3. Random Variables and Probability Distributions - Correlation rho

Lai Hui Shan
m5281022

1. Why We Use the Correlation Coefficient ρ

The covariance between two random variables X and Y is defined as:

$$\operatorname{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)].$$

Covariance measures the direction and strength of the linear relationship between X and Y. However, it has a key limitation: it depends on the scales (units) of X and Y. For example, if we change the units of X or Y, the covariance changes as well, which makes it difficult to interpret across different datasets.

To create a standardized measure of the relationship between X and Y that is independent of their scales, we use the correlation coefficient ρ, defined as:

$$\rho(X, Y) = \frac{\operatorname{Cov}(X, Y)}{\sigma_X \sigma_Y},$$

where σ_X and σ_Y are the standard deviations of X and Y, respectively.

Why Divide by the Standard Deviations?

Dividing by σ_X and σ_Y normalizes the covariance, transforming it into a unit-free measure that reflects only the strength and direction of the linear relationship, without being affected by the scales of X and Y. This normalized value, called the **correlation coefficient** ρ, provides a consistent basis for comparing relationships between different pairs of variables and enables us to interpret the strength of their relationship on a standard scale.

2. Showing that $\rho \in [-1, 1]$

To show that the correlation coefficient $\rho(X, Y)$ takes values in the interval $[-1, 1]$, we use the **Cauchy–Schwarz inequality**. For random variables X and Y in $\mathcal{L}^2$ (i.e., with finite variance), the Cauchy–Schwarz inequality states:

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2) \cdot \mathbb{E}(Y^2)}.$$

Applying this to $X - \mu_X$ and $Y - \mu_Y$, we get:

$$|\operatorname{Cov}(X, Y)| = |\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]| \leq \sqrt{\mathbb{E}[(X - \mu_X)^2] \cdot \mathbb{E}[(Y - \mu_Y)^2]} = \sigma_X \sigma_Y.$$

Dividing both sides by $\sigma_X \sigma_Y$ (assuming $\sigma_X, \sigma_Y > 0$), we obtain:

$$\left|\frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}\right| \leq 1.$$

Thus,

$$-1 \leq \rho(X,Y) \leq 1.$$

Interpretation of ρ Values

- $\rho = 1$: X and Y have a perfect positive linear relationship.
- $\rho = -1$: X and Y have a perfect negative linear relationship.
- $\rho = 0$: X and Y have no linear relationship (they may still be dependent in a nonlinear way).

This bounded range of ρ makes it a valuable and interpretable measure for the linear relationship between X and Y.