V-0. Review of Probability

Random Variable

- **Definition**

  Numerical characterization of outcome of a random event

- **Examples**

  1) Number on a rolled die or dice
  2) Temperature at specified time of day
  3) Stock Market at close
  4) Height of wheel going over a rocky road

Random Variable

- **Non-examples**
  1) ‘Heads’ or ‘Tails’ on coin
  2) Red or Black ball from urn

But we can make these into RV’s

- **<u>Basic Idea</u>** – don’t know how to completely determine what value will occur
  – Can only specify probabilities of RV values occurring.

Two types of Random Variables

Random Variable

Discrete RV

- Die
- Stocks

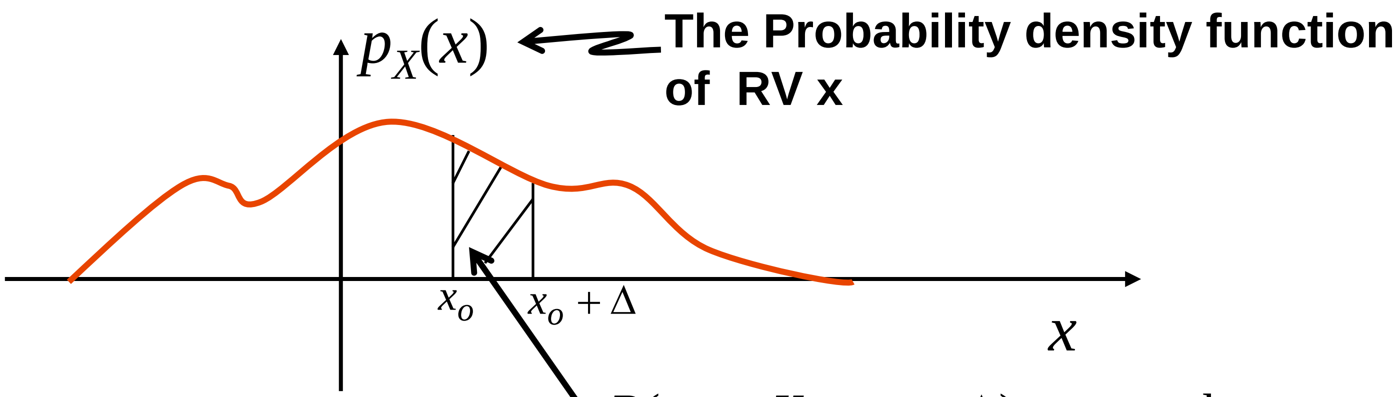
Continuous RV

- Temperature
- Wheel height

Given CRV X, What is the probability that $X = x_0$?

- Oddity : $P(X = x_0) = 0$
  Otherwise the Prob. “Sums” to infinity

- Need to think of <u>Prob. *Density* Function</u> (PDF)



$$P(x_0 < X < x_0 + \Delta) = \text{ area shown}$$

$$= \int p_X(x)dx$$

Most Commonly Used PDF: Gaussian PDF

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2 / 2\sigma^2}$$

A RV with this pdf is called a Gaussian RV

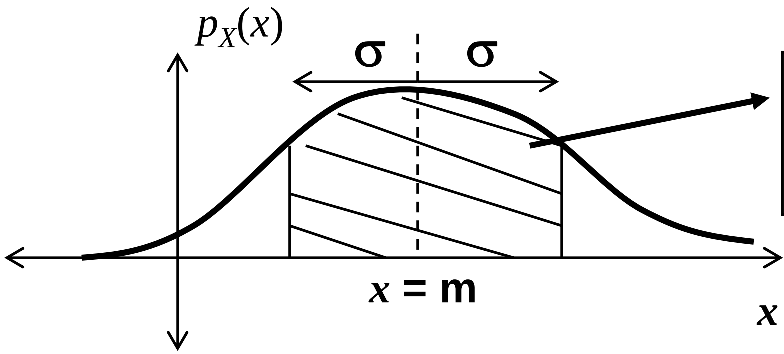
m & σ are parameters describing one of the many Gaussian pdf, where

m = mean of RV x

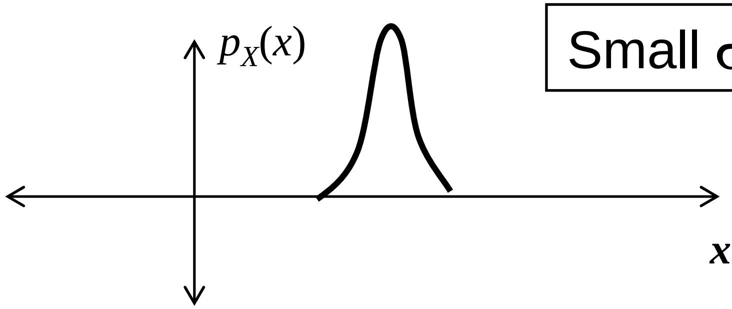
σ = std. Deviation of RV x (Note: σ > 0)

σ^2 = Variance of RV x

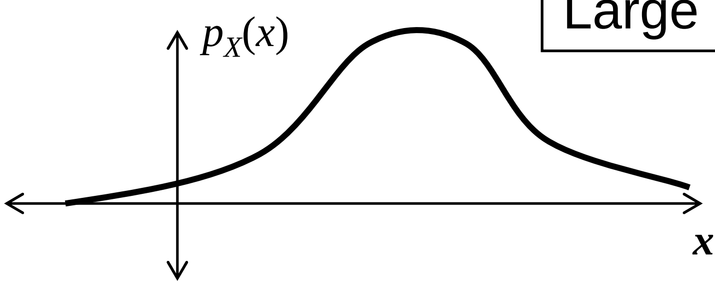
Three views of Guassian PDF’s



Area= 0.683 (i.e. 68.3% of area is within 1 σ of mean)



Small σ → Small variability/uncertainty



Large σ → Large variability/uncertainty

Why Is Gaussian Used?

- **<u>Central Limit theorem (CLT)</u>**

  The sum of N independent RV's has a pdf that tends to be Gaussian as N $\rightarrow \infty$

- **<u>So What! Here is what</u> : Electronic systems generate internal noise due to random motion of atoms in electronic components. The noise is the result of summing the random effects of lots of atoms.**

$\Rightarrow$ **CLT applies** $\Rightarrow$ **Guassian Noise**

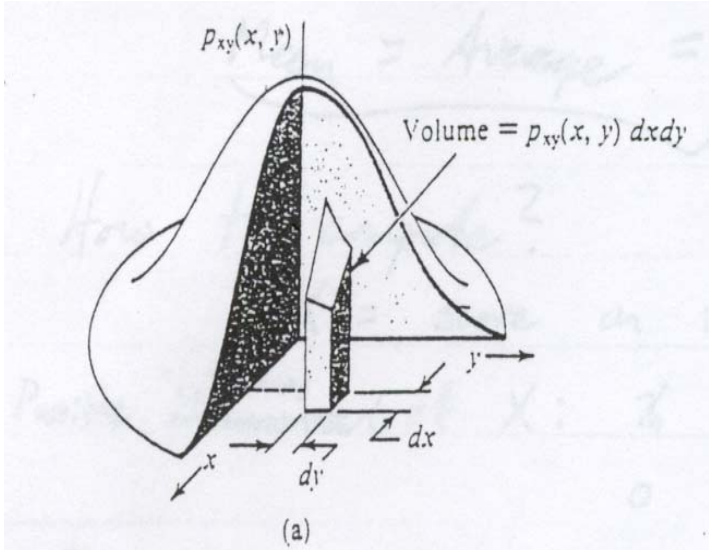
- **<u>Generally</u>: take the noise to be Zero Mean**

$$p_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{x^2 / 2\sigma^2}$$

Joint PDF of X and Y: $p_{XY}(x, y)$

Describes probabilities of joint events concerning X and Y. For example, the probability that X lies in interval [a,b] and Y lies in interval [a,b] is given by:

$$\Pr\{(a < X < b) \text{ and } (c < Y < d)\} = \int_a^b \int_c^d p_{XY}(x, y)\,dx\,dy$$



This graph shows the Joint PDF

Conditional PDF

When you have two RVs it is often necessary to ask questions like: What is the PDF of Y if X is constrained to take on a specific value.

In other words: What is the PDF of Y conditioned on the fact X is constrained to take on a specific value.

As an example consider the husband/wife salaries above: What is the PDF of the husband salary X conditioned on the wife salary is $100K?

So you first find all wives who make EXACTLY $100K and look at how that set of husband salaries are distributed.

Clearly the result depends on the joint PDF since that captures all the probabilistic details of how X and Y interact – and clearly it should only depend on the slice of the joint PDF at the value of Y=$100K.

Now… we have to adjust this to account for the fact that the joint PDF (even its slice) reflects how likely it is that X=$100K will occur (e.g., if X=100000 is unlikely then $p_{XY}(100000,y)$ will be small); so… if we divide by $p_X(100000)$ we adjust for this.

Conditional PDF (cont.)

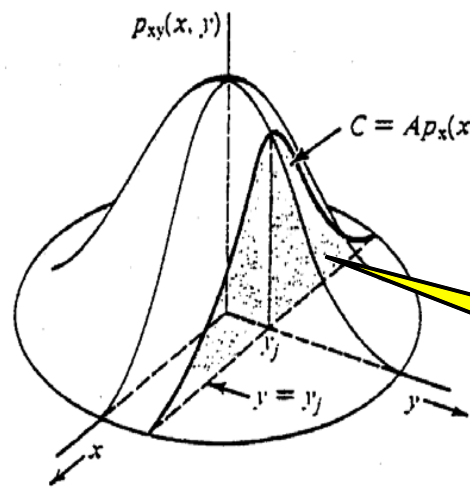
Thus, the conditional PDFs are defined as ("slice and normalize"):

$$p_{Y|X}(y|x) = \begin{cases} \dfrac{p_{XY}(x,y)}{p_X(x)}, & p_X(x) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$p_{X|Y}(x|y) = \begin{cases} \dfrac{p_{XY}(x,y)}{p_Y(y)}, & p_Y(y) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

x is held fixed

y is held fixed



(b)

"slice and normalize"
y is held fixed

This graph shows the Conditional PDF

Graph from B. P. Lathi's book: *Modern Digital & Analog Communication Systems*

<u>Independent RV's</u>

Independence should be thought of as saying that: neither RV impacts the other statistically – thus, the values that one <u>*will* likely take</u> should be irrelevant to the value that the other <u>*has* taken</u>.

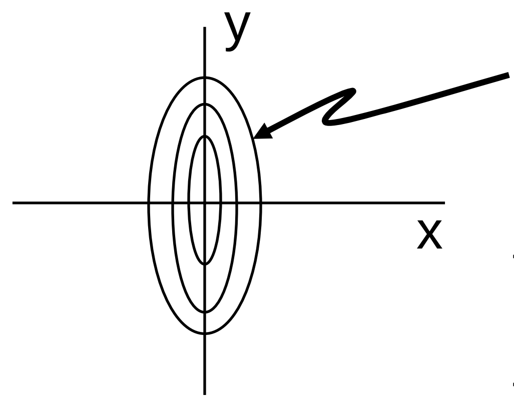
In other words: conditioning doesn't change the PDF!!!

$$p_{Y|X=x}(y\,|\,x) = \frac{p_{XY}(x,y)}{p_X(x)} = p_Y(y)$$

$$p_{X|Y=y}(x\,|\,y) = \frac{p_{XY}(x,y)}{p_Y(y)} = p_X(x)$$

Example: Independent Gaussian RVs

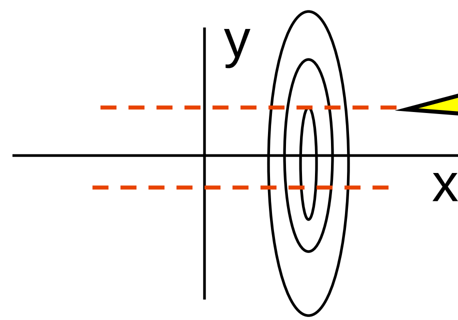
<u>Independent</u>
(zero mean)



Contours $p_{\mathbf{xy}}(x,y)$.

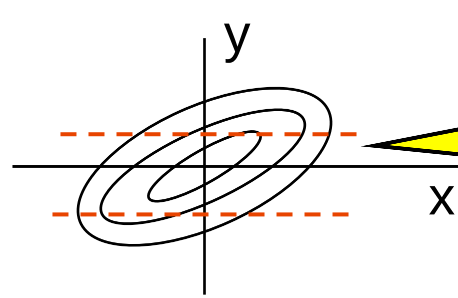
If X & Y are independent, then the contour ellipses are aligned with either the x or y axis

<u>Independent</u>
(non-zero mean)



<u>Different</u> slices give <u>same</u> normalized curves

<u>Dependent</u>



<u>Different</u> slices give <u>different</u> normalized curves

An “Independent RV” Result

RV’s X & Y are independent if:

$$p_{XY}(x, y) = p_X(x)\, p_Y(y)$$

Here’s why:

$$p_{Y|X=x}(y \mid x) = \frac{p_{XY}(x, y)}{p_X(x)} = \frac{\cancel{p_X(x)}\, p_Y(y)}{\cancel{p_X(x)}} = p_Y(y)$$

Characterizing RVs

- PDFs tell everything about RVs
  - but sometimes they are “more than we need/know”
- So… we make due with a few Characteristics
  - Mean of an RV (Describes the centroid of PDF)
  - Variance of an RV (Describes the spread of PDF)
  - Correlation of RVs (Describes “tilt” of joint PDF)

Mean of RV

Mean = Average = Expected Value

Call it E{X}

Motivation First w/ Data Analysis View

Consider RV X = Score on a test   Data: $X_1, X_2, \ldots X_N$

Possible values of X : $V_0 \;\; V_1 \;\; V_2 \ldots \;\; V_{100}$

0   1   2 … 100

Test Average $= \dfrac{\sum_{i=1}^{N} X_i}{N} = \dfrac{N_0 V_0 + N_1 V_1 + N_2 V_2 + \ldots N_n V_{100}}{N} = \sum_{i=0}^{100} V_i \dfrac{N_i}{N}$

$\dfrac{N_i}{N} \approx P(X=V_i)$

N_i = # of scores of value V_i

$N = \sum_{i=1}^{n} N_i$ (Total # of scores)

This is called Data Analysis or Empirical View → Statistics

Theoretical View of Mean

Data Analysis View leads to Probability Theory:

- For Discrete random Variables :

$$E\{X\} = \sum_{n=1}^{n} x_i P_X(x_i)$$

Probability

- This Motivates form for Continuous RV:

$$E\{X\} = \int_{-\infty}^{\infty} x \, p_X(x) dx$$

PDF

Notation: $E\{X\} = \overline{X}$

Aside: Probability vs. Statistics

Probability Theory
» Given a PDF Model
» Predict how the data will behave

Statistics
» Given a set of data
» Determine how the data did behave

$$E\{X\} = \int_{-\infty}^{\infty} x \, p_X(x) dx$$

Dummy Variable

PDF

≈

"Law of Large Numbers"

$$Avg = \frac{1}{N}\sum_{i=1}^{n} X_i$$

Data

There is no DATA here!!!
The PDF models how data will behave

There is no PDF here!!!
The Statistic measures how the data did behave

Variance of RV

There are similar Data vs. Theory Views here… Let's go to the theory

Variance measures extent of Deviation Around the Mean

Variance: $\sigma^2 = E\{(X - m_x)^2\}$

$$= \int (x - m_x)^2 p_X(x) dx$$

Note : If zero mean…

$$\sigma^2 = E\{X^2\}$$

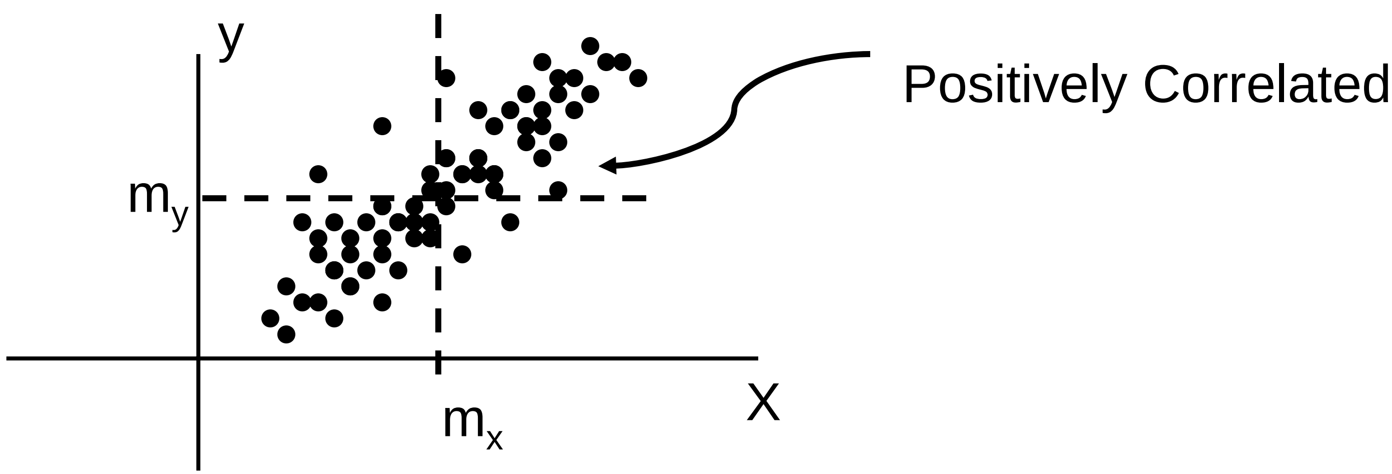
$$= \int x^2 p_X(x) dx$$

Correlation Between RV's

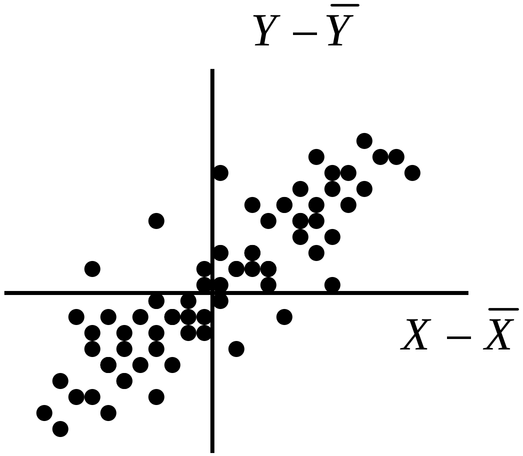
Motivation First w/ Data Analysis View

Consider a random experiment with two outcomes

$\Rightarrow$ 2 RVs X and Y of height and weight respectively

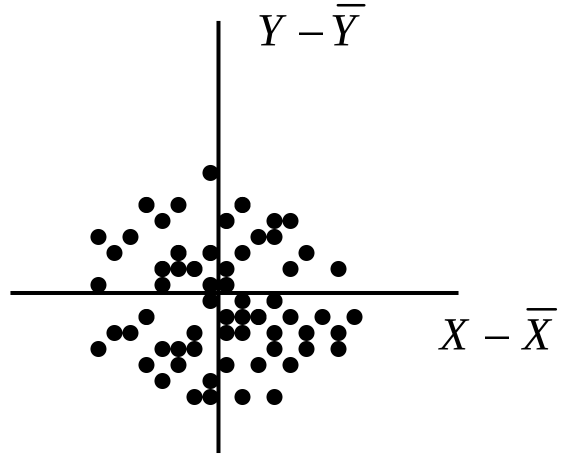


Three main Categories of Correlation



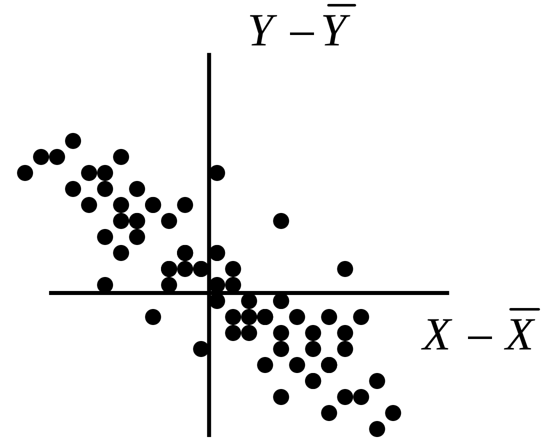
Positive correlation
“Best Friends”

Height
&
Weight



Zero Correlation
i.e. uncorrelated
“Complete Strangers”

Height
&
$ in Pocket



Negative Correlation
“Worst Enemies”

Student Loans
&
Parents’ Salary

Now the Theory…

To capture this, define Covariance :

$$\sigma_{XY} = E\{(X - \overline{X})(Y - \overline{Y})\}$$

$$\sigma_{XY} = \int\int (x - \overline{X})(y - \overline{Y})\, p_{XY}(x, y)\, dx dy$$

If the RVs are both Zero-mean : $\sigma_{XY} = \mathrm{E}\{XY\}$

If X = Y: $\sigma_{XY} = \sigma_X^2 = \sigma_Y^2$

If X & Y are independent, then: $\sigma_{XY} = 0$

If $\sigma_{XY} = E\{(X - \overline{X})(Y - \overline{Y})\} = 0$

Say that X and Y are "uncorrelated"

If $\sigma_{XY} = E\{(X - \overline{X})(Y - \overline{Y})\} = 0$

Then $\underbrace{E\{XY\}}_{\text{Called "Correlation of X \&Y"}} = \overline{X}\,\overline{Y}$

So… RVs X and Y are said to be uncorrelated

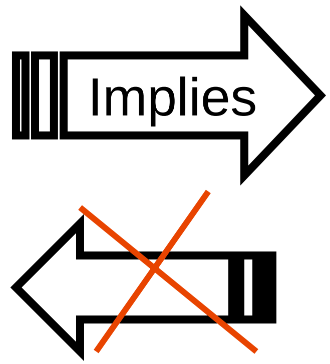
if $E\{XY\} = E\{X\}E\{Y\}$

Independence vs. Uncorrelated

X & Y are Independent

$$f_{XY}(x,y) = f_X(x)\,f_Y(y)$$

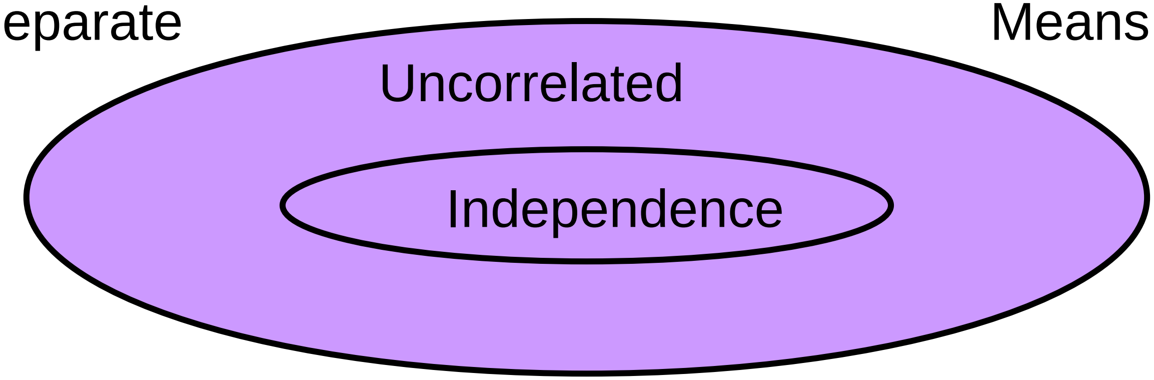
PDFs Separate



X & Y are Uncorrelated

$$E\{XY\} = E\{X\}E\{Y\}$$

Means Separate



INDEPENDENCE IS A STRONGER CONDITION !!!!

Confusing Terminology…

Covariance : $\sigma_{XY} = E\{(X - \bar{X})(Y - \bar{Y})\}$

Correlation : $E\{XY\}$

Same if zero mean

Correlation Coefficient : $\rho_{XY} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$

$-1 \le \rho_{XY} \le 1$

For Random Vectors…

$$\mathbf{x} = [X_1 \; X_1 \; \cdots \; X_N]^T$$

Correlation Matrix :

$$\mathbf{R}_\mathbf{x} = E\{\mathbf{x}\mathbf{x}^T\} = \begin{bmatrix} E\{X_1 X_1\} & E\{X_1 X_2\} & \cdots & E\{X_1 X_N\} \\ E\{X_2 X_1\} & E\{X_2 X_2\} & \cdots & E\{X_2 X_N\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{X_N X_1\} & E\{X_N X_2\} & \cdots & E\{X_N X_N\} \end{bmatrix}$$

Covariance Matrix :

$$\mathbf{C}_\mathbf{x} = E\{(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^T\}$$