4.3 Covariance and Correlation

Definition Let X and Y be two random variables with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2, repectively.

Consider the mathematical expectation

$$E[(X - \mu_X)(Y - \mu_Y)] = E[XY - \mu_Y X - \mu_X Y + \mu_X \mu_Y]$$

$$= E[XY] - \mu_X \mu_Y$$

This expectation is called the **covariance** of X and Y.

If σ_X and σ_Y are positive,

$$\rho = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

is called the **correlation coefficient** of X and Y.

ρ measures the linear association of X and Y and has the following properties.

- $-1 \leq \rho \leq 1$.
- $\rho = 1$ if and only if $Y = a + bX$ with probability equal to 1 for some $b > 0$
- $\rho = -1$ if and only if $Y = a + bX$ with probability equal to 1, for some $b < 0$
- ρ remains the same under location-scale transforms of X and Y.
- ρ measures the extent of linear association.
- ρ is 0 if X and Y are independent.

Let $f(x, y)$ denote the joint pdf of X and Y. If $E[e^{t_1 X + t_2 Y}]$ exists in a neighborhood of $(0, 0)$, it is denoted by $M(t_1, t_2)$ and is called the moment generating function of X and Y.

In the case of continuous random variables

$$\frac{\partial^{k+m} M(t_1, t_2)}{\partial t_1^k \partial t_2^m} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^m e^{t_1 x + t_2 y} f(x, y) dx dy$$

so that we can compute various moments by evaluating these partial derivatives at $(t_1, t_2) = (0, 0)$.

$$\mu_X = \frac{\partial M(0, 0)}{\partial t_1}$$

$$\sigma_X^2 = \frac{\partial^2 M(0, 0)}{\partial t_1^2} - \mu_X^2$$

$$Cov(X, Y) = \frac{\partial^2 M(0, 0)}{\partial t_1 \partial t_2} - \mu_X \mu_Y$$

4.4 Condtional Expectation

Recall the example:

Let X_1 and X_2 have joint pdf $f(x_1, x_2) = 6x_2$ for $0 < x_2 < x_1 < 1$.

The marginal pdf of X_1 is

$$f_1(x_1) = \int_0^{x_1} 6x_2 dx_2 = 3x_1^2$$

for $0 < x_1 < 1$. The conditional pdf of X_2 given x_1 is

$$f_{2|1}(x_2|x_1) = \frac{2x_2}{x_1^2}$$

The conditional mean or **conditional expectation** of X_2 given $X_1 = x_1$ is

$$E[X_2|x_1] = \int_0^{x_1} x_2(2x_2/x_1^2)dx_2 = 2x_1/3$$

We can view $Y = 2X_1/3$ as a random variable and find its distribution function

$$F(y) = P(Y \leq y) = P(X_1 \leq \frac{3y}{2})$$

$$= \int_0^{3y/2} 3x_1^2 dx_1 = \frac{27y^3}{8}$$

for $0 \leq y < 2/3$.

By taking its derivative we can find the pdf of Y.

$$f(y) = F'(y) = \frac{81y^2}{8}$$

Thus,

$$E[Y] = \int_0^{2/3} y(81y^2/8)dy = \frac{1}{2}$$

and

$$Var(Y) = \int_0^{2/3} y^2(81y^2/8)dy - \frac{1}{4} = \frac{1}{60}$$

The marginal pdf of X_2 is

$$f_2(x_2) = \int_{x_2}^{1} 6x_2 dx_1 = 6x_2(1 - x_2)$$

Knowing this, we can easily find that $E[X_2] = 1/2$ and $Var(X_2) = 1/20$.

In general, if $Y = E[X_2|X_1]$ then

$$E[Y] = E[E[X_2|X_1]] = E[X_2]$$

and

$$Var(Y) = Var(E[X_2|X_1]) \leq Var(X_2)$$

To be specific

$$Var(X_2) = E[(X_2 - E[X_2|X_1])^2] + E[(E[X_2|X_1] - \mu_2)^2]$$