Lecture 8: Expectation, Covariance and Conditional Expectation

Here we shall look at finding the expectation of a ~~random~~ real-valued function of a random vector ~~Ω~~ $X : \Omega \to \mathbb{R}^k$ whose joint distribution is known.

Let $Y = g(X)$ be a function of random vector X and it is real valued, i.e. $g(X(\omega)) \in \mathbb{R}$, and $g : Ran(X) \to \mathbb{R}$.

We shall also study the way to measure the relation between two random variables X and Y, using an idea called covariance and correlation coefficient.

~~Sect Ex~~ Section 1: Expectation of $g(x)$

Let $X : \Omega \to \mathbb{R}^k$ be random-vector given as

$$X = (X_1, X_2, \dots, X_k)$$

If each $X_1, X_2 \dots X_k$ be discrete random variables and let $g : \mathbb{R}^k \to \mathbb{R}$, then expectation of $g(X)$ is given as.

$$E[g(X_1 \dots X_k)] = \sum g(x_1, x_2, \dots x_k)\, f_X(x_1 \dots x_k)$$

If $X_1, \dots, X_k$ ~~be~~ are continuous random variables, then

$$E[g(X_1, \dots, X_k)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(x_1 \dots x_k)\, f_X(x_1 \dots x_k)\, dx_1 \dots dx_k.$$

Theorem 8.1: Let X and Y be two continuous random variable with $f_{XY}(\cdot,\cdot)$, representing their joint density function. Then if $E(X) < \infty$ & $E(Y) < \infty$ and $E(X+Y) < \infty$ then

$$E(X+Y) = E(X) + E(Y).$$

[• Here $g(X,Y) = X+Y$. For two variables the density function is ~~give~~ written as $f_{X,Y}(x,y)$.]

Proof: $$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y) f_{X,Y}(x,y)\,dx\,dy$$

$$\therefore E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{X,Y}(x,y)\,dx\,dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty} x\left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\right]dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty}\left[f_{X,Y}(x,y)\,dx\right]dy$$

$$= \int_{-\infty}^{\infty} x\, f_X(x)\,dx + \int_{-\infty}^{\infty} y\, f_Y(y)\,dy$$ [Using the defn of marginal distribution]

$$\therefore E(X+Y) = E(X) + E(Y)$$ $\square$

Think how will you extend it to more than 2 variables. In fact one can show that

$$E\left[\sum_{i=1}^{k} X_i\right] = \sum_{i=1}^{k} E[X_i]. \longrightarrow (\#)$$

We urge the reader to prove the above result in Thm 8·1, for the discrete case.

What about $Var(X+Y)$, where X and Y are two random variables, with finite expectations.

$$\begin{aligned}
Var(X+Y) &= E\left[(X+Y - E(X+Y))^2\right] \\
&= E\left[(X+Y-E(X)-E(Y))^2\right] \quad \text{Using Thm 8·1} \\
&= E\left[(X-E(X) + (Y-E(Y))^2\right] \\
&= E\left[(X-E(X))^2 + (Y-E(Y))^2 + 2(X-E(X))(Y-E(Y))\right] \\
&= E\left[(X-E(X))^2\right] + E\left[(Y-E(Y))^2\right] + 2E\left[(X-E(X))(Y-E(Y))\right]
\end{aligned}$$

$\hookrightarrow$ Using (#)

$\therefore \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2E\left[(X-E(X))(Y-E(Y))\right]$

We shall now focus on the term

$$E\left[(X-E(X))(Y-E(Y))\right]$$

$$= E\left[XY - XE(Y) - YE(Y) + E(X)E(Y)\right]$$

$$= E[XY] - E[Y]\cancel{[X]}E[X] - E[Y]E[X] + E[X]E[Y]$$

$$= E[XY] - E[X]E[Y]$$

$$\therefore \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\left[E[XY] - E[X]E[Y]\right]$$

What happens if X and Y are independent. What happens to the term $E[XY] - E[X]E[Y]$ in such a case.

Let just check out this for the continuous case. We will first compute $E[XY]$, Thus

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_X(x) f_Y(y)\,dx\,dy \qquad \left[\begin{array}{l} f_{XY}(x,y) \\ = f_X(x) f_Y(y) \\ \text{as } X \text{ and } Y \\ \text{are independent} \end{array}\right]$$

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

$$= E[X]E[Y]$$

$$\therefore \boxed{E[XY] = E[X]E[Y]}$$

Theorem 8.2 Let X and Y are random variables, with finite expectations. ~~Then~~ If X and Y are independent then

$$E[XY] = E[X]E[Y].$$

Thus $E[XY] - E[X]E[Y] = 0$ if X and Y are independent

So we have the following: If X and Y are independent random variables, with finite expectations, then

$$\boxed{Var(X+Y) = Var(X) + Var(Y)}$$

In general we shall call the term

$$E[(X - E(X))(Y - E(Y))]$$

as the Covariance of X and Y as it seems to measure the expected joint variation of X and Y from ~~the mean~~ their mean values. In symbols

$$Cov(X,Y) = E[(X-E(X))(Y-E(Y))]$$

$$\therefore Cov(X,Y) = E[XY] - E[X]E[Y]$$

Thus:

$$\boxed{\text{If } X \text{ and } Y \text{ are independent, then } Cov(X,Y) = 0}$$

The converse is not true. We shall provide examples but before that we state, that for any two random variables X and Y

$$\boxed{Var(X+Y) = Var(X) + Var(Y) + 2\,Cov[X,Y]}$$

- Note Covariance is symmetric in X and Y. Thus

$\cancel{Cov[X,Y]}$

$$\boxed{Cov(X,Y) = Cov(Y,X)}$$

~~Cof~~

Here we give two examples. One for the discrete case and other for the continuous case to show that $E[XY] = E[X]E[Y]$ is a necessary condition for X and Y to be independent, but not sufficient.

Eg 1: (~~from Math. St~~ Freund's Math. Stat, by Miller & Miller)
Consider two discrete random variables X and Y, taking the values $-1, 0, +1$.

$y \backslash x$	-1	0	1	$P(Y=y)$
-1	$1/6$	$1/3$	$1/6$	$2/3$
0	0	0	0	0
1	$1/6$	0	$\frac{1}{6}$	$\frac{1}{3}$
$P(X=x)$	$1/3$	$1/3$	$1/3$	1 → Total prob.

The above probability table gives the joint distribution of X and Y.

We have $E(X) = -1 \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = 0$

& $E(Y) = -1 \times \frac{2}{3} + 0 \times 0 + 1 \times \frac{1}{3} = -\frac{1}{3}$

$$E(X,Y) = (-1).(-1)\frac{1}{6} + (-1).(0) \times \frac{1}{3} + (-1)(-1)\frac{1}{6}$$
$$+ 1(-1).\frac{1}{6} + (1.1).\frac{1}{6} = 0$$

$$\therefore \boxed{E(X,Y) - E(X)E(Y) = 0.}$$

Now consider $x = -1$, $y = -1$, then ~~$f(x,y) = \frac{1}{6}$~~

$f_X(x,y) = f_X(-1,-1) = \frac{1}{6}$

Now $f_X(x) = f_X(-1) = \frac{1}{3}$

$f_Y(y) = f_Y(-1) = 2/3$

$$\therefore \boxed{f_X(x) f_Y(y) = \frac{2}{9} \neq \frac{1}{6} = f_X(x,y)}$$ ⇒ X and Y are not independent.

E.g2: Let X & Y be two continuous random variables given as

$$X = \sin 2\pi U$$
$$Y = \cancel{\sin} \cos 2\pi U$$

where U is a uniformly distributed random variable in $(0,1)$; i.e.

$$f_U(u) = \begin{cases} 1 & \text{if } u \in (0,1) \\ 0 & \text{otherwise.} \end{cases}$$

If X is known then $U = \frac{1}{2\pi} \sin^{-1} X$, and thus Y is known.

Now $E[Y] = \int_0^1 \cos 2\pi u \, du = 0$ & $E[X] = \int_0^1 \sin 2\pi u \, du = 0$

Now $E[XY] = \int_0^1 \sin 2\pi u \cos 2\pi u \, du = 0$

Note that, $XY = \sin 2\pi U \cos 2\pi U = g(U)$

$$\therefore \quad E[XY] = E[g(U)] = \int_0^1 g(u) f_U(u) \, du.$$

Here $E[XY] - E[X]E[Y] = 0$, though X and Y are not independent. $\square$

Let us call two random variables X and Y un-correlated if $\operatorname{cov}(X,Y) = 0$.

From the above discussion we have learned the following

$$\boxed{\begin{array}{c} \text{If } X \text{ and } Y \text{ are independent} \Rightarrow X \text{ and } Y \text{ are uncorrelated} \\ \text{But the converse is not true} \end{array}}$$

We can normalize the correlation measure, i.e. bring in the value of $\operatorname{cov}(X,Y)$ with in the interval. This leads to the introduce the notion of a Pearson's correlation coefficient, denoted by σ_{XY} and ~~denoted~~ given as.

Thus we can write

$$\boxed{\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}}$$

Let us now focus on the following result, from Rohatgi [Statistical Inference Dover Pub. 2003.

Theorem ~~8.7~~ 8.3: If $E(X^2) < \infty$ & $E(Y^2) < \infty$, then $\mathrm{cov}(XY)$ ~~σ_{XY}~~ exists and X and Y ~~have~~ have finite mean, then.

$$[\mathrm{Cov}(XY)]^2 \leq \sigma_X^2 \sigma_Y^2$$

ie $$\boxed{[\mathrm{Cov}(XY)]^2 \leq \mathrm{Var}(X)\,\mathrm{Var}(Y)} \longrightarrow (\#)$$

Proof: Set $E(X) = \mu_1$ & $E(Y) = \mu_2$.

$$\therefore (x-\mu_1)(y-\mu_2) \leq \frac{(x-\mu_1)^2 + (y-\mu_2)^2}{2}$$

$\therefore$ In terms ~~or andom~~ of random variables, then

$$(X-\mu_1)(Y-\mu_2) \leq \frac{(X-\mu_1)^2 + (Y-\mu_2)^2}{2}.$$

$$\therefore E[(X-\mu_1)(Y-\mu_2)] \leq \frac{1}{2}\mathrm{Var}(X) + \frac{1}{2}\mathrm{Var}(Y)$$

$$\Rightarrow E[(X-\mu_1)(Y-\mu_2)] \leq \frac{1}{2}[\mathrm{Var}(X) + \mathrm{Var}(Y)].$$

$$\therefore \mathrm{Cov}(X,Y) \leq \frac{1}{2}[\sigma_X^2 + \sigma_Y^2],$$ where $\sigma_X^2 = \mathrm{Var}(X)$ & $\sigma_Y^2 = \mathrm{Var}(Y)$

Thus if $\sigma_X^2 < +\infty$ and $\sigma_Y^2 < +\infty$, then $\mathrm{Cov}(X,Y)$ is bounded above, and is also finite.

Let a & $b \in \mathbb{R}$, then

$$E[a(X-\mu_1) + b(X-\mu_2)]^2 = a^2\sigma_X^2 + 2ab\,\mathrm{Cov}(X,Y) + b^2\sigma_Y^2 \rightarrow (\odot)$$

Let us assume that either $\sigma_X^2 = 0$ or $\sigma_Y^2 = 0$, then (#) holds automatical

Now let us assume that $\sigma_X^2 > 0$ & $\sigma_Y^2 > 0$. The equation ⊙ holds for any a & $b \in \mathbb{R}$. Thus set

$$a = -\frac{Cov(X,Y)}{\sigma_X^2}$$

Now from ⊙

$$\frac{(Cov(X,Y))^2}{\sigma_X^4} \cdot \sigma_X^2 + b^2 \sigma_Y^2 - 2b \frac{Cov(X,Y)^2}{\sigma_X^2} \geq 0 \rightarrow (\because)$$

The above expression is non-negative since

$$E\left[a(X-\mu_1) + b(X-\mu_2)\right]^2 \geq 0$$

Put $b=1$ in $(\because)$, and then we get

$$\frac{(Cov(X,Y))^2}{\sigma_X^2} + \sigma_Y^2 \geq 2 \frac{Cov(X,Y)^2}{\sigma_X^2}$$

$$\Rightarrow \frac{Cov(X,Y)^2}{\sigma_X^2} \leq \sigma_Y^2$$

$$\Rightarrow \boxed{(Cov(X,Y))^2 \leq \sigma_X^2 \sigma_Y^2}$$

This completes the proof. □

<u>Theorem 8.4:</u> $-1 \leq \rho_{XY} \leq +1$, if $\sigma_X^2 > 0$, $\sigma_Y^2 > 0$ and finite.

Proof: From Theorem 8.3, $(Cov(X,Y))^2 \leq \sigma_X^2 \sigma_Y^2$

$$\Rightarrow \left(\frac{Cov(X,Y)}{\sigma_X \sigma_Y}\right)^2 \leq 1$$

$$\Rightarrow \rho_{XY}^2 \leq 1$$

$$\Rightarrow -1 \leq \rho_{XY} \leq 1. \quad □$$

Section 8.2: Bivariate Normal Distribution

Can a normal distribution be defined in a joint manner for random variables X and Y. The answer turns out to be yes and before we proceed let us say little more about covariance. Sometimes one uses the symbol ~~Co~~ σ_{XY} to denote $\mathrm{Cov}(X,Y)$. In this formalism, we have $\mathrm{Var}(X) = \mathrm{Cov}(X,X) = \cancel{\sigma_{X,X}}\,\sigma_{XX}$. Thus given any two random variables X and Y we call the following matrix Σ, ~~as the variance~~ given as

$$\Sigma = \begin{bmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{YX} & \sigma_{YY} \end{bmatrix}$$

as the variance-covariance matrix and often an helpful tool to represent variance. Let $a, b \in \mathbb{R}$, then

$$\begin{aligned} \mathrm{Var}(aX+bY) &= a^2 \mathrm{Var}(X) + b^2 \mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X,Y) \\ &= a^2\sigma_{XX} + b^2\sigma_{YY} + 2ab\,\sigma_{XY} \\ &= \langle w, \Sigma w\rangle \end{aligned}$$

where $w = \begin{bmatrix} a \\ b \end{bmatrix}$. Note that as $\mathrm{Var}(aX+bY) \geq 0$, we have $\langle w, \Sigma w\rangle \geq 0$, $\forall\, w \in \mathbb{R}^2$. Thus Σ is a positive semidefinite matrix. Another important idea is that of the moment generating function of more than one random variable. This is defined as follows

$$m_{X,Y}(t_1, t_2) = E\left[e^{t_1 X + t_2 Y}\right]$$

Note that if X and Y are independent, then

$$\begin{aligned} m_{XY}(t_1, t_2) &= E\left[e^{t_1 X} e^{t_2 Y}\right] \\ &= E\left[e^{t_1 X}\right] E\left[e^{t_2 Y}\right] \\ &\cancel{= m_X(t_1)\, m_X t} \\ &= m_X(t_1)\, m_Y(t_2). \end{aligned}$$

Let us now see how do we ~~mo~~ compute the expectations in this case.

We have

$$E[X] = \frac{\partial m_{X,Y}(t_1,t_2)}{\partial t_1}\bigg|_{(t_1,t_2)=(0,0)}$$

$$E[Y] = \frac{\partial m_{X,Y}(t_1,t_2)}{\partial t_2}\bigg|_{(t_1,t_2)=(0,0)}$$

&

$$E[X,Y] = \frac{\partial^2 m_{X,Y}(t_1,t_2)}{\partial t_1 \partial t_2}\bigg|_{(t_1,t_2)=(0,0)}.$$

I am not providing the details here which I believe ~~are ab~~ can be checked easily by the reader. Further the above expressions can be written down by even thinking in an intutive way. Till now the most important learning of the previous section is ~~foll~~ the following:

X and Y are independent $\Rightarrow$ X and Y are uncorrelated

However the converse is not true. Now let us ask the following question

$$\boxed{\text{Under what situation we can have X and Y are independent} \underset{\text{if and only if}}{\Longleftrightarrow} \text{X and Y are uncorrelated}}$$

The answer is as follows: If X and Y jointly follow the ~~bi-vari~~ bivariate normal distribution then the above assertion is true. So we shall start with the ~~bi-varie~~ bivariate normal density.

The bivariate normal distribution has a joint pdf given by

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}}\, e^{-\frac{1}{2(1-\rho_{XY}^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho_{XY}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}}$$

where $-\infty < x < +\infty$, $-\infty < y < +\infty$, and $\sigma_X > 0$, $\sigma_Y > 0$ $\rho \in [-1,+1]$ are finite ~~finite~~ ~~constan~~ real-valued constants and further μ_X & μ_Y are also constants.

We shall first show that it is a pdf.

- $f_{XY}(x,y) \geq 0, \quad \forall (x,y) \in \mathbb{R}^2$ is clear.
- We will now have to show that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\,dx = 1$$

To begin with let us substitute

$$u = \frac{x-\mu_x}{\sigma_X} \quad \& \quad v = \frac{y-\mu_Y}{\sigma_Y}$$

Hence

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho_{XY}^2}}\, e^{-\frac{1}{2}\cdot\frac{1}{1-\rho_{XY}^2}\left(u^2 - 2\rho_{XY} uv + v^2\right)}\,du\,dv$$

(Check it yourself)

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho_{XY}}}\, e^{-\frac{1}{2(1-\rho_{XY}^2)}\left[(u-\rho_{XY}v)^2 + (1-\rho_{XY}^2)v^2\right]}\,du\,dv \quad \cancel{dx\,dw}$$

- Again let us substitute keeping v fixed.

$$w = \frac{u-\rho_{xy}v}{\sqrt{1-\rho_{xy}^2}}, \quad \therefore dw = \frac{du}{\sqrt{1-\rho^2}}$$

$$\therefore \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}(x,y)\,dy\,dx = \left(\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-w^2/2}\,dw\right)\left(\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\,dv\right) = 1$$

(Think why??)

(Think about the standard normal distribution).

Thus f_{XY} is a p.d.f.

Let us now compute the mgf of the ~~bi-v~~ bivariate normal distribution

$$m_{X,Y}(t_1,t_2) = E\left[e^{t_1X+t_2Y}\right]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1x+t_2y} f_{XY}(x,y)\,dy\,dx$$

Set $u = \dfrac{x-\mu_x}{\sigma_x}$, & $v = \dfrac{y-\mu_y}{\sigma_y}$, hence

$$m_{X,Y}(t_1,t_2) = e^{t_1\mu_X+t_2\mu_Y}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{t_1\sigma_x u+t_2\sigma_y v}}{2\sqrt{1-\rho_{XY}^2}}\, e^{-\frac{1}{2}\frac{1}{(1-\rho_{XY}^2)}\left(u^2-2\rho_{XY}uv+v^2\right)}\,dv\,du$$

$$\therefore m_{X,Y}(t_1,t_2) = e^{t_1\mu_X+t_2\mu_Y}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\sqrt{1-\rho_{XY}^2}}\, e^{-\frac{1}{2(1-\rho_{XY}^2)}\left[u^2-2\rho_{xy}uv+v^2-2(1-\rho^2)t_1\sigma_x u-2(1-\rho^2)t_2\sigma_y v\right]}$$

$$= \frac{e^{t_1\mu_X+t_2\mu_Y}}{2\sqrt{1-\rho_{XY}^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho_{XY}^2)}\left[\left(u-\rho_{xy}v-(1-\rho_{xy}^2)t_1\sigma_x\right)^2+(1-\rho^2)\left(v-\rho_{xy}t_1\sigma_x-t_2\sigma_y\right)^2-(1-\rho_{xy}^2)\left(t_1\sigma_x^2+t_2\sigma_y^2+2\rho_{xy}t_1t_2\sigma_x\sigma_y\right)\right]}\,dv\,du$$

Let

$$w = \frac{u-\rho_{xy}v-(1-\rho_{xy}^2)t_1\sigma_x}{\sqrt{(1-\rho_{xy}^2)}}$$

$$z = v-\rho_{xy}t_1\sigma_x-t_2\sigma_y$$

$$\therefore m_{X,Y}(t_1,t_2) = e^{t_1\mu_x+t_2\mu_Y}\, e^{\left[\frac{1}{2}\left[t_1^2\sigma_X^2+2\rho_{xy}t_1t_2\sigma_x\sigma_y+t_2^2\sigma_y^2\right]\right]} \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-w^2/2-z^2/2}\,dw\,dz$$

$$\therefore m_{X,Y}(t_1,t_2) = e^{t_1\mu_x+t_2\mu_y+\frac{1}{2}\left(t_1^2\sigma_x^2+2\rho_{xy}t_1t_2\sigma_x\sigma_y+t_2^2\sigma_y^2\right)}$$

[Note that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-w^2/2-z^2/2}\,dw\,dz = 1$]

Think why

Set $h = t_1\mu_x+t_2\mu_y+\frac{1}{2}\left(t_1^2\sigma_x^2+2\rho_{xy}t_1t_2\sigma_x\sigma_y+t_2^2\sigma_y^2\right)$

$$\frac{\partial m_{X,Y}(t_1,t_2)}{\partial t_1} = \left(\mu_X + t_1\sigma_X^2 + \rho_{XY} t_2 \sigma_X \sigma_Y\right) e^h$$

$$\frac{\partial m_{X,Y}(t_1,t_2)}{\partial t_2} = \left(\mu_Y + t_2\sigma_Y^2 + \rho_{XY} t_1 \sigma_X \sigma_Y\right) e^h$$

$$\frac{\partial^2 m_{XY}(t_1,t_2)}{\partial t_1 \partial t_2} = \rho_{XY}\sigma_X\sigma_Y e^h + \cancel{(\mu_X + t_2\sigma_Y^2 + \rho_{XY} t_1\sigma_X\sigma_Y}$$
$$\left(\mu_Y + t_2\sigma_Y^2 + \rho_{XY} t_1 \sigma_X\sigma_Y\right)\left(\mu_X + t_1\sigma_X^2 + \rho_{XY} t_2\sigma_X\sigma_Y\right)e^h.$$

$$\therefore \left.\frac{\partial m_{X,Y}(t_1,t_2)}{\partial t_1}\right|_{(t_1,t_2)=(0,0)} = \mu_X = E(X)$$

$$\therefore \left.\frac{\partial m_{X,Y}(t_1,t_2)}{\partial t_2}\right|_{(t_1,t_2)=(0,0)} = \mu_Y = E(Y)$$

$$\left.\frac{\partial^2 m(t_1,t_2)}{\partial t_1 \partial t_2}\right|_{(t_1,t_2)=(0,0)} = \cancel{\mu_X}\, \rho_{XY}\sigma_X\sigma_Y + \mu_X\mu_Y$$

$$\therefore \quad Cov(X,Y) = \sigma_{XY} = E[XY] - E[X]E[Y]$$
$$= \rho_{XY}\sigma_X\sigma_Y + \mu_X\mu_Y - \mu_X\mu_Y$$

$$\therefore \quad Cov(X,Y) = \rho_{XY}\sigma_X\sigma_Y$$

$\therefore$ Correlation coefficient $= \rho_{XY}$.

We leave it to the reader to check that $Var(X) = \sigma_X^2$ & $Var(Y) = \sigma_Y^2$ of this section.

Now we come to the main result of this section.

<u>Theorem 8.5</u>: If X and Y are two random variables which jointly follow the bivariate normal distribution, Then X and Y are independent $\Leftrightarrow$ X and Y are uncorrelated.

Proof: If X and Y are independent, then X and Y are uncorrelated is a known fact.

For the converse, let X and Y be uncorrelated

$$\Rightarrow \operatorname{Cov}(X,Y)=0$$

$$\Rightarrow \rho_{XY}=0$$

Thus

$$f_{X,Y}(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y}e^{-\frac{1}{2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$

$$=\left[\frac{1}{\sqrt{2\pi}\,\sigma_X}e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2}\right]\left[\frac{1}{\sqrt{2\pi}\,\sigma_Y}e^{-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}\right]$$

$$=f_X(x)\,f_Y(y),$$

where $X\sim N(\mu_X,\sigma_X^2)$ & $Y\sim N(\mu_Y,\sigma_Y^2)$.

Of course one can easily check that when $\rho_{X,Y}=0$, the marginal distributions are ~~f_X~~ $f_X(x)$ and $f_Y(y)$. This shows that X and Y are independent random variable □.

—— x ——

Section 8.3: Conditional Expectation

Conditional Expectation is a natural outcome from the fact that we have defined the notion of a conditional density. Let us define it formally

Conditional Expectation of the r.v.'s X and Y

Let (X,Y) be a two-dimensional random-vector. Let $Z = g(X,Y)$ be a function of the two random variables.

Then the conditional expectation of $g(X,Y)$, given $X=x$ is given as

$$E\left[g(X,Y)\,|\,X=x\right] = \int_{-\infty}^{\infty} g(x,y)\, f_{Y|X}(y|x)\, dy$$

This formula holds good if X and Y are jointly continuous

If X and Y are jointly discrete we have

$$E\left[g(X,Y)\,|\,X=x\right] = \sum_j g(x, y_j)\, f_{Y|X}(y_j|x).$$

E.g: Let $f_{X,Y}$ be the joint density of two continuous random variable given as

$$f_{X,Y}(x,y) = \begin{cases} x+y; & \text{if } \cancel{x>0 \,\&\, y>0}\ x\in(0,1)\ \&\ y\in(0,1) \\ 0 & \text{otherwise.} \end{cases}$$

How shall we now proceed to compute $E[Y|X=x]$ (i.e. $g(X,Y)=Y$)

First we compute

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{x+y}{f_X(x)}, \quad x,y\in(0,1).$$

$$f_X(x) = \int_0^1 (x+y)\,dy = x+\tfrac{1}{2} \quad \therefore \quad \boxed{\frac{x+y}{x+\frac{1}{2}} = f_{Y|X}(y,x)}$$

[Here $g(X,Y)=Y$. In fact $E[g(X,Y)\,|\,X=x]$ is a function of X]

Thus

$$E[Y \mid X = x] = \int_0^1 y \, f_{Y|X}(y|x)\, dy$$

$$= \int_0^1 y \, \frac{x+y}{x + 1/2}\, dy$$

$$= \frac{1}{x+1/2}\left[\int_0^1 yx\, dy + \int_0^1 y^2\, dy\right]$$

$$= \frac{1}{x+1/2}\left[\frac{x}{2} + \frac{1}{3}\right]. \qquad \square$$

For simplicity, ~~Note~~ consider $E[g(Y)|x]$, which ~~inge~~ in general a function of x. Let us denote it as

$$\phi(x) = E[g(Y)|x]$$

$\therefore$ $\phi(x)$ can also be written as

$$\phi(X) = E[g(Y)|X]$$

$\therefore$ For a given $\omega \in \Omega$

$$\phi(X(\omega)) = E[g(Y)|X(\omega)]$$

$\therefore$ If $X(\omega) = x$, then we have

$$\phi(x) = E[g(Y)|x]$$

Note that in general $\phi(X) = E[g(Y)|X]$ can be viewed as a random variable.

So our learning here is as follows:

Conditional Expectation $E[g(Y)|X]$ is a random variable.

It might be sometimes relevant to ask, "What is the expectation of $\phi(X)$". The answer may surprise you. We do the calculation for $g(Y) = Y$. Assume $f_X(x) > 0$, $\forall x$

Set

$$E[\phi(X)] = E[E[Y|X]]$$

$$= E[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) f_X(x)\,dx$$

$$= \int_{-\infty}^{\infty} E[Y|x]\, f_X(x)\,dx$$

$$\cancel{= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y f_{Y|X} (\right.}$$

$$= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy \right] f_X(x)\,dx$$

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

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \cdot \frac{f_{X,Y}(x,y)}{f_X(x)} \cdot f_X(x)\,dy\,dx$$

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

$$= E[Y].$$

$$\therefore \boxed{E[E[Y|X]] = E[Y]}$$

(The reader can prove it for the discrete case).

We shall now define the notion of conditional variance of Y when the random variable X takes a particular value say ~~x~~ x.

$$Var(Y|X=x) = E[Y^2|X=x] - \left[E[Y|X=x]\right]^2$$

In way we discussed before the case of the conditional expectation, the conditional variance can also be viewed as a r.v. $Var(Y|X)$, whose value at any $\omega \in \Omega$, is computed as

$$\boxed{Var(Y|X)(\omega) = Var(Y|X = X(\omega))}$$

$X(\omega)$ can of course take the value of x.

We end our discussion with the following interesting result.

<u>Theorem 8.4</u> Let X and Y be random variables. Then

$$Var(Y) = E[Var(Y|X)] + Var[E[Y|X]).$$

<u>Proof:</u>
$$E[Var(Y|X)] = E[E[Y^2|X] - (E[Y|X])^2]$$
$$= E[E[Y^2|X]] - E[(E[Y|X])^2]$$
$$= E[Y^2] - E[E[Y|X]^2]$$

$\therefore$ In fact $E[E[Y^2|X]] = E[Y^2]$, can be proved from the proof style of showing $E[E[Y|X]] = E[Y]$

$$\therefore E[Var(Y|X)] = E[Y^2] - (E[Y])^2 \cancel{E[E[Y} - E[(E[Y|X])^2] + (E[Y])^2$$

$$\therefore E[Var(Y|X)] = Var(Y) - \left[E[(E(Y|X))^2] - (E(E(Y|X)))^2\right]$$
$$= Var(Y) - Var(E[Y|X])$$

$$\therefore Var(Y) = E[Var(Y|X)] + Var(E[Y|X]).$$

—x—x—