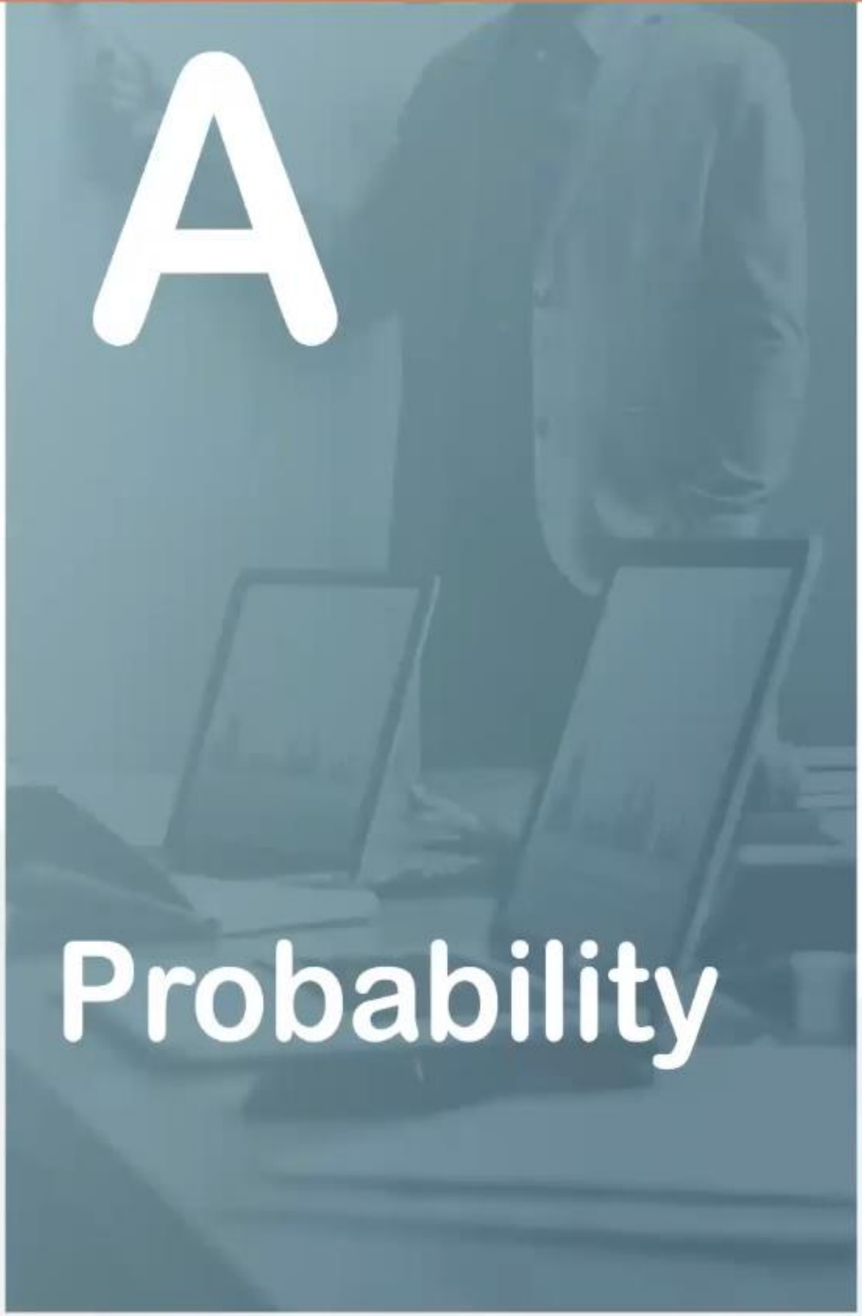
Probability & Bayes Theorem

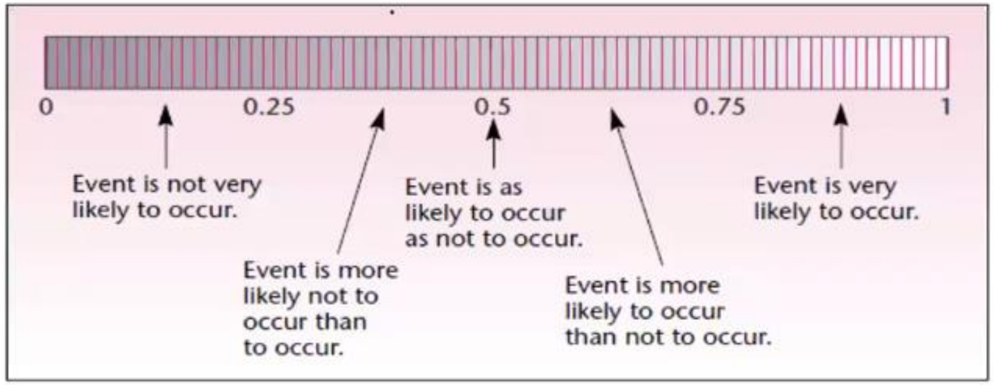
Agenda..

#	Topics
1	Probability
2	Bayes theorem


A
Probability

Probability

- Probability is a quantitative measure of uncertainty
  - a number that conveys the strength of our belief in the occurrence of un-certain event
- Theory of probability was largely developed by European mathematicians such as Galileo Galilei (1564-1642), Blaise Pascal (1623 -1662), Pierre de Fermat (1601-1665), Abraham de Moivre (1667 -1754).
- Development of probability theory in Europe is often associated with gamblers, who pursued their interests in the famous European Casinos, such as the one at Monte Carlo.



Examples

- Probability of even no when we throw a dice once.
- Chance of rolling a 4 with a single throw of dice.
- There are 5 marbles in a bag, 4 are blue and 1 is red. What is the probability that blue marble gets picked up.

Probability & Rules

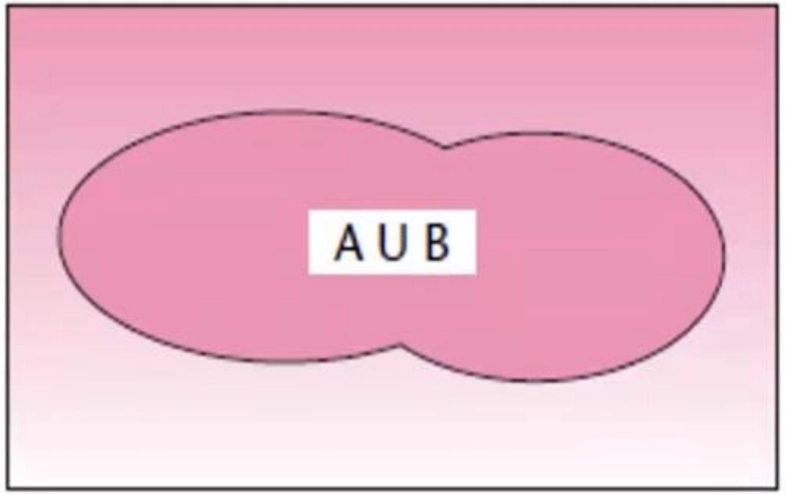
- Probability of event A: $P(A) = n(A) / n(S)$
- where:
- $n(A)$ = number of elements in the set of the event A
- $n(S)$ = number of elements in the sample space S

- For any event A, the probability satisfies $0 <= P(A) <= 1$

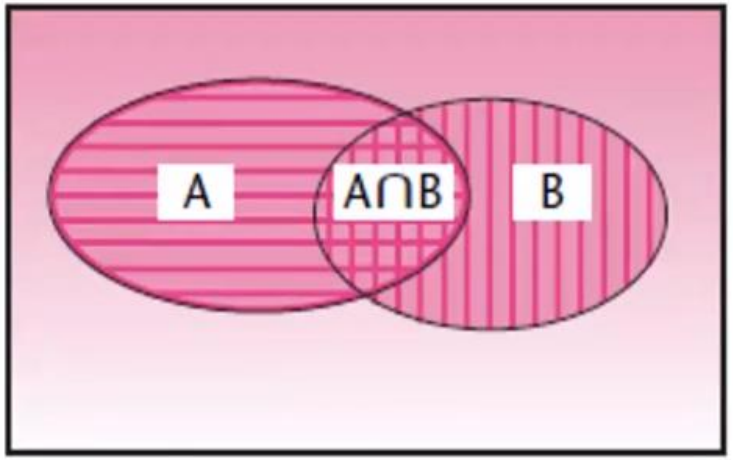
- Probability of the complement $P(\bar{A}) = 1 – P(A)$

Set Operations

UNION OF A OR B



INTERSECTION – A & B



Types of Probability

Probability

Objective or Classical Probability

- Probability based on chance. Here no personal judgement is involved.
- Ex. The numbers 1,2,3,4,5,6 on a fair die are equally likely to occur.
- Probability that a head will appear on any one toss of a coin is ½.

Subjective or Personal Probability

- Probability involves personal judgement.
- Ex. A physician assessing the probability of a patient's recovery is based on personal judgement based on what he / she know and feel about the situation.

Addition Rule of Probability P(A U B) or P(A or B)

Two Events
(A or B)

Mutually Exclusive
(No Intersection, Disjoint)

- **$P(A \cup B) = P$ (A or B) $= P(A) + P(B)$**
  - For mutually exclusive events A & B $P(A \cap B) = 0$

- Examples
  - Toss of a coin
  - Roll of a Die
  - Set of Even Integers & Set of odd Integers

Mutually Non-Exclusive

- **$P(A \cup B) = P$ (A or B) $= P(A) + P(B) - P(A \cap B)$**

- Example
  - Set of positive integers from 1 to 8 & set of even no's from 1 to 12

Independent & Dependent Events

Independent Event → One event is not affected by another event

- For example, if you choose a card from a deck of 52 cards, your probability of getting a Jack is 4 out of 52.
- P(Jack) = number of Jacks in a deck of cards / total number of cards in a deck = 4/52 = 1/13 ≈ 7.69%.
- **If you *replace* the jack and choose again (assuming the cards are shuffled), the events are independent.**
- Your probability remains the same (1/13). Choosing a card repeatedly would be an independent event, because each time you choose a card (a "trial" in probability) it's a **separate, non-connected event**.

Dependent Event → is also called conditional, where one event is dependent on another event

- **But what if the card was kept out of the pack the next time you choose?** Let's say you pulled the three of hearts, but you're still searching for that jack. The *second* time you pull out a card, the deck is now 51 cards, so:
- P(Jack) = number of Jacks in a deck of cards / total number of cards in a deck = 4/51 = 1/13 ≈ 7.84%
- The probability has increased from 7.69% (with replacement of the jack) to 7.84% (the jack isn't replaced), so choosing cards in this manner is an example of a **dependent event**.

Multiplication Rule of Probability (Joint Probability) P(A ∩ B) or P (A & B)

Two Events
(A & B)

Independent
(Probability of B has **no** bearing on A or vice versa)

- $P(A|B) = P(A)$, $P(B|A) = P(B)$
- Probability that both A & B occur
- $P(A \cap B) = P(A \text{ and } B) = P(AB) = P(A) * P(B)$

Dependent
(Probability of B has a bearing on A or vice versa; Conditional)

- Probability that both A & B occur
- $P(A \cap B) = P(A \text{ and } B) = P(AB) = P(A) * P(B|A)$

Example Basic Probability

HBO Subscribers

TV Shows	Male	Female	Total
Game of Thrones	0.16	0.24	**0.4**
West World	0.2	0.05	**0.25**
Other	0.1	0.25	**0.35**
Total	**0.46**	**0.54**	**1**

What is the probability that HBO subscriber is a male? (0.46)

What is the probability that subscriber preferring West world? (0.25)

What is the probability of HBO subscriber being a Male **AND** preferring West World? P(Male ∩ West World =0.2)

What is the probability of HBO subscriber being a Male **OR** preferring West World? P(Male U West World)

- Applying the formula $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.46 + 0.25 - 0.2 = 0.51$

Simple or Marginal Probability

If we divide everything by grand total i.e., 500, we get probability distribution

TV Shows	Male	Female	Total
Game of Thrones	0.16	0.24	0.4
West World	0.2	0.05	0.25
Other	0.1	0.25	0.35
Total	0.46	0.54	1

- **Simple Probability or Marginal Probability** (because in Margin)
  - P(GOT)=0.4
  - P(Westworld)=0.25
  - P(other)=0.35
  - P(Male)=0.46
  - P(Female)=0.54

- This probability ignores gender as if does not exist

Joint Probability

What Is a Joint Probability?

Joint probability is a statistical measure that calculates the likelihood of two events occurring together and at the same point in time. Joint probability is the probability of event Y occurring at the same time that event X occurs.

The Formula for Joint Probability Is

Notation for joint probability can take a few different forms. The following formula represents the probability of events intersection:

$$P\left(X \bigcap Y\right)$$

where:

$X, Y = \text{Two different events that intersect}$

$P(X \text{ and } Y), P(XY) = \text{The joint probability of X and Y}$

Joint Probability (And - ∩)

Joint Events: Depends on two classes M/F. Each of them is called joint events

TV Shows	Male	Female	Total
Game of Thrones	0.16	0.24	0.4
West World	0.2	0.05	0.25
Other	0.1	0.25	0.35
Total	0.46	0.54	1

- **Joint Probability**
  - P(Female And GOT) =0.24
  - P (Female ∩ GOT) =0.24
  - 0.16,0.24,0.2,0.05,0.1,0.25 all are joint probability distribution.
  - It sums to 1

Conditional Probability

Conditional probability is defined as the likelihood of an event or outcome occurring, based on the occurrence of a previous event or outcome. Conditional probability is calculated by multiplying the probability of the preceding event by the updated probability of the succeeding, or conditional, event.

For example:

- Event A is that it is raining outside, and it has a 0.3 (30%) chance of raining today.
- Event B is that you will need to go outside, and that has a probability of 0.5 (50%).
- P(A|B) = (0.3*0.5)/0.3 = 50%

A conditional probability would look at these two events in relationship with one another, such as the probability that it is both raining *and* you will need to go outside.

Multiplication Rule for Dependent Events
P (A) * P(B|A) = P(A∩B) = P (A and B) →

Rearranging: Gives Conditional Probability (B given A)
P(B|A) = P(A∩B) / P(A)

Conditional Probability = Joint Probability / Marginal Probability

Conditional Probability

Noni (Female) just got a subscription of HBO. What is the probability that her favorite show will be game of thrones. → Given a Female what is the probability that she likes GOT

TV Shows	Male	Female	Total
Game of Thrones	0.16	0.24/0.54 =0.444	0.4
West World	0.2	0.05/0.54 =0.093	0.25
Other	0.1	0.25/0.54 =0.463	0.35
Total	0.46	0.54	1

- **Conditional Probability**
  - $P(A|B) = P(A \cap B) / P(B)$ where $P(B) \neq 0$
  - $P(B|A) = P(A \cap B) / P(A)$ where $P(A) \neq 0$
  - P(GOT | Female) = GOT given Female
  - P(GOT ∩ Female) / P(Female)
  - 0.24/0.54 = 0.444

- **Comparing Conditional Probability and Simple Probability** tells us that Female like Game of thrones slightly more than general population. (0.444>0.4)
- Female like less the west world from the general population. (0.093<0.25)
- Female like other more than the general population. (0.463>0.35)

2

Bayes Theorem

Bayes Theorem

Prior Probabilities → New Information → Application of Bayes' Theorem → Posterior Probabilities

Bayes' theorem is a formula that describes how to update the probabilities of hypotheses when given evidence. It follows simply from the axioms of conditional probability, but can be used to powerfully reason about a wide range of problems involving belief updates.

Given a hypothesis H and evidence E, Bayes' theorem states that the relationship between the probability of the hypothesis before getting the evidence $P(H)$ and the probability of the hypothesis after getting the evidence $P(H \mid E)$ is

$$P(H \mid E) = \frac{P(E \mid H)}{P(E)} P(H).$$

- Where
  - P(H|E) : Posterior Probability
  - P(H) : Prior Probability
  - P(E) : Marginal Probability
  - P(E|H) : Conditional Probability
  - P(E|H) * P(H) : Joint Probability

Probability Distribution is a list of the possible outcomes wit corresponding probabilities

Lets take a Example

Two Dice are thrown simultaneously.

Q1) What is the probable outcomes. Draw the sample space of all the probable outcomes.
Q2) What are the probabilities of Sum of Two Dice being 0,1,2,3,4,5,6,7,8,9,10,11 and 12.

Probability Distribution

Plotting Probability Distribution

Dice Sum (Random Variable)	Probability (for the Sum of two Dice)	Probability in Decimals
2	1/36	0.027
3	2/36	0.055
4	3/36	0.083
5	4/36	0.111
6	5/36	0.138
7	6/36	0.166
8	5/36	0.194
9	4/36	0.111
10	3/36	0.083
11	2/36	0.055
12	1/36	0.027

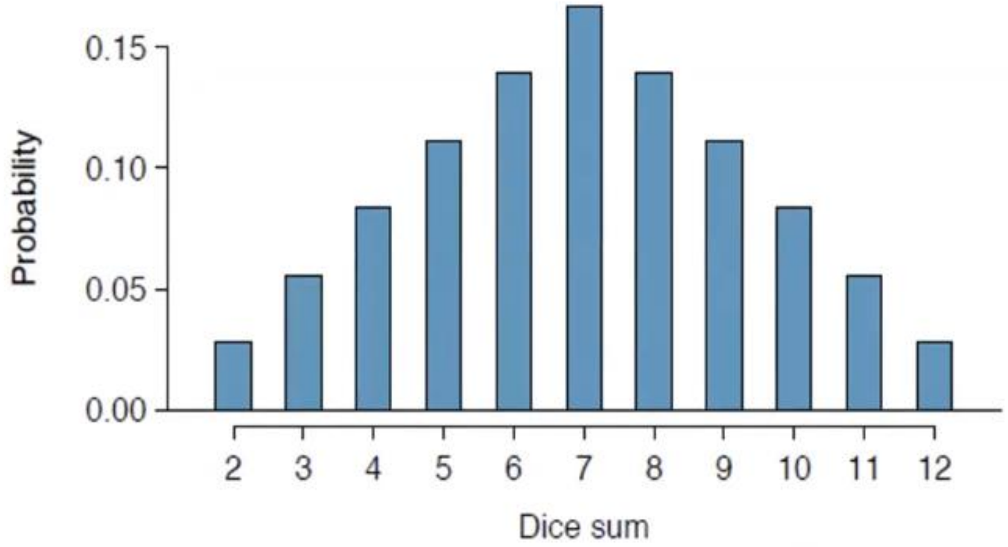


Figure 2.8: The probability distribution of the sum of two dice.

Here Dice Sum is a Random variable

Types of Probability Distribution

PROBABLITY DISTRIBUTION

Discrete Probability Distribution

Also known as Probability Mass Function

Example: No of complaints per day, Whether piece is defective or not

- Bernoulli Distribution
- Binomial Distribution
- Poisson Distribution

Distribution

Also known as Probability Density Function

Example: Time required to complete a Task, Temperature of a Solution, Height in inches, Weight in Kgs

- Normal Distribution
- Standard Normal Distribution

Bernoulli Distribution

- If X is a random variable that takes value 1 with a probability of success p and 0 with probability 1-p, then X is a Bernoulli random variable with mean and standard deviation.
- Mean for Bernoulli Distribution (Mu) = p
- Standard Deviation for Bernoulli Distribution (Sigma) = (p*1-p)^1/2

Bernoulli Random Variable	Probability
X	P(X)
1	P (Probability of success)
0	1-p (Probability of Failure)

Binomial Distribution- In the real world we often make several trials one to achieve on or more successes e.g. Suppose Toss of a coin is multiple times.

Conditions that need to be satisfied for a binomial experiment:

- Process consists of n trials
- Process has two outcomes Success or Failure (mutually exclusive)
- Probability of success denoted by p, does not change from trial to trial and probability of failure is 1-p and is also fixed from trial to trial
- Trials are independent

Binomial Distribution Formula

Let us consider cases where there are *n* number of Bernoulli trials.

Binomial Distribution

If $X \sim B(n, p)$, then

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{(n-x)} \qquad x = 0, 1, 2, \ldots, n$$

$$E(X) = np$$

$$V(X) = np(1 - p)$$

For example, if $n = 5$ and $p = 0.6$, then

$$P(X = 3) = 10 * 0.6^3 * 0.4^2 = 0.3456$$

$$E(X) = 5 * 0.6 = 3$$

$$V(X) = 5 * 0.6 * 0.4 = 1.2$$

- Where
  - P(X) = binomial probability
  - n=no of trials

    x = total number of "successes"

    p = probability of a success
  - q=1-p (probability of failure)

  - Mean = n* p
  - Variance = n*p*(1-p) = n* p*q

Examples Binomial Distribution

Example 1: A salesperson finds that, in the long run, two out of three sales calls are successful. Twelve Sales calls are to be made; let X be the number of concluded sales. Is X a binomial random variable? Explain.

Example 2: A new treatment for baldness is known to be effective in 70% of the cases treated. Four bald members of the same family are treated; Let X be the number of successfully treated members of the family. Is X a binomial Random Variable? Explain

Example 3: A coin is tossed 10 times. What is the probability of getting exactly 6 heads? Make its Probability Distribution table?

Example 4: Eight Coins are tossed simultaneously. Find the probability to get a)Exactly 6 heads b)up-to 6 heads c) more than 6 heads

Example 5: 10 Coins are tossed simultaneously. Find the probability of getting a)Exactly 7 heads b) up-to 6 heads c) more than 7 heads

Poisson Distribution: In general if we count the number of times a rare event occurs during a fixed interval, then that number would follow a Poisson Distribution

- Poisson Distribution focusses on the no of discrete events or occurrences over a specified interval or continuum (time, length, distance). It represents probability of k arrivals in time t
  - Some examples of Poisson Variables are
    - No of people in the retail counter or bank queue or Visa counter at customs
    - No of car arriving at a toll both.
    - The number of blinds born in a town in a particular year.
    - Number of mistakes committed in a typed page.
    - The number of students scoring very high marks in all subjects.
    - The number of plane accidents in a particular week.
    - The number of defective screws in a box of 100, manufactured by a reputed company.
    - Number of suicides reported in a particular day
- **Here we have to answer the Questions as follows:**
  - What is the probability that exactly 7 customers enter your line between 4:30 pm and 4:45 pm
  - What is the probability that more than 10 people arrive in your check line

Conditions for Poisson Distribution

- Discrete Outcomes
  - No of occurrences in each interval can range from 0 to infinity
  - Describes the distribution of infrequent or rare events
  - Each event is independent of other events
  - Expected no of occurrences are assumed to be constant throughout out the experiment i.e.
  - *λ (Mean)* = no of occurrences / specified interval = 10 customers / per 15 min

Example: Poisson Distribution

- Example 1:No of typographical errors in a book is a Poisson distribution with a mean of 1.5 per 100 pages Suppose 100 pages of book are randomly selected, What is the probability that there are no typos?

  Example 2: Suppose on an average 1 house in 1000 in a certain district has a fire during a year. If there are 2000 houses in that district, what is the probability that exactly 5 houses will have a fire during the year? (given that e^-2 =0.13534)

  Example 3: If 2% of electric bulbs by a certain company are defective. Find the probability that in a sample of 200 bulbs a) less than 2 bulbs b) more than 3 bulbs are defective (e^-4 = 0.0183)

Continuous Probability Distribution: The probabilities associated with continuous random variable X are determined by the probability densi function of the random variable.

- Conditions for Continuous Probability Distribution
- The function, denoted $f(x)$, has the following properties.
  - $f(x) = 0$ for all x. (Point Probability is zero)
  - The probability that X will be between two numbers a and b is equal to the area under $f(x)$ between a and b.
  - The total area under the entire curve of $f(x)$ is equal to 1.00.

- Continuous Probability Distribution is of two types
  - Normal Probability Distribution
  - Standard Normal Probability Distribution

- Continuous Random variable: A continuous random variable is a random variable that can take on any value in an interval of numbers

Normal Distribution

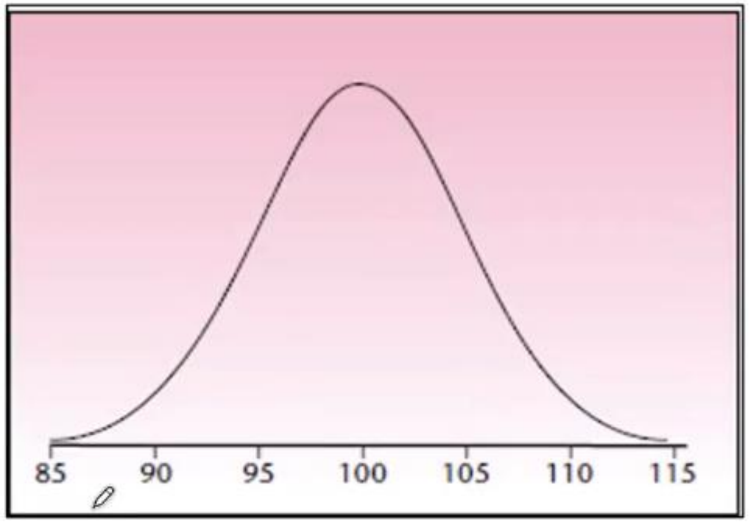
- The **normal distribution** is a continuous distribution.
- For **example**, heights, blood pressure, measurement error, and IQ scores follow the **normal distribution**. It is also known as the Gaussian **distribution** and the bell **curve**.
  - Example :Marks of students in a class follow normal distribution, Heights of students in a class
- For a normal distribution with mean and standard deviation , the probability density function *f* (*x*) is given by the complicated formula

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < +\infty \qquad (4\text{–}1)$$

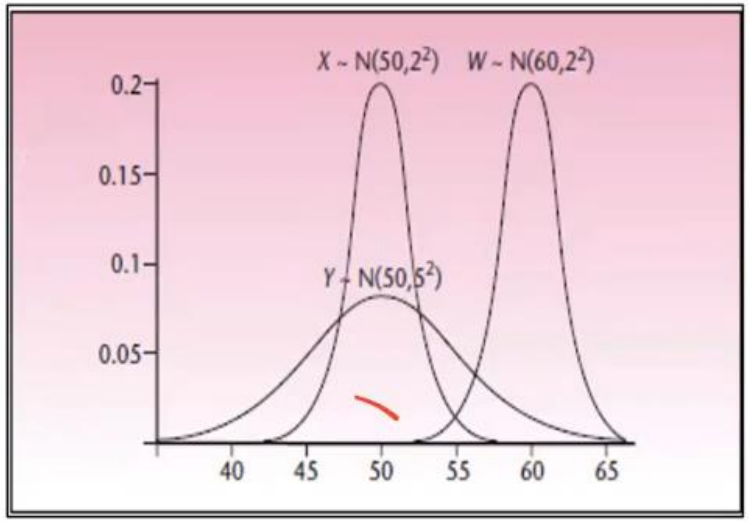
- In equation *e* is the natural base logarithm, equal to 2.71828 . . .

Normal Distribution

- Normal distribution is symmetric about its mean.
- Mean=Median=Mode



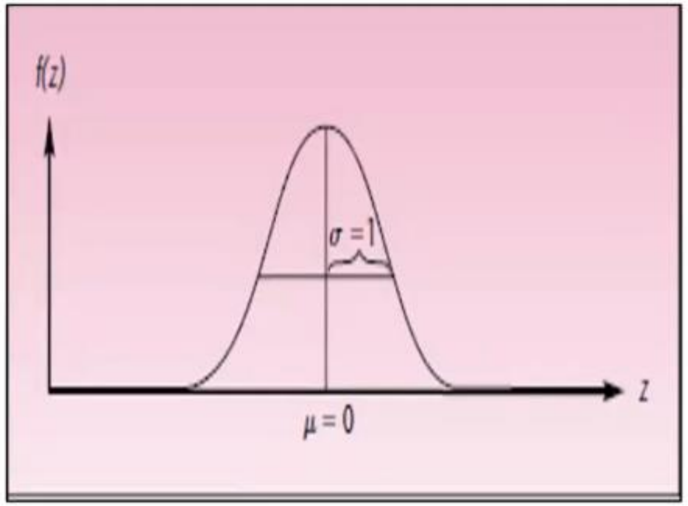
Mean = 100, SD =5



Mean = 50, 60; SD= 2 & 5

Standard Normal Distribution

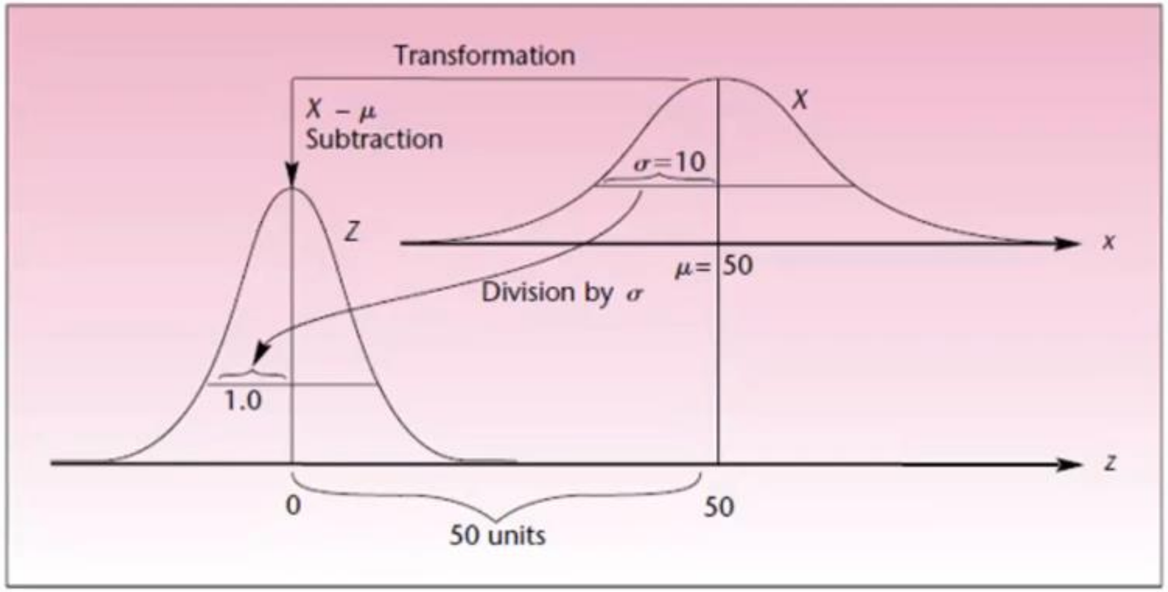
- We define the **standard normal random variable *Z*** as the normal random variable with mean 0 and standard deviation



- Notation to express Standard Normal Distribution

$$Z \sim N(0, 1^2) \quad (4\text{–}3)$$

Transformation of Normal Distribution to Standard Normal Distribution



Transforming a Normal Random variable with mean 50 and standard deviation 10 into Standard Normal Random Variable

The transformation of X to Z:

$$Z = \frac{X - \mu}{\sigma}$$