Correlation between Two Real-Valued Random Variables

Prof. Joon Ho Cho

Department of Electrical Engineering, POSTECH

- Review. The received signal contains a noise $N(t)$ often modeled by an **additive white Gaussian noise (AWGN)**. To understand it, we started to review a pre-requisite of this course, *Probability, Random Variables, and Random Processes*.
  - CDF and PDF of a random variable
  - joint CDF and joint PDF of multiple random variables
  - how to handle multiple complex-valued random variables by using multiple real-valued random variables
  - conditional probability
  - statistical independence among multiple events
  - statistical independence among multiple random variables
  - extension of the above notion of **statistical independence** to those among
    - multiple random vectors,
    - multiple random matrices,
    - multiple DT random processes,
    - multiple CT random processes, etc.?

- Preview. We will examine the notion of **correlation between two real-valued random variables**. This is to later handle the correlation functions of random processes and their PSDs.
  - expected value of a random variable
  - correlation, covariance, and correlation coefficient between two real-valued random variables
  - uncorrelatedness and orthogonality between two real-valued random variables

- Q. Given a real-valued random variable X, what is the expectation or the expected value of it?
- A. It is denoted by $\mathbb{E}\{X\}$ and defined as

$$\mathbb{E}\{X\} \triangleq \int_{-\infty}^{\infty} x dF_X(x) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

  - Specialize. If X is discrete, then

$$\mathbb{E}\{X\} \triangleq \sum_i x_i \Pr(X = x_i) = \sum_{x \in S} x p_X(x),$$

where S is the sample space and $p_X(x)$ is called the probability mass function of X. Thus, the expected value is the **sum of the value times the probability**.

- Q. What is the physical meaning of the expected value?
- A. Suppose that X_i's are i.i.d. (independent and identically distributed) random variables with the same CDF as X. Then, by the law of large numbers (LLN),
$$\mathbb{E}\{X\} \approx \frac{\sum_{i=1}^{N} X_i}{N}$$
for sufficiently large N. Thus, the expected value is approximately equal to the arithmetic average of the i.i.d. random variables X_i's.

- Properties of expectation
  - Linearity. $\mathbb{E}\{a_1 X_1 + a_2 X_2\} = a_1 \mathbb{E}\{X_1\} + a_2 \mathbb{E}\{X_2\}$
  - An indicator function relates the expectation and the probability. $\mathbb{E}\{1_A(X)\} = \Pr(X \in A)$.

- Q. What is the nth-order moment and the nth-order central moment of X?
- A.
  - The nth order moment is defined by $\mathbb{E}\{X^n\}$.
    - In particular, the first-order moment is called the mean of X.
    - The mean of X is often denoted by μ or μ_X.
    - If the PDF has even symmetry, then $\mathbb{E}\{X^1\} = \mathbb{E}\{X^3\} = \mathbb{E}\{X^5\} = \ldots = 0$.
  - The nth order central moment is defined by $\mathbb{E}\{(X - \mu_X)^n\}$.
    - In particular, the second-order central moment is called the variance of X.
    - The standard deviation of X is often denoted by σ or σ_X.
    - The variance of X is often denoted by σ^2 or σ_X^2.
    - $\sigma_X^2 = \mathbb{E}\{X^2\} - \mu_X^2$.

- Q. Given two real-valued random variables X and Y, what is the **correlation** between X and Y?
- A. It is denoted by R_{XY} and defined by

$$R_{XY} \triangleq \mathbb{E}\{XY\}$$

  - $R_{XX} = \mathbb{E}\{X^2\}$

- Q. Given two real-valued random variables X and Y, what is the **covariance** between X and Y?
- A. It is denoted by C_{XY} and defined by

$$C_{XY} \triangleq \mathbb{E}\{(X - \mu_X)(Y - \mu_Y)\}$$

  - $C_{XX} = \mathbb{E}\{(X - \mu_X)^2\}$
  - $C_{XY} = R_{XY} - \mu_X \mu_Y \Longleftrightarrow R_{XY} = C_{XY} + \mu_X \mu_Y$

- Q. Given two real-valued random variables X and Y, what is the **correlation coefficient** between X and Y?
- A. It is denoted by ρ_{XY} and defined by

$$\rho_{XY} \triangleq \mathbb{E}\left\{\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)\right\}.$$

  - $\rho_{XY} = R_{\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)}$, i.e., it is the correlation between the normalized versions of X and Y.
  - $\rho_{XY} = C_{\left(\frac{X}{\sigma_X}\right)\left(\frac{Y}{\sigma_Y}\right)} = \frac{C_{XY}}{\sqrt{C_{XX}C_{YY}}}$, i.e., it is the covariance between the normalized versions of X and Y.

- Q. Given two real-valued random variables X and Y, show that

$$-1 \leq \rho_{XY} \leq 1.$$

Moreover, show that

  - $\rho_{XY} = 1 \iff Y = aX + b$ for some $a > 0$ and b.
  - $\rho_{XY} = -1 \iff Y = aX + b$ for some $a < 0$ and b.

- A. Let $\hat{X}$ and $\hat{Y}$ be the normalized versions $(X - \mu_X)/\sigma_X$ of X and $(Y - \mu_Y)/\sigma_Y$ of Y, respectively. Consider the inequality

$$\mathbb{E}\{(c\hat{X} - \hat{Y})^2\} \geq 0, \forall c \in \mathbb{R}.$$

Then, the discriminant $D/4$ of the LHS must be less than or equal to 0, i.e.,

$$c^2 R_{\hat{X}\hat{X}} - 2cR_{\hat{X}\hat{Y}} + R_{\hat{Y}\hat{Y}} \leq 0 \iff ... \iff -1 \leq \rho_{XY} \leq 1,$$

where the equality holds iff $\mathbb{E}\{(c\hat{X} - \hat{Y})^2\} = 0$, i.e., $Y = aX + b$ for some a and b.

- Q. What does that mean by "Two real-valued random variables X and Y are strongly correlated."?
- A. Their correlation coefficient is almost ± 1.
  - When $\rho_{XY} \approx 1$, it is highly probable that $\hat{X}$ is large/small iff $\hat{Y}$ is large/small.
  - Example. $Y = X + 1$
  - When $\rho_{XY} \approx -1$, it is highly probable that $\hat{X}$ is large/small iff $\hat{Y}$ is small/large.
  - Example. $Y = -2X$

- Q. What does that mean by "Two real-valued random variables X and Y are uncorrelated."?
- A. Their correlation coefficient is zero, i.e., $\rho_{XY} = 0$.
  - Uncorrelated iff $C_{XY} = 0$.
  - If at least one has mean zero, then uncorrelated iff $R_{XY} = 0$, because $\mu_X \mu_Y = 0 \Longleftrightarrow C_{XY} = R_{XY}$.

- Q. What does that mean by "Two real-valued random variables X and Y are orthogonal."?
- A. Their correlation is zero.
  - Orthogonal iff $R_{XY} = 0$.
  - If at least one has mean zero, then orthogonality iff uncorrelatedness, because $\mu_X \mu_Y = 0 \Longleftrightarrow C_{XY} = R_{XY}$.

- Q. Show that if two real-valued random variables X and Y are independent then they are uncorrelated.
- A.

$$\rho_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)f_{X,Y}(x,y)dxdy = ... = 0$$

- In general, uncorrelatedness does not imply independence. Thus, they are not equivalent. An important exception is the Gaussian case. (We'll see Later once we define a joint Gaussian random vector.)