Systems and Signals 344

2026

Lecture 16

Independent Random Variables & Expected Value of a Function of Two Random Variables & Covariance, Correlation and Independence

Sections 5.6-5.8

Independent Random Variables

Random variables X and Y are independent if and only if $F_{X,Y}(x, y) = F_X(x)F_Y(y)$ for all real x and y.

For discrete random variables this is equivalent to $P_{X,Y}(x, y) = P_X(x)P_Y(y)$.

When the relevant densities exist, it is equivalent to $f_{X,Y}(x, y) = f_X(x)f_Y(y)$.

This is derived from $P(A \cap B) = P(A)P(B)$

Examples

For *Continuous* random variables. Independent if:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

Expected Value of a Function of Two Random Variables

There are many situations where we observe two random variables and use their values to compute a new random variable.

$$W = g(X, Y)$$

W is referred to as a **derived random variable**. The probability model for W is embodied by $P_W(w)$ or $f_W(w)$ which we will discuss later.

However, if we want to calculate $\mathbf{E}[W]$ we can actually do so directly from the joint probability models $P_{X,Y}(x, y)$ or $f_{X,Y}(x, y)$ without the need to derive $P_W(w)$ or $f_W(w)$.

Expected Value of a Function of Two RVs

For random variables X and Y, the expected value of $W = g(X, Y)$ is

Discrete

$$\mathrm{E}[W] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y) P_{X,Y}(x, y)$$

Continuous

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

Expected Value of a Function of Two RVs

From the previous slide we can also easily calculate the expected value of a linear combination of *several* functions.

$$E[a_1 g_1(X,Y) + \ldots + a_n g_n(X,Y)] =$$
$$a_1 \mathrm{E}[g_1(X,Y)] + \ldots + a_n \mathrm{E}[g_n(X,Y)]$$

Example

Expected Value of the sum of two RVs

- The expected value and the variance of the *sum of two random variables* are as follows:

$$\mathrm{E}[X+Y]=\mathrm{E}[X]+\mathrm{E}[Y]$$

- For the expected value, we only need the marginal probability models, and not the complete joint probability models.

$$\mathrm{Var}[X+Y]=\mathrm{Var}[X]+\mathrm{Var}[Y]+2\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]$$

- For the variance we do need the joint probability models.
- The last term is so common it has its own name, the **covariance**.

Covariance, Correlation and Independence

The covariance $\mathbf{Cov}[X,Y]$ *and correlation* $\boldsymbol{r}_{X,Y}$ *require the relevant moments to exist; the correlation coefficient* $\boldsymbol{\rho}_{X,Y}$ *additionally requires positive variances. Under these conditions, independent random variables satisfy* $\mathbf{Cov}[X,Y] = \rho_{X,Y} = \mathbf{0}$*.*

In previous years, different notation was used:

$$C_{XY} = \mathrm{Cov}[X,Y] \quad \text{and} \quad R_{XY} = r_{X,Y}$$

Covariance

The **covariance** of two random variables X and Y is

$$\text{Cov}[X, Y] = \mathbf{E}[(X - \mu_X)(Y - \mu_Y)]$$

Like $\mathbf{Var}[X]$ is a single number that describes the spread of samples around the expected value $\mathbf{E}[X]$,
$\text{Cov}[X, Y]$ is a single number that describes how the pair of random variables X and Y vary together.

Covariance

The **covariance** of two random variables X and Y is

$$\text{Cov}[X, Y] = \mathbf{E}[(X - \mu_X)(Y - \mu_Y)]$$

Like $\mathbf{Var}[X]$ is a single number that describes the spread of samples around the expected value $\mathbf{E}[X]$,
$\text{Cov}[X, Y]$ is a single number that describes how the pair of random variables X and Y vary together.

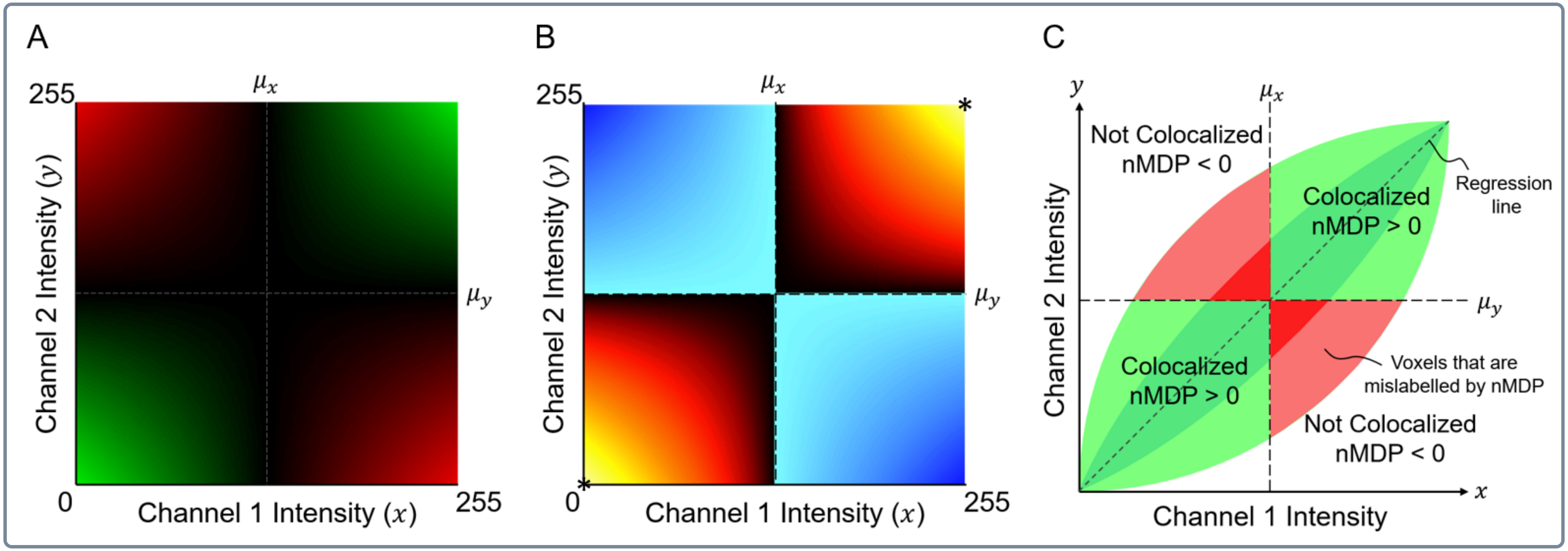
Note that sometimes the notation σ_{XY} is used for covariance.

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

- If $\mathbf{Cov}[X, Y] > 0$
  - Then the typical values of $(X - \mu_X)(Y - \mu_Y)$ are positive.
  - Typically $(X - \mu_X)$ and $(Y - \mu_Y)$ have the *same sign*.
    - i.e. $X < \mu_X$ and $Y < \mu_Y$ or $X > \mu_X$ and $Y > \mu_Y$
  - We expect that X and Y will go up or down together.
- If $\mathbf{Cov}[X, Y] < 0$
  - Typically $(X - \mu_X)$ and $(Y - \mu_Y)$ have *opposite signs*.
  - We expect that if X goes up then Y will go down (vice versa).
- If $\mathbf{Cov}[X, Y] \approx 0$
  - The sign of $(X - \mu_X)$ does not provide much information about the sign of $(Y - \mu_Y)$.

Covariance

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



For more click here.

Examples

Suppose we perform an experiment in which we measure X and Y in *centimetres* (for example, the heights of two sisters). Now we measure $\hat{X}$ and $\hat{Y}$, which are the same measurements in *metres*.

How does the covariance compare?

Correlation Coefficient

The **correlation coefficient** of two random variables X and Y is

$$\rho_{X,Y} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}} = \frac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y}$$

This is a parameter that indicates the relationship of two random variables regardless of the measurement units (since it is normalised).

Correlation Coefficient

$$\rho_{X,Y} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}} = \frac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y}$$

Therefore:

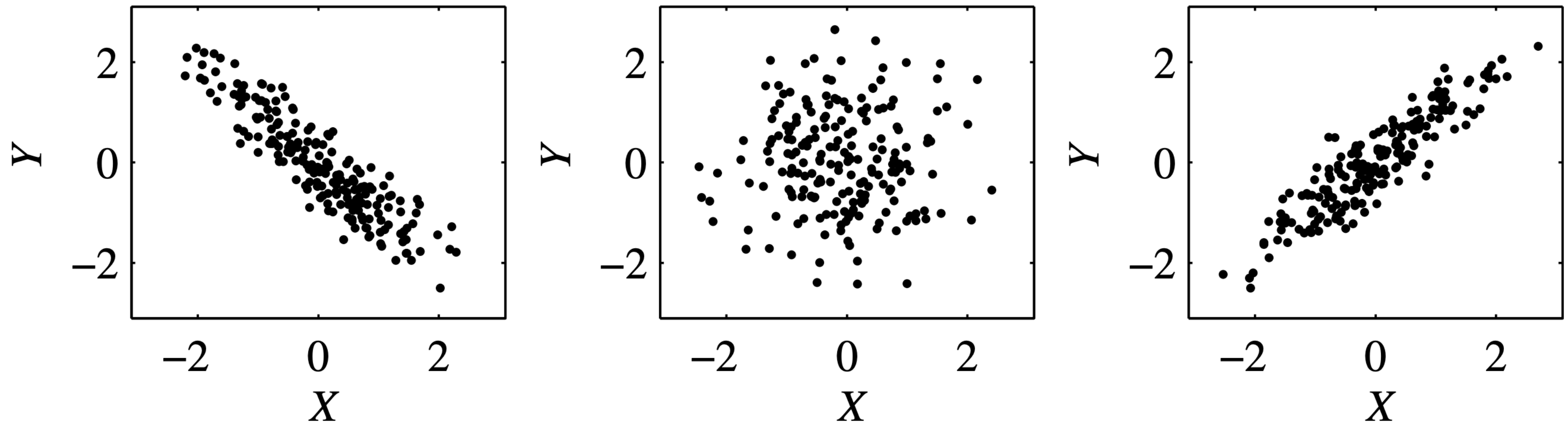
$$-1 \leq \rho_{X,Y} \leq 1$$

Also, for $\hat{X} = aX + b$ and $\hat{Y} = cY + d$, then

$$\rho_{\hat{X},\hat{Y}} = \mathrm{sign}(ac)\rho_{X,Y} \text{ for } a, c \neq 0, \quad \text{and}$$

$$\mathrm{Cov}[\hat{X}, \hat{Y}] = ac\mathrm{Cov}[X, Y]$$

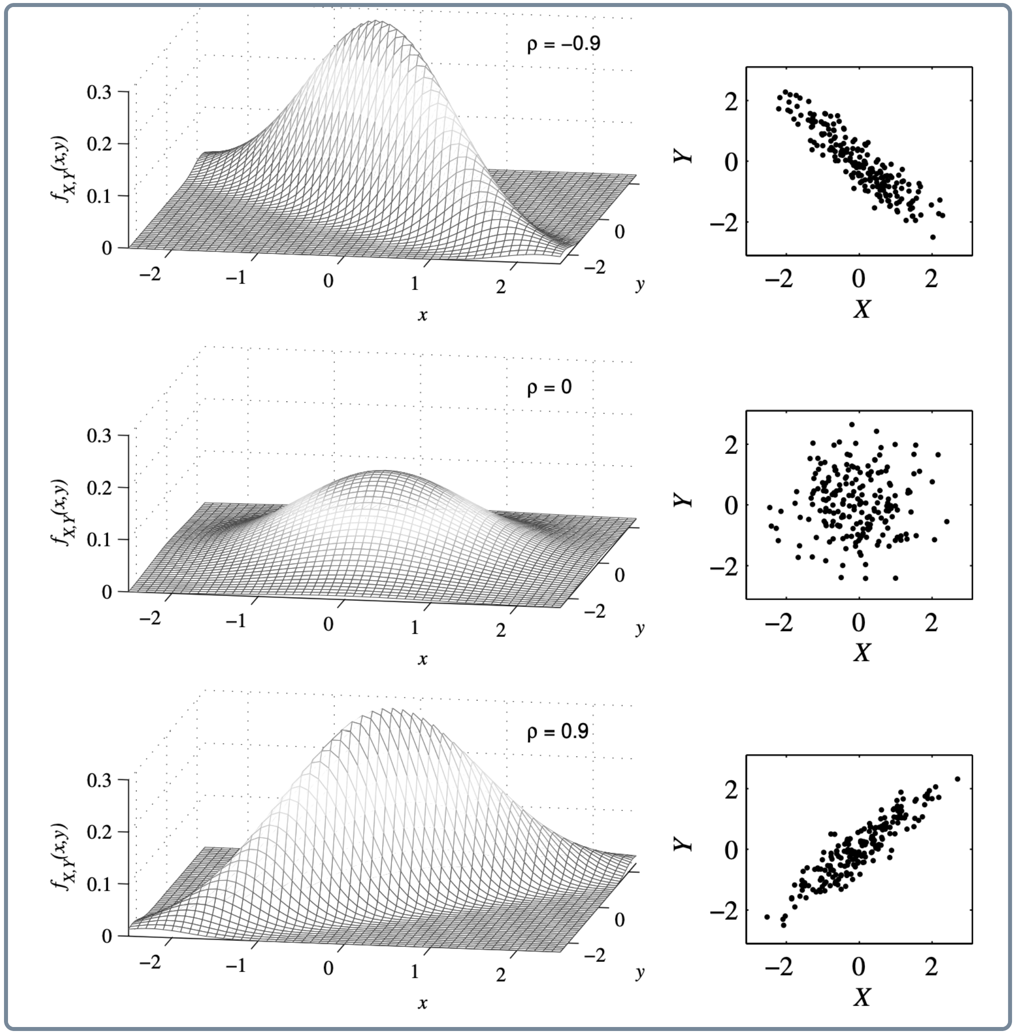
Correlation Coefficient



(a) $\rho_{X,Y} = -0.9$ **(b)** $\rho_{X,Y} = 0$ **(c)** $\rho_{X,Y} = 0.9$

Figure 5.5 Each graph has 200 samples, each marked by a dot, of the random variable pair (X, Y) such that $\mathrm{E}[X] = \mathrm{E}[Y] = 0$, $\mathrm{Var}[X] = \mathrm{Var}[Y] = 1$.

Correlation Coefficient



Correlation Coefficient

- When $\rho_{X,Y} > 0$, we say X and Y are *positively correlated*
- When $\rho_{X,Y} < 0$, we say X and Y are *negatively correlated*
- When $|\rho_{X,Y}| \approx 1$, we say X and Y are *highly correlated*
- When $\rho_{X,Y} = 0$, we say X and Y are *uncorrelated*

Correlation Coefficient

If X has positive variance and $Y = aX + b$, then

$$\rho_{X,Y} = \mathrm{sign}(a), \qquad a \neq 0.$$

When $a = 0$, Y is constant and the correlation coefficient is undefined.

Correlation

Another close relative of the *covariance* is **correlation**.

The **correlation** of X and Y is $r_{X,Y} = \mathbf{E}[XY]$.

It is a raw second-order moment and, unlike the correlation coefficient, depends on the means and scales of the variables. It does not by itself establish dependence, but can be useful as in *indication of dependence*.

Correlation vs Correlation Coefficient

- **Correlation coefficient ($\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$)**:

  Dimensionless measure of *linear association* between X and Y, ranging from (-1) to (+1). It is invariant under positive scale changes and changes sign if exactly one scale factor is negative.

- **Covariance ($\mathrm{Cov}(X, Y)$)**:
  Measures joint variability of X and Y. Keeps units of $X \times Y$, so not directly comparable across contexts.

- **Correlation ($E[XY]$)**:
  The raw cross-moment (expected product) of X and Y. Depends on scale and units, not bounded, and not by itself a measure of linear relationship.

Covariance, Correlation

Some useful identities

- $\mathrm{Cov}[X,Y] = r_{X,Y} - \mu_X \mu_Y$
- $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}[X,Y]$
- If $X = Y$, then $\mathrm{Cov}[X,Y] = \mathrm{Var}[X] = \mathrm{Var}[Y]$ and $r_{X,Y} = \mathrm{E}[X^2] = \mathrm{E}[Y^2]$

Covariance, Correlation

Random variables X and Y are **orthogonal** if $r_{X,Y} = 0$

Random variables X and Y are **uncorrelated** if $\mathrm{Cov}[X, Y] = 0$ and
$r_{X,Y} = \mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$

Note the somewhat confusing terminology:
Orthogonal means zero correlation ($r_{X,Y} = 0$), and
Uncorrelated means zero covariance.

Example 5.17

For the integrated circuits tests in Example 5.3, we found in Example 5.5 that the probability model for X and Y is given by the following matrix.

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$	$P_X(x)$
$x=0$	0.01	0	0	0.01
$x=1$	0.09	0.09	0	0.18
$x=2$	0	0	0.81	0.81
$P_Y(y)$	0.10	0.09	0.81	

Find $r_{X,Y}$ and $\mathrm{Cov}[X,Y]$.

Independence

- When X and Y are *independent* then they are also *uncorrelated*.
- However, uncorrelatedness does not necessarily imply statistical independence.
- Furthermore, for *independent* random variables X and Y:
    - $\mathrm{E}[g(X)h(Y)] = \mathrm{E}[g(X)]\mathrm{E}[h(Y)]$
    - $r_{X,Y} = \mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$
    - $\mathrm{Cov}[X,Y] = \rho_{X,Y} = 0$
    - $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$

Example 5.18

For the noisy observation $Y = X + Z$ of Example 5.1, find the covariances $\mathrm{Cov}[X, Z]$ and $\mathrm{Cov}[X, Y]$ and the correlation coefficients $\rho_{X,Z}$ and $\rho_{X,Y}$.