac{s^2}{2}+st)\ \Big|_0^x dt$$

$$=\int_0^y (\frac{x^2}{2}+xt) \quad dt \quad = t\frac{x^2}{2}+x\frac{t^2}{2}\ \Big|_0^y$$

$$= y\frac{x^2}{2}+x\frac{y^2}{2} \quad = \frac{1}{2}xy(x+y)$$

$$\therefore \qquad F(x,y)=\begin{cases} 0 & x\le 0, y\le 0 \\ \frac{1}{2}xy(x+y) & 0<x<1, 0<y<1 \\ 1 & x\ge 1, y\ge 1 \end{cases}$$

2. Find $p(0.5 < x < 2, -1 < y < 1.5)$.

$$p(\frac{1}{2}<x<2, -1<y<\frac{3}{2}) = \int_0^1\int_{\frac{1}{2}}^1 (x+y) \qquad dxdy$$

$$=\int_0^1 \frac{x^2}{2}+xy\ \Big|_{\frac{1}{2}}^1 dy \qquad = \int_0^1 (\frac{1}{2}+y-\frac{1}{8}-\frac{y}{2})dy$$

$$=\int_0^1 (\frac{3}{8}+\frac{y}{2})dy \quad = \frac{3}{8}y+\frac{y^2}{4}\ \Big|_0^1$$

$$=\frac{3}{8}+\frac{1}{4}=\frac{5}{8}$$

Example:

Find the joint pdf of the random variables X and Y whose joint distribution function is:

$$F_{X,Y}(x,y)=\begin{cases}(1-e^{-x})(1-e^{-y}) & x>0, y>0\\ 0 & elsewhere\end{cases}$$

Also find p(1 < x < 3, 1 < y < 2).

Since partial differentiation yields:

$$\frac{\partial^2 F(x,y)}{\partial x \partial y}=f(x,y)$$

$$\frac{\partial F(x,y)}{\partial x}=(1-e^{-y})e^{-x}$$

$$\frac{\partial^2 F(x,y)}{\partial x \partial y}=e^{-x}e^{-y}=e^{-(x+y)}$$

$$f(x,y)=\begin{cases}e^{-(x+y)} & x>0, y>0\\ 0 & elsewhere\end{cases}$$

$$p(1<x<3,1<y<2)=\int_1^2\int_1^3 e^{-(x+y)} \quad dxdy$$

$$=\int_1^2 -e^{-y}e^{-x}\ \Big|_1^3 dy \qquad =\int_1^2 -e^{-y}[e^{-3}-e^{-1}]dy$$

$$[e^{-3}-e^{-1}]e^{-y}\ \Big|_1^2 \qquad =[e^{-3}-e^{-1}][e^{-2}-e^{-1}]$$

$$=e^{-5}-e^{-4}-e^{-3}+e^{-2}=0.074$$