Lecture 1

1 Review of Some Statistical Concepts

The notation $\sum$ (sigma), in mathematical term, denotes the **summation**

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n \tag{Eq.1}$$

The noteworthy properties of summation include:

1. $\sum_{i=1}^{n} k = nk$
2. $\sum_{i=1}^{n} kx_i = k\sum_{i=1}^{n} x_i$, where k is a constant term.
3. $\sum_{i=1}^{n}(a + bx_i) = na + b\sum_{i=1}^{n} x_i$, where a and b are constants.
4. $\sum_{i=1}^{n}(X_i + Y_i) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i$, where a, b and k are constant.

Multiple summation is the summation of variable that is in the form of matrix, shown as,

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij} = \sum_{i=1}^{n}(x_{i1}+x_{i2}+...+x_{im}) = (x_{11}+x_{21}+...+x_{n1})+(x_{12}+x_{22}+...+x_{n2})+...(x_{1m}+x_{2m}+...+x_{nm}) \tag{Eq.2}$$

where

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}_{nXm}$$

The significant properties of multiple summations are:

1. $\sum_{i=1}^{n}\sum_{j=1}^{m} X_{ij} = \sum_{j=1}^{m}\sum_{i=1}^{n} X_{ij}$
2. $\sum_{i=1}^{n}\sum_{j=1}^{m} X_iY_j = \sum_{i=1}^{n} X_i \times \sum_{j=1}^{m} Y_j$
3. $\sum_{i=1}^{n}\sum_{j=1}^{m}(X_{ij} + Y_{ij}) = \sum_{i=1}^{n}\sum_{j=1}^{m} X_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij}$

4. $(\sum_{i=1}^{n} X_i)^2 = \sum_{i=1}^{n} X_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} X_i X_j$

The product operator $\prod$ is defined as:

$$\prod_{i=1}^{n} x_i = x_1 * x_2 * x_n \qquad \text{(Eq.3)}$$

2 Experiment

Sample space is the set of all possible results of an experiment. For example, if you toss the coin twice, all feasible outcomes are composed of head twice, head followed by tail, tail followed by head, and tail twice. Let H and T denotes head and tail, respectively. The sample space can be written as,

$$SS = \{HH, HT, TH, TT\}$$

Sample Point is the member of sample space, eg. the event that head occurs twice from tossing a coin twice. Specifically, sample point is,

$$SP = HH \text{ or } HT \text{ or } TH \text{ or } TT$$

Events are the set of specific consequences of the experiment such as the events that head occurs twice. Events are the subset of sample space.

$$A = \text{the event that head occurs twice} = \{HH\}$$

Events are **mutually exclusive**, if the occurrence of one event makes no other events in sample space possible. As an illustration, for the experiment of tossing two coins once, let C be the event that both turn head and D be the event that both turn tail. Since C and D cannot happen at the same time, these two events are said to be mutually exclusive. Another example is the experiment of drawing one card from the standard 52-card deck, let E be the event that the rank of card is King and F be the event that suit of card is Clubs. As the event E and F can occur simultaneously, namely the King of Clubs, the two events are not mutually exclusive.

Events are **collectively exhaustive** if they cover all possible outcomes in the sample space. With the experiment of tossing the coin twice, let A be the event that head appears twice, B be the event that tail appears twice, and C be the event that head and tail each appear once. In this case, A, B and C are collectively exhaustive since all events cover all possible results from sample space; that is, HH, HT, TH and TT.

3 Probability and Random Variable

Probability is the possibility that any event will occur, given some specific sample space.

Let A be the event occurring in the given sample space and $P(A)$ be the probability that A will happen. Then, $P(A)$ is defined as;

$$P(A) = \frac{\text{the number of times the event A will occur}}{\text{the number of all possible outcomes in sample space}} \tag{Eq.4}$$

For instance, to draw one card from the standard 52-card deck, let A be the event that the rank of card is 2. Times the event will occur is 4 and the amount of all possible outcomes is 52; hence, the probability of A is $\frac{4}{52}$ or $\frac{1}{13}$.

Some properties of probability are;

1. $0 \leq P(A) \leq 1$

2. If A, B and C are exhaustive set, then,

$$P(A) + P(B) + P(C) = 1$$

3. If A, B and C are mutually exclusive, then,

$$P(A + B + C) = P(A) + P(B) + P(C)$$

Suppose that the results of an experiment are in the form of value, the variable, whose value is determined by one of those results, is known as **Random Variable**. Random variable can be either **discrete** or **continuous value**.

For discrete random variable, the example is the sum of the values on the face of two dice, when rolling two dice once. In other word, the obtained sum will range from 2 to 12, and it is impossible to get 2.5 or 3.5.

For continuous random variable, the example is the height of the high-school student, constricted to the range from 160 to 180 centimetres. It can be seen that the value of the height need not be the integers and can take the value of 160.5 or 160.52 centimetres.

These two distinct characteristics of random variable enable us to classify them into different probability density functions, which would be stated in Section 2.4.

4 Probability Density Function

As the value of random variable depends on an experiment, the **probability density function** would portray the overall image of possible random results. The type of the probability density function relies on the characteristics of the random variable. In this section, many important types are discussed.

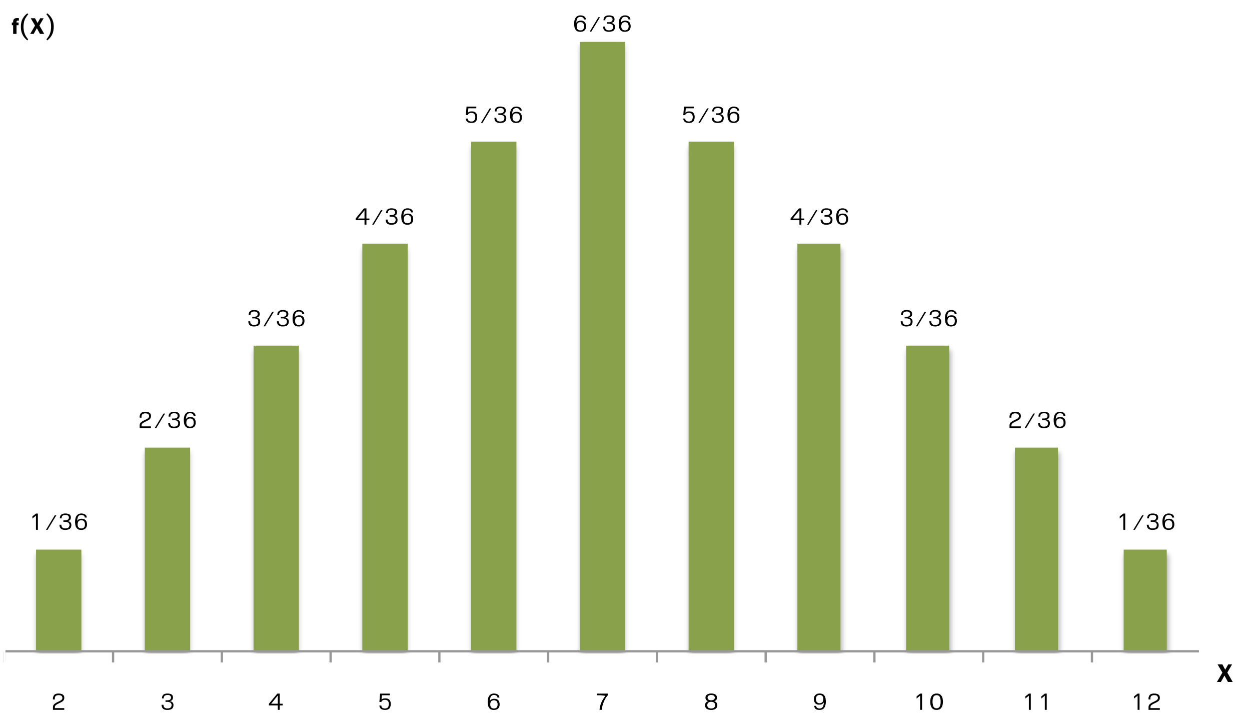
4.1 Probability Density Function for Discrete Random Variable

Let X be the discrete random variable with the value $x_1, x_2, \ldots, x_n$ and we get,

$$\begin{array}{rclcrcl} f(x) & = & P(X = x_i) & \text{for} & i & = & 1, 2, ..., n \\ f(x) & = & 0 & \text{for} & x & \neq & x_i \end{array}$$

Example: Let X be random variable of the sum of values on the face of two dices. The value might be 2 or 12, that is the value from both rolling round is 1 or 6, respectively. The Figure 2-1 summarizes all possible results#

Figure 2-1: Probability Density function of the Sum of Values on the Side of the Dice, Obtained from Rolling the Dice Twice



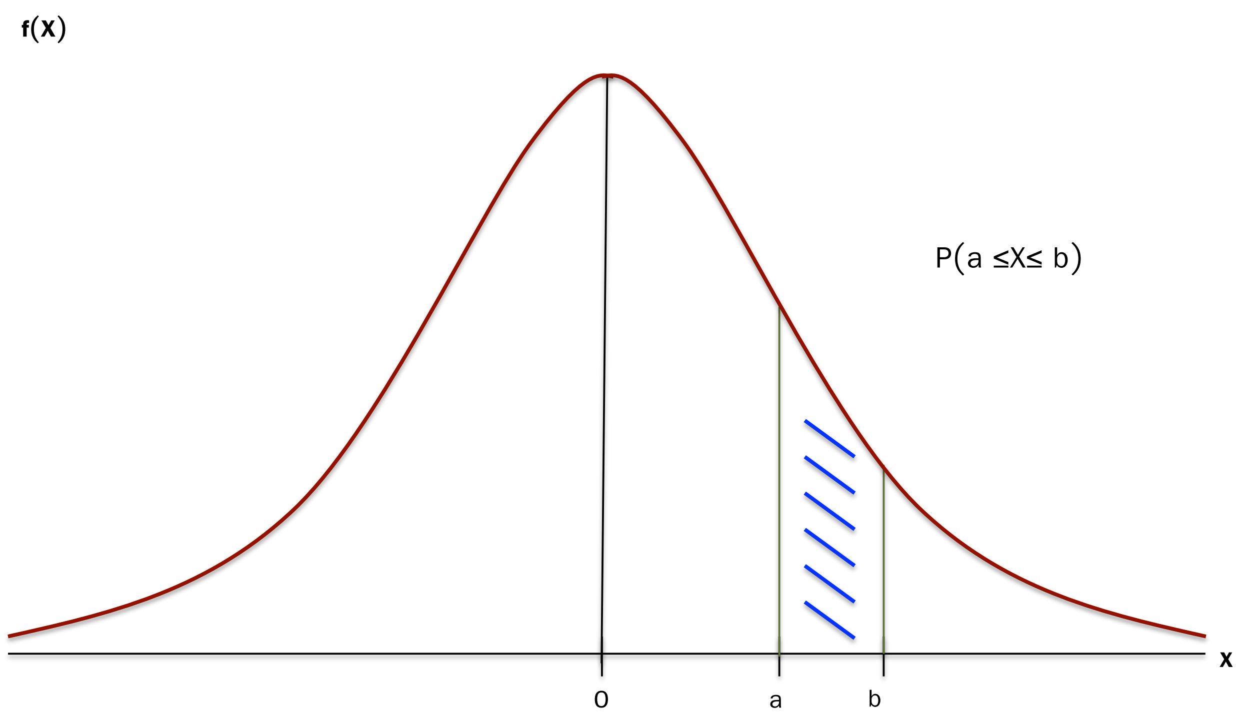
4.2 Probability Density Function for Continuous Random Variable

Let X be the continuous random variable. The probability density function of X satisfies the three following conditions.

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $\int_{a}^{b} f(x)dx = P(a \leq x \leq b)$

Figure 2-2 exhibits the probability density function for the continuous random variable, where the area under the curve represents the probability that the variable will lay on that range. Specifically, $P(a \leq X \leq b)$ means the probability that X will take the value between a and b.

Figure 2-2: Probability Density Function for Continuous Random Variable



Example-

4.3 Joint Probability Density Function

In this subsection, only **joint probability density function** for discrete variable is discussed. Let X and Y be discrete random variables. The joint probability density function, identifying the probability that X and Y happen simultaneously, is written as,

$$f(X, Y) = P(X = x \text{ and } Y = y)$$

Example: The following table explains the joint probability density function.

Table 2-1: The table illustrating the joint probability density function of X and Y

		X		
		-1	0	1
	1	0.11	0.08	0.05
Y	2	0.09	0.05	0.03
	3	0.35	0.07	0.17

According to the table, the probability that random variable X will be 0 and random variable Y will be 3 is 0.07 or 7 percent. In mathematical term, it can be written as $f(X = 0, Y = 3) = 0.07$.

4.4 Marginal Probability Density Function

The above joint probability density function $f(X, Y)$ shows the joint distribution of two variables. On the other hand, **marginal probability density function** with respect to joint probability function, displays the probability density function of single variable like $f(X)$, $f(Y)$, which can be derived from;

$$\begin{array}{lclll} f(X) & = & \sum_Y f(X,Y) & \text{called} & \text{marginal PDF of X} \\ f(Y) & = & \sum_X f(X,Y) & \text{called} & \text{marginal PDF of Y} \end{array}$$

where $\sum_Y$ or $\sum_X$ means the summation of probability over all values of X and Y respectively.
Example: According to Table 2-1 above, marginal PDF of X is obtained from

$$\begin{aligned} f(X=-1) &= \\ &= \\ &= \\ &= \\ f(X=0) &= \sum_Y f(X=0,Y) \\ &= f(X=0,Y=1)+f(X=0,Y=2)+f(X=0,Y=3) \\ &= 0.08+0.05+0.07 \\ &= 0.20 \\ f(X=1) &= \sum_Y f(X=1,Y) \\ &= f(X=1,Y=1)+f(X=1,Y=2)+f(X=1,Y=3) \\ &= 0.05+0.03+0.17 \\ &= 0.25 \end{aligned}$$

marginal PDF of Y is obtained from

$$\begin{aligned} f(Y=1) &= \\ &= \\ &= \\ &= \\ f(Y=2) &= \sum_X f(X,Y=2) \\ &= f(X=-1,Y=2)+f(X=0,Y=2)+f(X=1,Y=2) \\ &= 0.09+0.05+0.03 \\ &= 0.17 \\ f(Y=3) &= \sum_X f(X,Y=3) \\ &= f(X=-1,Y=3)+f(X=0,Y=3)+f(X=1,Y=3) \\ &= 0.35+0.07+0.17 \\ &= 0.59 \end{aligned}$$

According to the calculation above, the result can be summarized into Table 2-2.

Table 2-2: Table demonstrating joint probability of random variable X and Y

		X			
		-1	0	1	
	1	0.11	0.08	0.05	$f(Y=1)$ =
Y	2	0.09	0.05	0.03	$f(Y=2)$ =
	3	0.35	0.07	0.17	$f(Y=3)$ =
		$f(X=-1)$ =	$f(X=0)$ =	$f(X=1)$ =	$f(X)=$ $f(Y)=$

4.5 Conditional Probability Density Function

Conditional probability density function is the probability of one event given that some events have already occurred. The function is written as,

$$f(X|Y) = P(X = x|Y = y)$$

This function can be obtained from the joint probability density function through,

$$f(X|Y) = \frac{f(X,Y)}{f(Y)}$$

Example: According to Table 2.1, find $f(X = 1|Y = 2)$ and $f(Y = 2|X = 0)$

$$\begin{aligned} f(X=0|Y=1) &= \\ &= \\ &= \\ &= \\ f(Y=2|X=0) &= \\ &= \\ &= \end{aligned}$$

Example: Let event A be tossing the dice once and the point is odd number and B be the tossing the dice once and the point is at least 5. Find the probability that the point coming up is odd given that the point has to be at least 5.

Answer A and B will occur simultaneously if the point from tossing the dice is 5; so, the joint probability of A and B is $\frac{1}{6}$. The probability that B occurs is $\frac{2}{6}$. Hence, the conditional probability of A given B is

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{\frac{1}{6}}{\frac{2}{6}} = \frac{1}{2}\#$$

4.6 Statistical Independence

Two random variables are **independent** if the resulting value of one variable does not affect the resulting value of the other; namely,

$$f(X,Y) = f(X)f(Y)$$

Example: Consider Mr. Ake's expenditure for a meal and the Miss Somsri's expenditure for a dessert. Given that they do not know each other, the realization of Mr. Ake's expenditure does not imply the realization of Miss Somsri's expenditure. We can, thus, conclude that the expenditures of these two people are independent#

Example: Consider drawing cards sequentially from the standard 52-card deck without putting it back into the deck. Once the first card is drawn, the probability of drawing the second card will be influenced because the amount of cards in the deck is reduced. In this case, it can be concluded that drawing the first and second card are not independent#

5 Expectation, Variance, Covariance and Correlation

5.1 Mean or Expected Value

Because the value of random variable hinges on the value of random results of experiment which cannot be determined certainly, statisticians have invented the measures of central tendency of the random variable. One of them is **expected value**, indicating the mean of the random variable.

For discrete random variable, the expected value is calculated by;

$$E(X) = \sum_{i=i}^{n} x_i f(x_i) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_n f(x_n)$$

For continuous random variable, the expected value is calculated by,

$$E(X) = \int_a^b x f(x) dx$$

where;

$E(X)$ is the measure of central tendency of random variable, resulting from repeated trial of experiment.

$\sum_{i=i}^{n} x_i f(x_i)$ is the average of random variable weighted by the probability corresponding to each value.

a and b are the lowest and highest constant possible respectively.

Example: Find the expected value of rolling two dice once (Figure 2-1)

Example:

Crucial properties of expected value include:

1. $E(b) = b$
2. $E(aX + b) = aE(X) + b$
3. $E(XY) = E(X)E(Y)$; given that X and Y are independent
4. $E(g(X)) = \sum_x g(X)f(X)$

where a and b are constant.

Conditional expectation value is the expectation value of random variable under some conditions such as expected value of X conditional on Y or $E(X|Y = 5)$

Let $f(X, Y)$ be the joint probability function of X and Y. The expectation of X conditional on some value of Y is defined as,

For discrete random variable $E(X|Y = y) = \sum_X X_i f(X|Y = y)$

For continuous random variable $E(X|Y = y) = \int_{-\infty}^{\infty} X_i f(X|Y = y)$

Example

5.2 Variance

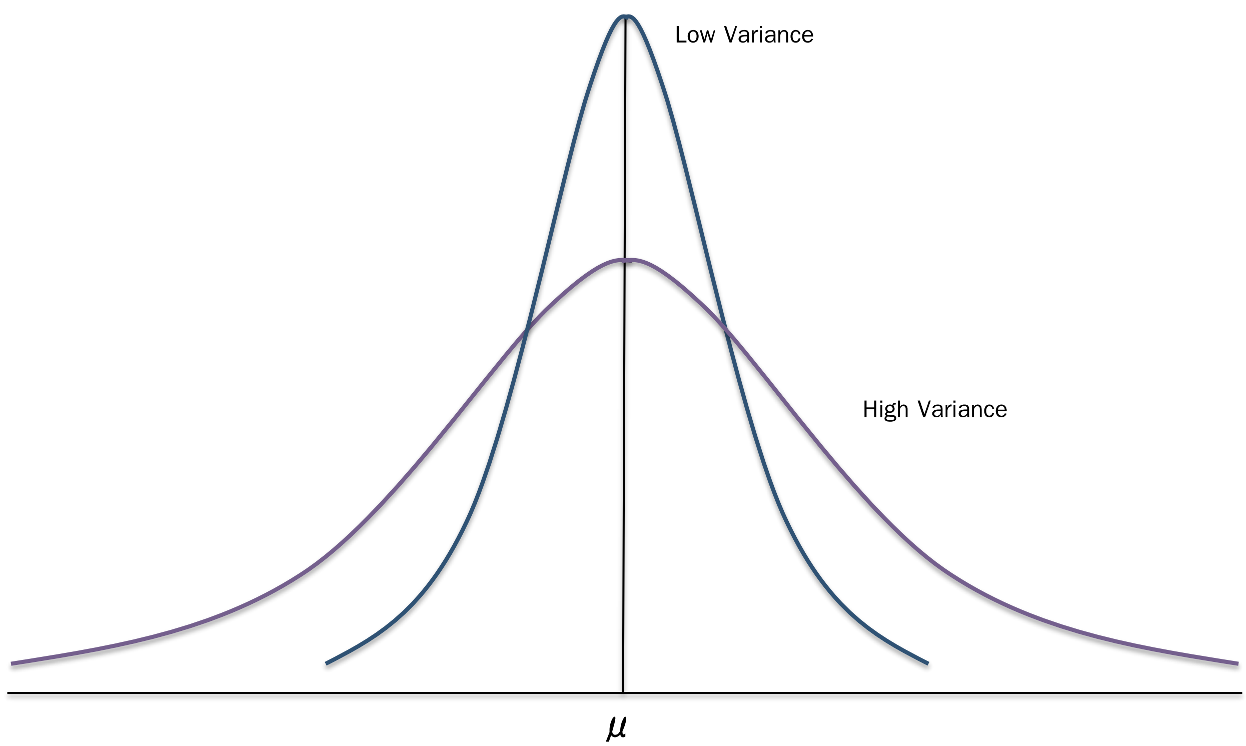
Variance is the measure of dispersion of the value of variable around the expected value. The higher the variance, the more dispersing the random variable (Figure 2-3). If X is the random variable with expected value μ, we get;

$$Var(X) = \sigma_X^2 = E[X - E(X)]^2 = E(X)^2 - \mu^2 \quad \text{(Eq.5)}$$

From,

$$
\begin{aligned}
Var(X) &= \sigma_X^2 \\
&= E[X - E(X)]^2 \\
&= E[X^2 - 2XE(X) + (E(X))^2] \\
&= E(X^2) - 2(E(X))^2 + (E(X))^2 \\
&= E(X)^2 - \mu^2
\end{aligned}
$$

Figure 2-3: Distribution of Random Variables with Different Variance



Important properties of expected value include;

1. $Var(b) = 0$

2. $Var(aX + b) = a^2 Var(X)$

3. $Var(X \pm Y) = Var(X) + Var(Y)$; given that X and Y are independent

4. $Var(aX \pm bY) = a^2 Var(X) + b^2 Var(Y)$

where a and b are constant.

Lecture 2

5.3 Conditional Variance

The conditional variance of X is given Y = y is defined as following:

$$\begin{aligned} var(X|Y=y) &= E\left\{[X - E(X|Y=y)]^2 | Y=y\right\} \\ &= \sum_x [X - E(X|Y=y)]^2 f(x|Y=y) \\ &= \int_{-\infty}^{\infty} [X - E(X|Y=y)]^2 f(x|Y=y)dx \end{aligned} \quad \text{(Eq.6)}$$

Example

Properties of conditional expectation and conditional variance

5.4 Covariance

Theorem. Let X and Y be two random variables with means μ_x and μ_y, respectively. Then, we can define the covariance between these two variables as following:

$$cov(X,Y) = E\left\{(X-\mu_x)(Y-\mu_y)\right\} = E(XY) - \mu_x\mu_y \qquad \text{(Eq.7)}$$

If X and Y are continuous random variables we can calculate their cov(X,Y):

$$\begin{aligned} cov(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (X-\mu_x)(Y-\mu_y) f(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} XYf(x,y)dxdy - \mu_x\mu_y \end{aligned} \qquad \text{(Eq.8)}$$

Properties of Covariance

1. If X and Y are independent, the covariance between X and Y is zero.

Proof:

2. $cov(a+bX, c+dY) = bd * cov(X,Y)$, where a,b,c, and d are constants.

Example Suppose the join PDF of random variables X and Y can be represented as in the below table. What is the covariance between X and Y?

		X			
		1	2	3	
	1	0.25	0.25	0	$f(Y=1)$ =
Y	2	0	0.25	0.25	$f(Y=2)$ =
		$f(X=1)$ =	$f(X=2)$ =	$f(X=3)$ =	$f(X) =$ $f(Y) =$

Next, we will turn our attention to seeing how we can apply the covariance to calculate the correlation between the random variables X and Y

5.5 Correlation

When we calculate the covariance of X and Y, it reflects the units of both random variables. However, it is useful to have **a dimensionless measure of dependency** by calculating the correlation instead.

Definition Let X and Y be any two random variables (discrete or continuous) with standard deviation σ_X and σ_Y, respectively. The **correlation coefficient** of X and Y, denoted **corr(X,Y)** or ρ_{XY} (the greek letter "rho") is defined as:

$$\rho_{XY} = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}} = \frac{cov(x,y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Example Suppose the join PDF of random variables X and Y can be represented as in the below table. What is the correlation between X and Y?

		X		
	1	2	3	
1	0.25	0.25	0	$f(Y=1)$ =0.5
Y 2	0	0.25	0.25	$f(Y=2)$ =0.5
	$f(X=1)$ $=0.25$	$f(X=2)$ $=0.5$	$f(X=3)$ =0.25	$f(X)=1$ $f(Y)=1$

From the definition, ρ_{XY} is measure of linear association between two random variables. The value of ρ lies between -1 and +1, $-1 \leq \rho_{XY} \leq +1$. We can interpret the value of correlation as:

▶ If $\rho_{XY} = 1$, then X and Y are perfectly, positively, linearly correlated.
▶ If $\rho_{XY} = -1$, then X and Y are perfectly, negatively, linearly correlated.
▶ If $\rho_{XY} = 0$, then X and Y are completely,un-linearly correlated. This means that X and Y may correlated in some other manner i.e. a parabolic manner., but NOT in a linear manner
▶ If $\rho_{XY} \leq 0$, then X and Y are positively, linearly correlated, but NOT perfectly.
▶ If $\rho_{XY} \geqq 0$, then X and Y are negatively, linearly correlated, but NOT perfectly.

Theorem. If X and Y are independent random variables, then:

$$corr(X, Y) = cov(X, Y) = 0$$

Example: Let X = the outcome of a fair, black, 6-sided die.
Let Y =outcome of a fair, red, 4-sided die.
What is the covariance of X and Y? What is the correlation of X and Y?

NOTE: The converse of the theorem is NOT NECESSARILY CORRECT!

Example: Let X and Y be two discrete random variables with the following join PDF:

		X -1	X 0	X 1	
Y	-1	0.20	0	0.20	$f(Y=-1)$ =
Y	0	0	0.20	0	$f(Y=0)$ =
Y	1	0.20	0	0.20	$f(Y=1)$ =
		$f(X=-1)$ =	$f(X=0)$ =	$f(X=1)$ =	$f(X)=$ $f(Y)=$

What is the correlation between X and Y? And, are X and Y independent?

5.6 Variances of Correlated Variables

Let X and Y be two random variables, then

$$\begin{aligned} var(X+Y) &= var(X)+var(Y)+2cov(X,Y) \\ &= var(X)+var(Y)+2\rho\sigma_x\sigma_y \\ var(X-Y) &= var(X)+var(Y)-2cov(X,Y) \\ &= var(X)+var(Y)-2\rho\sigma_x\sigma_y \end{aligned}$$

(Eq.9)

The generalized result:
Let $\sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$, then the variance of the linear combination $\sum X_i$ is:

$$\begin{aligned} var\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n} var(X_i) + 2\sum_{i<j}\sum cov(X_i, X_j) \\ &= \sum_{i=1}^{n} var(X_i) + 2\sum_{i<j}\sum \rho_{ij}\sigma_i\sigma_j \end{aligned}$$

(Eq.10)

Example:
what is the var$(X_1 + X_2 + X_3)$?

Lecture 3

5.7 Higher Moments of Probability Distributions

In the previous subsection, we have already discussed about mean, variance, and covariance as the measures of the first and second moments of univariate and multivariate PDFs. Besides the first two moments, we are occasionally interested in the higher moments such as the third and fourth moments which are normally applied in studying the "Shape" of the distribution. In general, the r^{th} moments about the mean is defined as

$$r^{th}\textbf{moment} : E(X-\mu)^r$$

By the definition of r^{th} moments, we can easily define the third and fourth moments as:

Third moment:

$$E(X-\mu)^3$$

Fourth moment:

$$E(X-\mu)^4$$

We can study the shape of the distribution by calculating **skewness** and **kurtosis**.

SKEWNESS is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean.

One measure of skewness is defined as:

$$\begin{aligned} S &= \frac{E(X-\mu)^3}{\sigma^3} \\ &= \frac{\text{third moment about the mean}}{\text{cube of the standard deviation}} \qquad \text{(Eq.11)} \end{aligned}$$

KURTOSIS is a measure of the peakedness of the probability distribution of a real-valued random variable

We can also measure kurtosis as:

$$\begin{aligned} S &= \frac{E(X-\mu)^4}{\sigma^4} \\ &= \frac{\text{fourth moment about the mean}}{\text{square of the second moment}} \end{aligned} \qquad \text{(Eq.12)}$$

♣ **Platykurtic (fat or short-tailed)** $\Longrightarrow$ PDFs with Kurtosis < 3
♣ **Leptokurtic (slim or long-tailed)** $\Longrightarrow$ PDFs with Kurtosis > 3
♣ **Mesokurtic (which is the normal distribution)** $\Longrightarrow$ PDFs with Kurtosis $= 3$

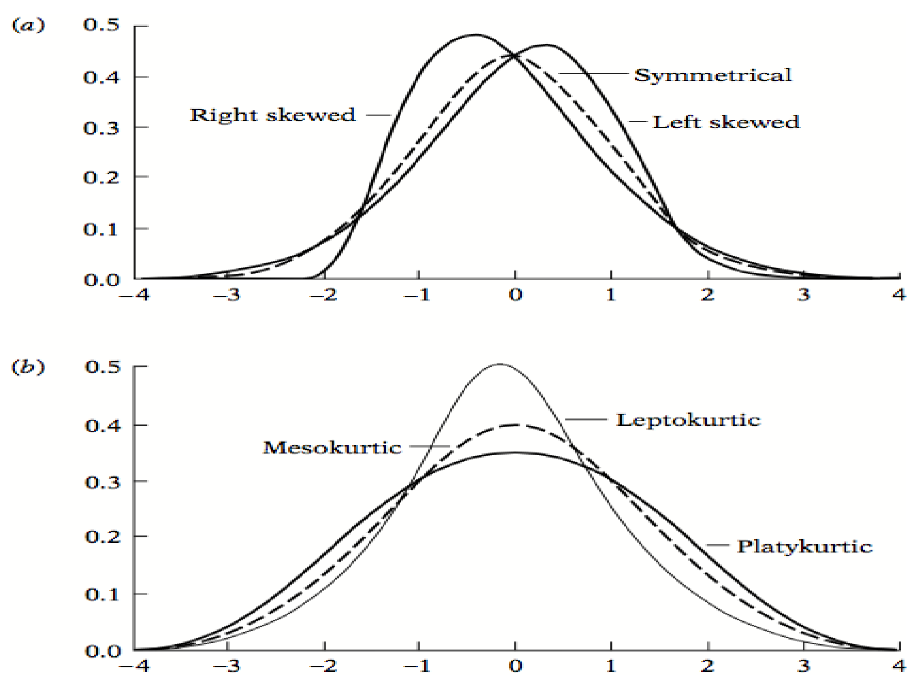


Figure 1. (a) Skewness; (b)Kurtosis

6 Some important probability distribution

6.1 Normal Distribution

A continuous random variable X has a normal distribution with mean μ and variance σ^2 if its probability density function (pdf) is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \quad \text{where} \quad -\infty < x < \infty$$

NOTE: The normal distribution can be described by two parameters

- $\mu =$ The mean of the distribution.
- $\sigma =$ The standard deviation of the distribution.

Therefore, changing the values of μ and σ alter the positions and shapes of the distributions.

If X is Normally distributed with mean μ and standard deviation σ, we can write it as:

$$X \sim N(\mu, \sigma^2)$$

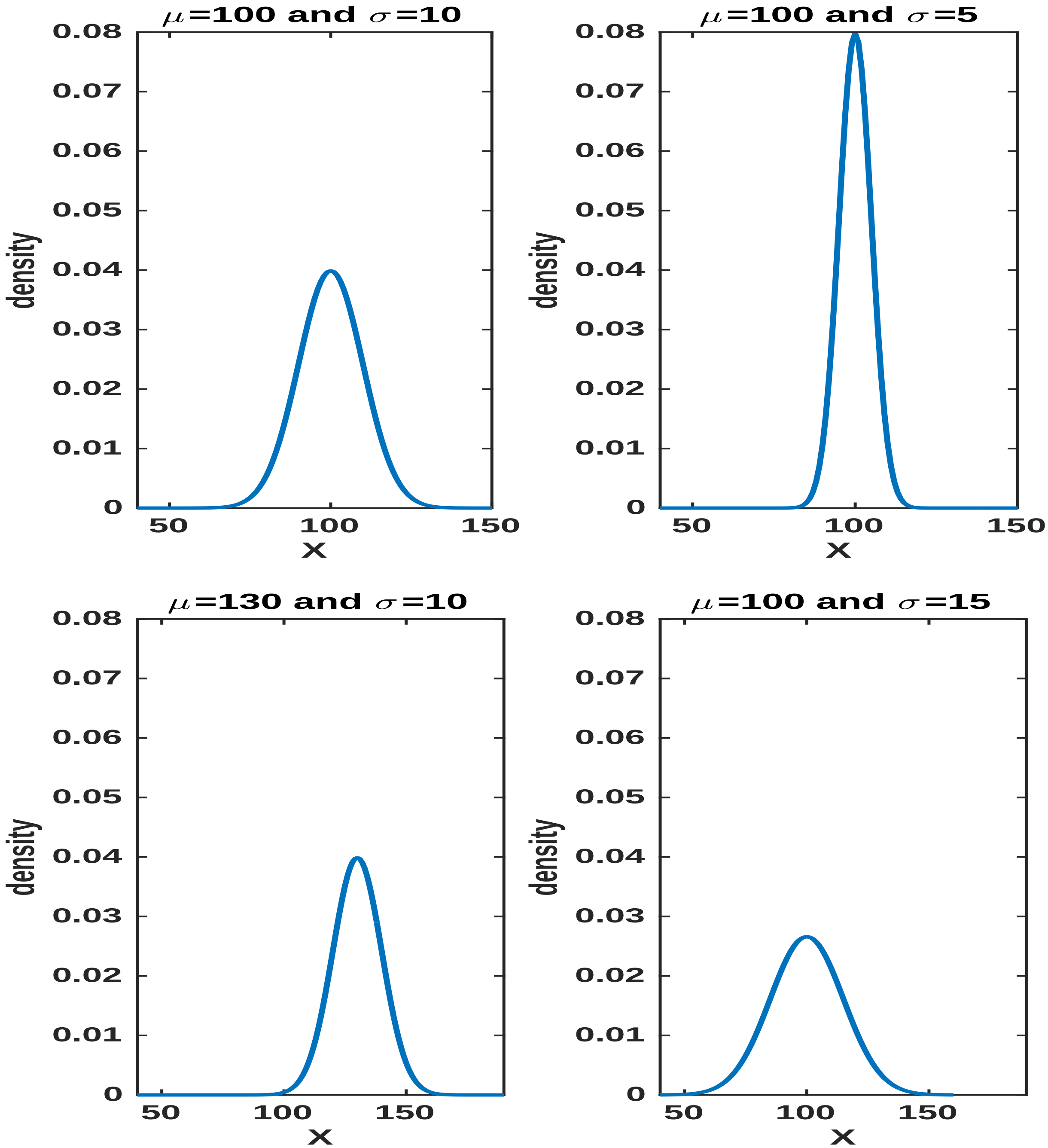


Figure 2. Compare the mean and standard deviation of the normal distribution

The properties of the normal distribution.

★ It is symmetrical around its mean value.

★ About 68 percent of the area under the normal distribution lies between the value $\mu \pm \sigma$
About 95 percent of the area under the normal distribution lies between the value $\mu \pm 2\sigma$
About 99.7 percent of the area under the normal distribution lies between the value $\mu \pm 3\sigma$
(as shown in figure 2)

★ We can convert the given normally distributed variable X with mean μ and σ^2 into the standardized normal variable Z by calculating Z where Z can be defined as:

$$Z =$$

With the standardized normal variable Z, we can rewrite the normal pdf as:

$$f(Z) =$$

In sum, you can see that we convert the given normally distributed variable X into the standardized normal variable by:
(i) Subtracting the mean μ
(ii)Dividing by the standard deviation σ

♡ Subtracting the mean re-centers the distribution on zero.
♡ Dividing by the standard deviation re-scales the distribution so it has standard deviation 1.

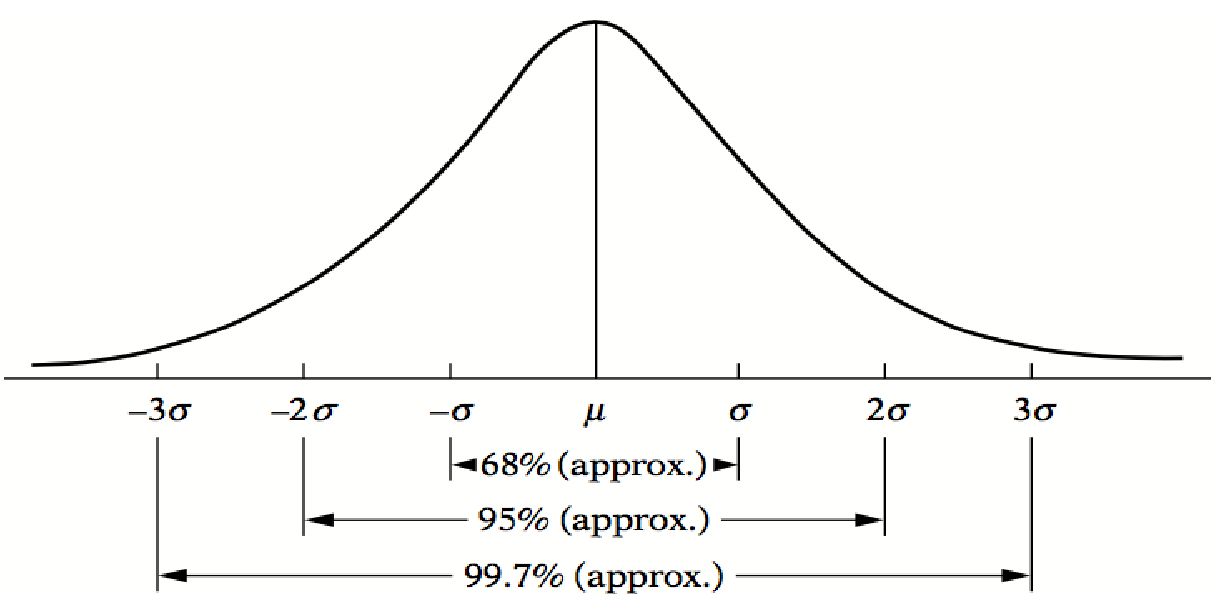


Figure 3. Areas under the normal distribution

It should be remarked that it mean value is zero and its variance is unity for any standardized variable.

By convention, we can denote a normally distributed variable X with zero mean and unit variance as

$$X \sim N(0,1)$$

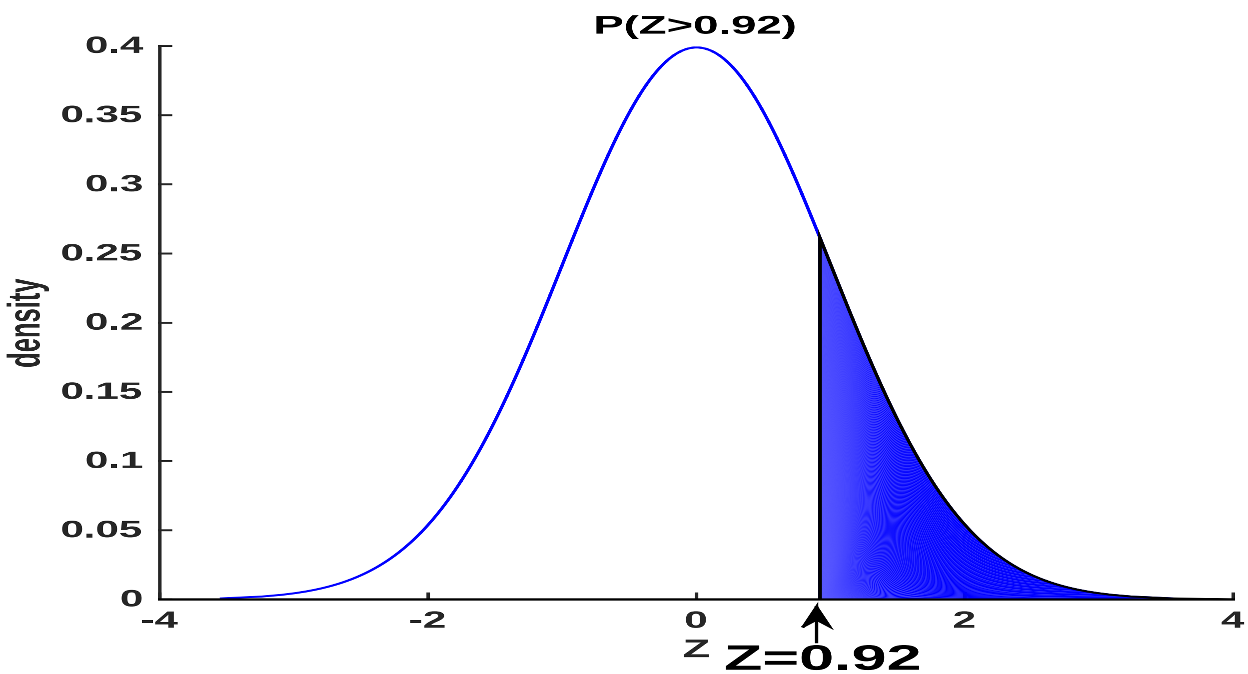


Figure 4. If Z ∼ N(0,1), the probability that P(Z>0.92)

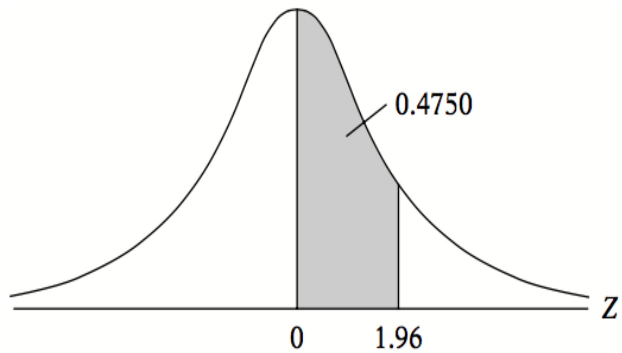
Example If Z ∼ N(0,1) what is P(Z>0.92)?

AREAS UNDER THE STANDARDIZED NORMAL DISTRIBUTION

Example

$\Pr(0 \leq Z \leq 1.96) = 0.4750$

$\Pr(Z \geq 1.96) = 0.5 - 0.4750 = 0.025$



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4454	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Example If $Z \sim N(0,1)$ what is $P(-0.64<Z<0.43)$?

Example If $X \sim N(3500,500^2)$ what is $P(X<3100)$?

Let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ and assume that X_1 and X_2 are independent. If we have the linear combination between X_1 and X_2 where we can write it as:

$$Y = aX_1 + bX_2,$$

where a and b are the constant terms. Then

$$Y \sim N\left[(a\mu_1 + b\mu_2), (a^2\sigma_1^2 + b^2\sigma_2^2)\right]$$

In other words, **a linear combination of normally distributed variables is itself normally distributed.**

Central limit theorem Let $X_1, X_2, ..., X_n$ denote n independent random variables and

$$X_i \sim N(\mu, \sigma)$$

Let $\bar{X}$=$\sum \frac{X_i}{n}$, then as n increases indefinitely (i.e, n $\rightarrow \infty$),

The third and fourth moments of the normal distribution:

Third moment: $E(X - \mu)^3 = 0$

Fourth moment: $E(X - \mu)^4 = 3\sigma^4$