Theorem:

Let X and Y are jointly distributed random variables such that the variance of X and Y exist finitely. Let μ_1, μ_2 are the means and σ_1^2, σ_2^2 are variances of X and Y respectively and ρ is the correlation coefficient between X and Y. If $E(Y|X)$ is linear in X, then

(a) $E(Y|X) = \mu_2 + \rho \frac{\sigma_2}{\sigma_1}(X-\mu_1)$

(b) $E(Var(Y|X)) = \sigma_2^2(1-\rho^2)$

Proof:

We shall prove this result for continuous random variables. Similarly it can be proved for discrete random variables as well.

(a) Suppose X and Y are continuous random variables having joint pdf $f(x,y)$ and marginal pdf's of X and Y are $f_1(x)$ and $f_2(y)$ respectively

$\because E(Y|X)$ is linear in X

$\therefore E(Y|X) = a + bx$

$\Rightarrow \int_{-\infty}^{\infty} y\, f(y|x)\, dy = a + bx$

$$\Rightarrow \int_{-\infty}^{\infty} y \frac{f(x,y)}{f_1(x)} dy = a+bx$$

$$\Rightarrow \int_{-\infty}^{\infty} y f(x,y)\, dy = (a+bx) f_1(x) \quad \text{——— } (1)$$

Integrate w.r.to x both sides, we have,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\, dy\, dx = \int_{-\infty}^{\infty} (a+bx) f_1(x)\, dx$$

$$\Rightarrow \int_{-\infty}^{\infty} y \left\{ \int_{-\infty}^{\infty} f(x,y)\, dx \right\} dy = a\int_{-\infty}^{\infty} f_1(x)\, dx + b\int_{-\infty}^{\infty} x f_1(x)\, dx.$$

$$\Rightarrow \int_{-\infty}^{\infty} y f_2(y)\, dy = a.1 + b\,E(x)$$

$$\Rightarrow E(y) = a + b\,E(x)$$

$$\Rightarrow \mu_2 = a + b\mu_1 \quad \text{——— } (2)$$

Now multiply eqn (1) with x, we have

$$\int_{-\infty}^{\infty} xy f(x,y)\, dy = x(a+bx) f_1(x)$$

Now integrate w.r.to x both sides, we have

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)\, dy\, dx = \int_{-\infty}^{\infty} x(a+bx) f_1(x)\, dx.$$

$$\Rightarrow \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)\, dx\, dy = a\int_{-\infty}^{\infty} x f_1(x)\, dx + b\int_{-\infty}^{\infty} x^2 f_1(x)\, dx$$

$$\Rightarrow E(XY) = aE(X) + bE(X^2) \quad \text{——— (3)}$$

Now, $\rho = \dfrac{\operatorname{Cov.}(X,Y)}{\sigma_x \sigma_y} = \dfrac{E(XY) - E(X)E(Y)}{\sigma_1 \sigma_2}$.

$$\Rightarrow E(X \cdot Y) = \rho\sigma_1\sigma_2 + E(X)E(Y)$$

$$\Rightarrow E(XY) = \rho\sigma_1\sigma_2 + \mu_1\mu_2.$$

using it in eqn (3), we have

$$\rho\sigma_1\sigma_2 + \mu_1\mu_2 = aE(X) + bE(X^2)$$

$$\Rightarrow \rho\sigma_1\sigma_2 + \mu_1\mu_2 = a\mu_1 + b(\sigma_1^2 + \mu_1^2) \quad \text{——— (4)}$$

$\left[\because \sigma_1^2 = E(X^2) - (E(X))^2 \Rightarrow \sigma_1^2 = E(X^2) - \mu_1^2 \Rightarrow E(X^2) = \sigma_1^2 + \mu_1^2\right]$

By solving eqn (2) and (4), we have

$$a = \mu_2 - \rho\frac{\sigma_2}{\sigma_1} \quad \text{and} \quad b = \rho\frac{\sigma_2}{\sigma_1}$$

$$\therefore E(Y|X) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X - \mu_1)$$

(b) Now, $\operatorname{Var}(Y|X) = E\left((Y) - E(Y|X)\right)^2.$

$$\Rightarrow \operatorname{Var}(Y|X) = E\left(Y - E(Y|X)\right)^2.$$

Now, conditional variance for Y given $X = x$ is

$$Var(Y|X=x) = E(Y - E(Y|X=x))^2$$

$$\Rightarrow V(Y|X) = E\left(Y - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(\cancel{X} - \mu_1)\right)^2$$

$$\Rightarrow V(Y|X) = \int_{-\infty}^{\infty}\left(y - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2 f(y|x)\,dy$$

$$\Rightarrow V(Y|X) = \int_{-\infty}^{\infty}\left(y - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2 \frac{f(x,y)}{f_1(x)}\,dy$$

Now multiply both sides by $f_1(x)$, we have.

$$V(Y|X)f_1(x) = \int_{-\infty}^{\infty}\left(y - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2 f(x,y)\,dy$$

Now, integrate both sides w.r.t. x, we have.

$$\int_{-\infty}^{\infty} V(Y|X)f_1(x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(y - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)^2 f(x,y)\,dy\,dx$$

$$\Rightarrow E(Var(Y|X)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{(y-\mu_2)^2 + \rho^2\frac{\sigma_2^2}{\sigma_1^2}(x-\mu_1)^2 - 2\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)(y-\mu_2)\right\}f(x,y)\,dx\,dy$$

$$\Rightarrow E(Var(Y|X)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-\mu_2)^2 f(x,y)\,dx\,dy + \rho^2\frac{\sigma_2^2}{\sigma_1^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)^2 f(x,y)\,dx\,dy$$
$$- 2\rho\frac{\sigma_2}{\sigma_1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)\,dx\,dy.$$

$$\Rightarrow E(Var(Y|X)) = E(Y-\mu_2)^2 + \rho^2 \frac{\sigma_2^2}{\sigma_1^2} \cdot E(X-\mu_1)^2$$
$$- 2\rho \frac{\sigma_2}{\sigma_1} E(X-\mu_1)(Y-\mu_2)$$

$$\Rightarrow E(Var(Y|X)) = \sigma_2^2 + \rho^2 \cdot \frac{\sigma_2^2}{\sigma_1^2} \cdot \sigma_1^2$$
$$- 2\rho \frac{\sigma_2}{\sigma_1} \cdot Cov.(X,Y)$$

$$\Rightarrow E(Var(Y|X)) = \sigma_2^2 + \rho^2 \sigma_2^2 - 2\rho \frac{\sigma_2}{\sigma_1} \cdot \rho \sigma_1 \sigma_2$$

$$\Rightarrow E(Var(Y|X)) = \sigma_2^2 + \rho^2 \sigma_2^2 - 2\rho^2 \sigma_2^2 \qquad \left[\because \rho = \frac{Cov.(X,Y)}{\sigma_1 \sigma_2} \Rightarrow Cov.(X,Y) = \rho \sigma_1 \sigma_2 \right]$$

$$\Rightarrow E(Var(Y|X)) = \sigma_2^2 - \rho^2 \sigma_2^2$$

$$\Rightarrow E(Var(Y|X)) = \sigma_2^2 (1-\rho^2)$$

Result:

Let X and Y are jointly distributed random variables such that the variance of X and Y exists finitely. Let μ_1, μ_2 are the means and σ_1^2, σ_2^2 are variances of X and Y respectively. Suppose ρ is the correlation coefficient between X and Y.

If $E(X|Y)$ is linear in x, then

(a) $E(X|Y) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (Y - \mu_2)$

(b) $E(Var(X|Y)) = \sigma_1^2 (1-\rho^2)$

Try to prove it yourself.

Scanned with CamScanner