5 Jointly Distributed Random Variables

Probability statements concerning two or more random variables.

- For any two random variables X & Y,
the joint cumulative probability distribution function of X & Y is

$$\mathrm{F(a,\ b) = P\{X \leq a,\ Y \leq b\}} \qquad -\infty < a, b < \infty$$

- $\mathrm{F}_X\mathrm{(a) = P\{X \leq a\}}$

$$= \mathrm{P\{X \leq a,\ Y \leq \infty\}}$$

$$= \mathrm{P}\left(\lim_{b\to\infty} \mathrm{\{X \leq a,\ Y \leq b\}}\right)$$

$$= \lim_{b\to\infty} \mathrm{P\{X \leq a,\ Y \leq b\}}$$

$$= \lim_{b\to\infty} \mathrm{F(a,\ b)}$$

$$= \mathrm{F(a,\ \infty)}$$

$\mathrm{F}_Y\mathrm{(b) = P\{Y \leq b\}}$

$$= \lim_{a\to\infty} \mathrm{F(a,\ b)}$$

$$= \mathrm{F(\infty,\ b)}$$

F_X, F_Y : called the "Marginal Distributions" of X & Y.

- All joint probability statements about X and Y can, in theory be answered in terms of their joint distribution funtions.

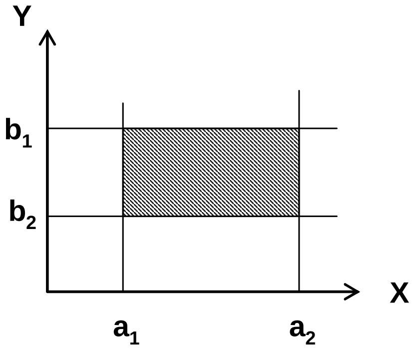
$$\mathrm{P\{X > a,\ Y > b\} = 1 - P(\{X > a,\ Y > b\}^c)}$$

$$= \mathrm{1 - P(\{X > a\}^c \cup \{Y > b\}^c)}$$

$$= \mathrm{1 - P(\{X \leq a\} \cup \{Y \leq b\})}$$

$$= 1 - [P\{X \leq a\} + P\{Y \leq b\} - P\{X \leq a, Y \leq b\}]$$

$$= 1 - F_X(a) - F_Y(b) + F_x(a, b)$$



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

• When X and Y are both discrete random variables, the joint probability mass funtion of X and Y

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

Then,

$$P_X(x) = P\{X = x\} = \sum_{y:p(x,y)>0} p(x,y)$$

$$P_Y(y) = P\{Y = y\} = \sum_{x:p(x,y)>0} p(x,y)$$

Example .

3 Red Balls
4 White Balls
5 Blue Balls

Select three balls Randomly

X: Numbers of Red chosen
Y: Numbers of White chosen

$$p(i, j) = P\{X = i, Y = j\}$$

$$p(0,\,0) = \frac{\binom{5}{3}}{\binom{12}{3}}$$

$$p(0,\,1) = \frac{\binom{4}{1}\binom{5}{2}}{\binom{12}{3}}$$

$$p(0,\,2) = \frac{\binom{4}{2}\binom{5}{1}}{\binom{12}{3}}$$

.

.

.

i/j	0 1 2 3	Column Sum = P{X=i}
0		
1		
2		
3		
Column Sum = P{Y=j}		

• X and Y are jointly continuous if their exist a function $f(x, y)$ define for all real x and y, and for every set C of pairs of real numbers

$$\mathrm{P}(\{x, y\}\ \epsilon\ \mathrm{C}) = \int\int_{(x,y)\epsilon C} f(x, y)dxdy$$

$f(x, y)$: The joint probability density funtion of X and Y.

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

- $\mathrm{F(a,\ b)} = \int_{-\infty}^{b}\int_{-\infty}^{a} f(x, y)dxdy$

After Differentiation

$$f(a,b) = \frac{\partial^2}{\partial a \partial b}\text{F(a,b)}$$

OR

- $\text{P}\{\text{a<X a+da, b<Y<b+db}\} = \int_b^{b+db}\int_a^{a+da} f(x,y)dxdy \approx f(a,b)dadb$

• If X and Y are jointly continuous, they are individually continuous, and their probability density functions are;

- $\text{P}\{\text{X}\epsilon\text{A}\} = \int_A\int_{-\infty}^{\infty} f(x,y)dxdy = \int_A f_X(x)dx$

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

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

Example .

$$f(x,y) = \begin{cases} e^{-(x,y)} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Density function of the Random Variable X/Y = ?

Solution.

$$F_{\frac{X}{Y}}(\text{a}) = P\left\{\frac{X}{Y} \leq a\right\}$$

$$= \int\int_{x/y\leq a} e^{-(x+y)}dxdy$$

$$= \int_0^{\infty}\int_0^{ay} e^{-(x+y)}dxdy$$

$$= \int_0^{\infty}(1-e^{-ay})e^{-y}dy$$

$$= \left(-e^{-y} + \frac{e^{-(a+1)y}}{a+1}\right)\Bigg|_0^{\infty}$$

$$= 1 - \frac{1}{a+1}$$

After Differentiation

$$f_{X/Y}(\text{a}) = \frac{1}{(a+1)^2} \qquad 0 < \text{a} < \infty$$

• Distribution function of the n random variables $X_1, X_2, ..., X_n$.

$$\text{F}(a_1, a_2, ..., a_n) = \text{P}\{X_1 \leq a_1,\ ......,\ X_n \leq a_n\}$$

• n random variables are jointly continues if there exist a function $f(x_1, x_2, ..., x_n)$ such that for any set C in n space.

$$\text{P}\{(X_1, ..., X_n)\epsilon\ \text{C}\} = \int\int .....\int_{(x_1,....,x_n)\epsilon C} f(x_1, ..., x_n)dx_1....dx_n$$

5.1 Independent Random Variables

Definition:

The Random Variables X and Y are independent if for any two sets of real numbers A and B,

$$\text{P}\{\text{X}\epsilon\text{A},\ \text{Y}\epsilon\text{B}\} = \underbrace{\text{P}\{\text{X}\epsilon\text{A}\}\ \text{P}\{\text{Y}\epsilon\text{B}\}}_{these\ events\ are\ independent}$$

$$\text{P}\{\text{X}\leq\text{a},\ \text{Y}\leq\text{b}\} = \text{P}\{\text{X}\leq\text{a}\}\ \text{P}\{\text{Y}\leq\text{b}\}$$

$$\text{i.e. } \text{F(a, b)} = \text{F}_X(a)\text{F}_Y(b) \qquad \text{for all } a, b$$

• For independent discrete random variables X and Y,

$$p(x, y) = \text{P}_X(x)\text{P}_Y(y) \qquad \text{for all } x, y$$

$$\because\ \text{P}\{\text{X}\epsilon\text{A},\ \text{Y}\epsilon\text{B}\} = \sum_{y\epsilon B}\sum_{x\epsilon A} p(x, y) = \sum_{y\epsilon B}\sum_{x\epsilon A} \text{P}_X(x)\text{P}_Y(y)$$

$$= \sum_{y\epsilon B} \text{P}_Y(y) \sum_{x\epsilon A} \text{P}_X(x) = \text{P}\{\text{Y}\epsilon\text{B}\}\ \ \text{P}\{\text{X}\epsilon\text{A}\}$$

• In the jointly continuous case, independence $\Leftrightarrow f(x, y) = f_X(x)\ f_Y(y)$ for all x, y.

• Random Variables that are not independent are said to be dependent.

Example .

A & B decide to meet at a certain location.
Each independently arrives at a time uniformly distributed between 12 noon and 1 pm.
Probability that the first to arrive has to wait longer than 10 minutes ?

Solution.
X : The time (min) past 12 that A arrives
Y : The time (min) past 12 that B arrives
X & Y are independent, each is uniform over (0, 60).

$$\mathrm{P}\{\mathrm{X}+10<\mathrm{Y}\} + \mathrm{P}\{\mathrm{Y}+10<\mathrm{X}\}$$

$$\because \text{Symmetry}$$

$$= 2\mathrm{P}\{\mathrm{X}+10<\mathrm{Y}\}$$

$$= 2\int\int_{x+10<y} f(x,y)dxdy$$

$$= 2\int\int_{x+10<y} f_X(x)f_Y(y)dxdy$$

$$= 2\int_{10}^{60}\int_{0}^{y-10}\left(\frac{1}{60}\right)^2 dxdy$$

$$= \frac{25}{36}.$$

Example 2.e

A bullet is fired at a target
X : horizontal miss distance
Y : vertical miss distance

Assume,

1. X & Y are independent continuous random variables with differentiable density functions

2. $f(x,y) = f_X(x)f_Y(y) = g(x^2+y^2)$ (for some funtion g)

Differentiate

$$f_X'(x)f_Y(y) = 2xg'(x^2+y^2)$$

$$\frac{f'_X(x)}{2xf_X(x)} = \frac{g'(x^2+y^2)}{g(x^2+y^2)} = \text{Constant}$$

$$\because \text{ for any } x_1, x_2, y_1, y_2, \text{ such that } \; {x_1}^2 + {y_1}^2 = {x_2}^2 + {y_2}^2$$

$$\frac{f'_X(x_1)}{2x_1 f_X(x_1)} = \frac{f'_X(x_2)}{2x_2 f_X(x_2)} = \text{Constant}$$

$$\therefore \log f_X(x) = a + \frac{cx^2}{2}$$

$$\Rightarrow f_X(x) = \text{K } e^{\frac{cx^2}{2}}$$

$$\int_{-\infty}^{\infty} f_X(x)dx = 1 \quad \Rightarrow \quad \text{C needs to be negative.}$$

$$\text{Let, C = -1}/\sigma^2 \quad \Rightarrow \quad \text{X} \sim N(0,\, \sigma^2)$$

$$\text{Similarly,} \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\overline{\sigma}}\, e^{\frac{-y^2}{2\sigma^{-2}}}$$

Assumption $\Rightarrow$ $\overline{\sigma} = \sigma$.

$\therefore$ X and Y are independent identically distributed (iid) normal random variables $\sim N(0,\, \sigma^2)$.

Proposition

The continuous (discrete) random variables X&Y are independent, then their joint probability density function can be expressed as;

$$f_{X,Y}(x,y) = h(x)g(y) \qquad -\infty < x < \infty, \; -\infty < y < \infty$$

Proof.

- Independent $\Rightarrow$ Then factorization will hold.
- Suppose $f_{X,Y}(x,y) = h(x)g(y)$, then

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy \\ &= \underbrace{\int_{-\infty}^{\infty} h(x)dx}_{C_1}\underbrace{\int_{-\infty}^{\infty} g(y)dy}_{C_2} \end{aligned}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = C_2\ h(x)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = C_1\ g(y)$$

$$C_1C_2 = 1 \quad \Rightarrow \quad f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

• The n random variables $X_1, ..., X_n$ are independent if, for all sets of real numbers $A_1, ..., A_n$,

$$\mathrm{P}\{X_1 \epsilon A_1,\ .....,\ X_n \epsilon A_n\} = \prod_{i=1}^{n} \mathrm{P}\{X_i \epsilon A_i\}$$

This is equivalent to,

$$\mathrm{P}\{X_1 \le a_1,\ .....,\ X_n \le a_n\} = \prod_{i=1}^{n} \mathrm{P}\{X_i \le a_i\}$$

• An infinite collection of random variables are independent if every finite sub-collection of them are independent.

Example .

X, Y & Z are independently uniform over (0, 1).
P{X≥YZ} = ?

Solution.

$$f_{X,Y,Z}(x,y,z) = f_X(x)f_Y(y)f_Z(z) = 1 \qquad 0 \le x, y, z \le 1$$

$$\mathrm{P}\{\mathrm{X} \ge \mathrm{YZ}\} = \int\int\int_{x \ge yz} f_{X,Y,Z}(x,y,z)dxdydz$$

$$= \int_0^1 \int_0^1 \int_{yz}^1 dxdydz$$

$$= \int_0^1 \int_0^1 (1 - yz)dydz$$

$$= 3/4.$$