The correlation function of X and Y is defined as

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

$$= \frac{\mathbb{E}\left[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right]}{\sqrt{\mathbb{E}\left[(X-\mathbb{E}(X))^2\right]\,\mathbb{E}\left([Y-\mathbb{E}(Y)]^2\right)}}.$$

Conditional moments.

As conditional distribution function $F_{X|Y}(x)$ is a random variable (randomness is induced by the random variable Y), it is natural to imagine that conditional moments are also random. Consider a function (continuous if the bivariate random vector is of continuous type) $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Then we define the random variable (of continuous type)

$$\mathbb{E}(g(X,Y)\,|\,Y) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})).$$

Such that $\mathbb{E}(g(X,Y)|Y)$ takes the value

$$\mathbb{E}_{Y=y}(g(X,Y)\,|\,Y) = \mathbb{E}(g(X,Y)\,|\,y)$$

$$= \mathbb{E}(g(X,Y)\,|\,Y=y)$$

$$= \begin{cases} \int_{-\infty}^{\infty} dx \; f_{X|Y}(x|y)\, g(x,y) & (X,Y) \text{ is a random vector of continuous type} \\ \sum_{i=1}^{\infty} g(x_i,y)\, p_{X|Y}(x_i|y) & (X,Y) \text{ is a random vector of discrete type and } p_{X|Y}(x_i|y)>0. \text{ that is, } p_Y(y)>0. \end{cases}$$

with conditional probability density function $f_{X|Y}(x|y)$ and mass function $p_{X|Y}(x_i|y)$.

As we have seen that there is a relation between conditional probability density/mass function, marginal probability density/mass function and joint probability density/mass function, it is natural to ~~see~~ imagine that those relations can be used to obtain some ~~forma~~ formulae ~~&~~ which ~~reat~~ relate the moments of (X,Y) to the conditional moments.

We can show that $\mathbb{E}(g(X,Y))$ exists that is, $\mathbb{E}(|g(X,Y)|)<\infty$, then

$$\mathbb{E}(g(X,Y)) = \mathbb{E}\left[\mathbb{E}(g(X,Y)\,|\,Y)\right].$$

$$= \int_{-\infty}^{\infty} dy \, f_Y(y) \, \mathbb{E}\left(g(X,Y) \mid Y=y\right)$$

$$= \begin{cases} \int_{-\infty}^{\infty} dy \, f_Y(y) \left[\int_{-\infty}^{\infty} dx \, f_{X|Y}(x|y) \, g(x,y) \right] \\ \sum_{j=1}^{\infty} p_Y(y_j) \sum_{i=1}^{\infty} p_{X|Y}(x_i|y_j) \, g(x_i, y_j) \end{cases}$$

Similarly, we can use the following formula

$$\mathbb{E}\left(g(X,Y)\right) = \mathbb{E}\left[\mathbb{E}\left(g(X,Y) \mid X\right)\right].$$

Properties of conditional expectation

i) For any constant c, $\mathbb{E}(c|Y) = c$ almost surely.

"$X = c$ almost surely" means "$\mathbb{P}(X = c) = 1$".

ii) **Linearity**.

$$\mathbb{E}[a_1 g_1(X) + a_2 g_2(X) \mid Y] = a_1 \mathbb{E}(g_1(X)|Y) + a_2 \mathbb{E}(g_2(X)|Y)$$

almost surely (with probability one).

iii) $\mathbb{P}(X \geq 0) = 1$ then $\mathbb{E}(X|Y) \geq 0$. with probability 1.

iv) If $\mathbb{P}(g(X,Y) \geq 0) = 1$, then $\mathbb{E}(g(X,Y)|Y) \geq 0$ with probability 1.

v) If $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $\mathbb{E}(|\phi(X,Y)|) < \infty$, then

$$\mathbb{E}[\phi(X,Y) \mid Y = y] = \mathbb{E}[\phi(X,y) \mid Y = y]$$

vi) $\mathbb{E}[\psi(X)\, \phi(X,Y) \mid X] = \psi(X)\, \mathbb{E}(\phi(X,Y)|X)$

with probability 1.

vii) The r^{th} conditional moment ~~given~~ of X given Y is defined to be $\mathbb{E}(X^r | Y)$.

viii) Let $\mathbb{E}(h(X))$ exists that is, $\mathbb{E}(|h(X)|) < \infty$, then

$$\mathbb{E}(h(X)) = \mathbb{E}[\mathbb{E}(h(X)|Y)].$$

ix) If $\mathbb{E}(X^2) < \infty$, then

$$\mathrm{Var}(X) = \mathrm{Var}(\mathbb{E}(X|Y)) + \mathbb{E}(\mathrm{Var}(X|Y)).$$

Lecture - XI   Independent random vectors, moments and moment generating functions.

Consider two distribution functions $F_1 : \mathbb{R} \to [0,1]$ and $F_2 : \mathbb{R} \to [0,1]$. Define a function

$F : \mathbb{R} \times \mathbb{R} \longrightarrow [0,1]$ such that $F(x,y) = F_1(x) F_2(y)$

for all $(x,y) \in \mathbb{R}^2$.

<u>Exercise</u> Check that $F : \mathbb{R}^2 \to [0,1]$ is a two-dimensional distribution function.

According to this exercise, there exists a bivariate random vector (X, Y) such that

$\mathbb{P}(\{X \le x\} \cap \{Y \le y\}) = F(x,y) = F_1(x) F_2(y)$.

for all $(x,y) \in \mathbb{R}^2$. This means that there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $(X,Y) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ is a bivariate random vector.

As the distribution function uniquely charaterizes the law of a random vector, it is immediate to see that for any $A \times B \in \mathcal{B}(\mathbb{R}^2)$, we should have

$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(X \in A)\, \mathbb{P}(Y \in B). \quad \longrightarrow (*)$$

Let $(X, Y) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ be a bivariate random vector. Then we call X and Y to be independent if for every $A \times B \in \mathcal{B}(\mathbb{R}^2)$

$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B) \longrightarrow (*_1)$$

The components of an n-dimensional random vector $(X_1, X_2, \ldots, X_n) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ are called independent if for every $A_1 \times A_2 \times \cdots \times A_n \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i \in A_i\}\right) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i) \longrightarrow (*_2)$$

Let $F_{X_1, \ldots, X_n} : \mathbb{R}^n \longrightarrow [0,1]$ be the distribution function of the random vector $(X_1, X_2, \ldots, X_n)$.

Then it follows from $(*_2)$ that

$$F_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i) \longrightarrow (*_3)$$

where $F_{X_i} : \mathbb{R} \longrightarrow [0,1]$ is the distribution function of X_i for all $i = 1, 2, \ldots, n$.

<u>Theorem</u>:

Consider n distribution functions $(F_i : 1 \le i \le n)$.

Define a function $F : \mathbb{R}^n \longrightarrow [0,1]$ such that

$$F(x_1, x_2, \cdots x_n) = \prod_{i=1}^{n} F_i(x_i) \text{ for every } (x_1 \ldots, x_n) \in \mathbb{R}^n.$$

Then there exists an n-dimensional random vector $(X_1, \ldots X_n) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ with distribution function given by F.

~~Probabab~~ Probability mass/density function of independent random variables.

Let (X, Y) be a random vector of discrete type.

Then

$$\mathbb{P}(X = x_i, \cancel{\&} Y = y_j) = \mathbb{P}(X = x_i)\,\mathbb{P}(Y = y_j)$$
$$= p_X(x_i)\, p_Y(y_j)$$

Let (X, Y) be a random vector of continuous type with density function $f_{X,Y}(x,y)$. Then it follows that

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y) \quad \forall \ (x,y) \in \mathbb{R}^2$$

where f_X and f_Y are the density functions of X and Y respectively.

joint mass/~~des~~ density function is the product of marginal mass/density functions if the random variables are independent.

Pairwise independent random variables.

Let $(X_1, X_2, \dots, X_n) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ be an n-dimensional random vector. We call the random variables $(X_1 \dots X_n) = (X_i : 1 \le i \le n)$ to be pairwise independent if ~~and only if~~ for every $i \ne j$, we have

$$\mathbb{P}(\{X_i \in A\} \cap \{X_j \in B\}) = \mathbb{P}(X_i \in A)\, \mathbb{P}(X_j \in B)$$

for every $A \times B \in \mathcal{B}(\mathbb{R}^2)$.

Remark Pairwise independent random variables might not be independent completely.

Theorem

Let $(X, Y) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ be a random vector. Suppose that $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are Borel measurable maps. Then $f(X)$ and $g(Y)$ are also independent random variables.

Identically distributed random variables

Let X and Y be two random variables (may not be defined on the same probability space). Then we say X and Y are identically distributed if for every $A \in \mathcal{B}(\mathbb{R})$, $\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$.

Independently and identically distributed random variables

We call components of an n-dimensional random vector $(X_1, \dots, X_n) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ to be ~~indepoently~~ ~~indepently~~ independently and identically distributed if the components are independent and identically distributed.

Independent random vectors.

Let $(\underline{X}, \underline{Y}) = (X_1, \dots, X_m, Y_1, Y_2, \dots, Y_n) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}^{m+n}, \mathcal{B}(\mathbb{R}^{m+n}))$ be random vector. We call the random vectors $\underline{X} = (X_1, \dots, X_m)$ and $\underline{Y} = (Y_1 \dots Y_n)$ to be independent if for every $A \in \mathcal{B}(\mathbb{R}^m)$ and $B \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\mathbb{P}\left((\underline{X}, \underline{Y}) \in A \times B \right) = \mathbb{P}\left(\{\underline{X} \in A\} \cap \{\underline{Y} \in B\} \right)$$

$$= \mathbb{P}(\underline{X} \in A)\, \mathbb{P}(\underline{Y} \in B).$$

Some properties of i.i.d. random variables.

Suppose that $(X_1, X_2, \dots, X_n)$ be i.i.d. random variables. Then we can see that

$$F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} F_{X_1}(x_i) = F_{X_{i_1} \dots X_{i_n}}(x_1, x_2, \dots, x_n)$$

that is, $(X_1, \dots, X_n) \stackrel{d}{=} (X_{i_1}, X_{i_2}, \dots, X_{i_n})$