Interpolation

Computer Oriented Numerical and Statistical Methods

Outline

- Polynomial Interpolation
- Difference Tables
- Netwon’s Forward and Backward Interpolation Formula
- Lagrange's Formula
- Divided Difference Formula
- Inverse Interpolation

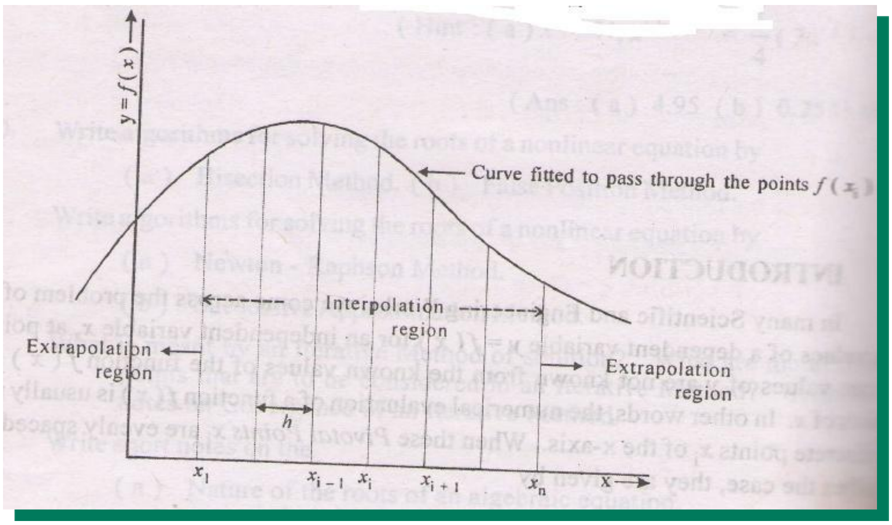
Introduction

- Suppose x and y are two variables and their relation can be expressed as $y = f(x)$; $x_1 \le x \le x_n$. Then we say that x is an independent variable and y is a dependent variable.
- When the form of f(x) is known, then the value of y can be computed directly corresponding to any value of x in the range $x_1 \le x \le x_n$.
- However, if the form of f(x) is not known, and only a set of values $(x_1,y_1), (x_2,y_2), \ldots, (x_n,y_n)$ satisfying the relation $y = f(x)$ are known, then the process of estimating the value of independent variable y for a given value of x in the range $x_1 \le x \le x_n$ is known as interpolation.

Introduction

- However, if we move in opposite direction i.e. estimate the value of dependent variable x for a given value of independent variable y, the process is known as inverse interpolation.
- The process of estimating the value of independent variable y for a given value of x outside the range $x_1 \le x \le x_n$ is known as extrapolation.

Introduction



Methods Of Interpolation

- The decision of using a particular method depends in tabulation of the functions.
- The tabulated points (x_i, y_i) i = 1,2,…,n of function y = f(x), can be equally spaced or unequal spaced.
- Methods for equally spaced functions
  - Netwon's forward interpolation formula
  - Netwon's backward interpolation formula

Methods Of Interpolation

- Methods for unequally spaced functions
  - Netwon’s divided difference interpolation formula
  - Lagrangian interpolation (Lagrange’s)
  - These methods also works well for equally spaced function.

Finite Differences

- The finite differences are either difference between the values of the function or the differences between the past differences.
- There are 3 types of differences
  - Forward Differences
  - Backward Differences
  - Divided Differences

Forward Differences

- If $y_1, y_2, y_3, \ldots, y_n$ denotes the values of the function of type $y = f(x)$ at $x = x_1, x_2, \ldots, x_n$ then $y_2 - y_1, y_3 - y_2, y_4 - y_3, \ldots, y_n - y_{n-1}$ are called the forward differences of y
- These differences are denoted as $\Delta y_1, \Delta y_2, \Delta y_3, \ldots, \Delta y_{n-1}$
- $\therefore\ \Delta y_1 = y_2 - y_1,\ \Delta y_2 = y_3 - y_2, \ldots, \Delta y_{n-1} = y_n - y_{n-1}$ where Δ is called forward difference operator and $\Delta y_1, \Delta y_2, \Delta y_3, \ldots, \Delta y_{n-1}$ are called first order forward differences.

Forward Differences

- The differences of the first order forward differences are called second order forward differences and are denoted as $\Delta^2 y_1, \Delta^2 y_2, \Delta^2 y_3, \ldots, \Delta^2 y_{n-1}$
- $\therefore\ \Delta^2 y_1 = \Delta y_2 - \Delta y_1 = y_3 - 2y_2 + y_1$
- $\Delta^2 y_2 = \Delta y_3 - \Delta y_2 = y_4 - 2y_3 + y_2$
- In the similar manner, the third order forward differences are
- $\Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1 = y_4 - 3y_3 + 3y_2 - y_1$
- $\Delta^3 y_2 = \Delta^2 y_3 - \Delta^2 y_2 = y_5 - 3y_4 + 3y_3 - y_2$
- In general, the first order forward differences at the i^{th} point is $\Delta y_i = y_{i+1} - y_i$ and the j^{th} order forward differences at the i^{th} point is $\Delta^j y_i = \Delta^{j-1} y_{i+1} - \Delta^{j-1} y_i$

Backward Differences

- If $y_1, y_2, y_3, \ldots, y_n$ denotes the values of the function of type $y = f(x)$ at $x = x_1, x_2, \ldots, x_n$ then $y_2 - y_1, y_3 - y_2, y_4 - y_3, \ldots, y_n - y_{n-1}$ are called the backward differences of y
- These differences are denoted as $\nabla y_2, \nabla y_3, \ldots, \nabla y_n$
- $\therefore\ \nabla y_2 = y_2 - y_1,\ \nabla y_3 = y_3 - y_2, \ldots, \nabla y_n = y_n - y_{n-1}$ where ∇ is called backward difference operator and $\nabla y_2, \nabla y_3, \ldots, \nabla y_n$ are called first order backward differences.

Backward Differences

- The differences of the first order backward differences are called second order backward differences and are denoted as $\nabla^2 y_3, \nabla^2 y_4, \ldots$ etc.
- $\therefore\ \nabla^2 y_3 = \nabla y_3 - \nabla y_2 = y_3 - 2y_2 + y_1$
- $\nabla^2 y_4 = \nabla y_4 - \nabla y_3 = y_4 - 2y_3 + y_2$
- In the similar manner, the third order backward differences are
- $\nabla^3 y_4 = y_4 - 3y_3 + 3y_2 - y_1$
- $\nabla^3 y_5 = y_5 - 3y_4 + 3y_3 - y_2$
- In general, the first order backward differences at the i^{th} point is $\nabla y_i = y_i - y_{i-1}$ and the j^{th} order backward differences at the i^{th} point is $\nabla^j y_i = \nabla^{j-1} y_i - \nabla^{j-1} y_{i-1}$

Divided Differences

- If $y_1, y_2, y_3, \ldots, y_n$ denotes the values of the function of type $y = f(x)$ at $x = x_1, x_2, \ldots, x_n$ then

$$\frac{y_2 - y_1}{x_2 - x_1}, \quad \frac{y_3 - y_2}{x_3 - x_2}, \quad \frac{y_4 - y_3}{x_4 - x_3}, \ldots, \quad \frac{y_n - y_{n-1}}{x_n - x_{n-1}}$$

- are called the divided differences of y and are denoted as $\Delta_d y_1, \Delta_d y_2, \Delta_d y_3, \ldots, \Delta_d y_{n-1}$
- $\therefore\ \Delta_d y_1 = (y_2 - y_1) / (x_2 - x_1) = [x_1, x_2]$
- $\Delta_d y_2 = (y_3 - y_2) / (x_3 - x_2) = [x_2, x_3]$
- $\Delta_d y_{n-1} = (y_n - y_{n-1}) / (x_n - x_{n-1}) = [x_{n-1}, x_n]$
- Where Δ_d is called divide difference operator and $\Delta_d y_1, \Delta_d y_2, \ldots, \Delta_d y_n$ are called first order divided differences.

Divided Differences

- The differences of the first order divided differences are called second order divided differences and are denoted as $\Delta^2{}_d y_1, \Delta^2{}_d y_2, \ldots, \Delta^2{}_d y_n$ etc.
- $\therefore\ \Delta^2{}_d y_1 = (\Delta_d y_2 - \Delta_d y_1) / (x_3 - x_1)$
- $\Delta^2{}_d y_2 = (\Delta_d y_3 - \Delta_d y_2) / (x_4 - x_2)$
- In the similar manner, the third order forward differences are
- $\Delta_d^3 y_1 = (\Delta_d^2 y_2 - \Delta_d^2 y_1) / (x_4 - x_1)$
- $\Delta_d^3 y_2 = (\Delta_d^2 y_3 - \Delta_d^2 y_2) / (x_5 - x_2)$
- In general, the first order divided differences at the i^{th} point is $\Delta_d y_i = (y_{i+1} - y_i) / (x_{i+1} - x_i)$ and the j^{th} order forward differences at the i^{th} point is $\Delta_d^j y_1 = (\Delta_d^{j-1} y_{i+1} - \Delta_d^{j-1} y_i) / (x_{i+1} - x_i)$

Differences Tables

- A difference table is a table that lists the differences of the function values and the differences of differences in succession.

Forward Difference Table

- Let us consider the values y_1, y_2, y_3, y_4 of the function type y = f(x) tabulated at equally spaced points x_1, x_2, x_3, x_4. The forward difference table along with tabulated points will look like

Forward Difference Table					
i	x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$
1	x_1	y_1	$\Delta y_1 = y_2 - y_1$	$\Delta^2 y_1 = \Delta y_2 - \Delta y_1$	$\Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1$
2	x_2	y_2	$\Delta y_2 = y_3 - y_2$	$\Delta^2 y_2 = \Delta y_3 - \Delta y_2$	
3	x_3	y_3	$\Delta y_3 = y_4 - y_3$		
4	x_4	y_4			

Forward Difference Table

- The forward difference table for function tabulated at n equally spaced points can be represented by a matrix of size (n-1) * (n-1) where j^{th} order frequency at the i^{th} point ($\Delta^j y_i$) is represented by the element d_{ij} of matrix D.
- Note that only the elements in the column 1 to (n – i), for rows i = 1, 2, …., n-1 are of interest.

Backward Difference Table

- Let us consider the values y_1, y_2, y_3, y_4 of the function type y = f(x) tabulated at equally spaced points x_1, x_2, x_3, x_4. The Backward difference table along with tabulated points will look like

Backward Difference Table					
i	x_i	y_i	∇y_i	$\nabla^2 y_i$	$\nabla^3 y_i$
1	x_1	y_1			
2	x_2	y_2	$\nabla y_2 = y_2 - y_1$		
3	x_3	y_3	$\nabla y_3 = y_3 - y_2$	$\nabla^2 y_3 = \nabla y_3 - \nabla y_2$	
4	x_4	y_4	$\nabla y_4 = y_4 - y_3$	$\nabla^2 y_4 = \nabla y_4 - \nabla y_3$	$\nabla^3 y_4 = \nabla^2 y_4 - \nabla^2 y_3$

Divided Difference Table

- Let us consider the values y_1, y_2, y_3, y_4 of the function type $y = f(x)$ tabulated at points x_1, x_2, x_3, x_4 not necessarily equally spaced . The divide difference table along with tabulated points will look like

Divided Difference Table

Divided Difference Table					
i	x_i	y_i	$\Delta_d y_i$	$\Delta^2 y_i$	$\Delta^3 y_i$
1	x_1	y_1	$\Delta_d y_1 = (y_2 - y_1) / (x_2 - x_1)$	$\Delta_d^2 y_1 = (\Delta_d y_2 - \Delta_d y_1) / (x_3 - x_1)$	$\Delta_d^3 y_1 = (\Delta_d^2 y_2 - \Delta_d^2 y_1) / (x_4 - x_1)$
2	x_2	y_2	$\Delta_d y_2 = (y_3 - y_2) / (x_3 - x_2)$	$\Delta_d y_2 = (\Delta_d y_3 - \Delta_d y_2) / (x_4 - x_2)$	
3	x_3	y_3	$\Delta_d y_3 = (y_4 - y_3) / (x_4 - x_3)$		
4	x_4	y_4			

Netwon's Methods Of Interpolation

- It is divided into following methods depending on the type of differences being used.

1. Netwon's Forward Difference Interpolation Formula
2. Netwon's Backward Difference Interpolation Formula
3. Netwon's Divided Difference Interpolation Formula

- If the function is tabulated at equal intervals, then we can use either Netwon's Forward Difference Interpolation Formula or Netwon's Backward Difference Interpolation Formula.

Netwon's Forward Difference Interpolation Formula

- Let us assume that the function y(x) is tabulated at (n+1) equally spaced (interval size h) points.
- To derive the formula for netwon's forward difference interpolation assume a polynomial of type
  $y(x) = a_1 + a_2(x-x_1) + a_3(x-x_1)(x-x_2) + \ldots + a_{n+1}(x-x_1)(x-x_2) \ldots (x-x_n)$ ……………………..[a]
- It is the n^{th} order polynomial in x.
- It is an interpolation polynomial for the table values x_i and y_i then the polynomial must pass through all points.

∀ $\therefore$ we can obtain y_i by substituting the corresponding x_i for x

Netwon's Forward Difference Interpolation Formula

- At $x = x_1$ $y_1 = a_1$
- $x = x_2$ $y_2 = a_1 + a_2(x_2-x_1)$
- $x = x_3$ $y_3 = a_1 + a_2(x_3-x_1) + a_3(x_3-x_1)(x_3-x_2)$
- And so on.
- Since x_i's are equally spaced therefore we can write $x_{i+1} - x_i = h$ and $x_{i+m} - x_i = mh$ then we can express the function values y_i's in terms of intervals values as

  $y_1 = a_1$

  $y_2 = a_1 + a_2h$

  $y_3 = a_1 + a_2(2h) + a_3(2h)(h)$

  …….

$y_{n+1} = a_1 + a_2(nh) + a_3(nh)((n-1)h)+\ldots.+a_{n+1}(nh)((n-1)h)\ldots(h)$

Netwon's Forward Difference Interpolation Formula

- Now for a_1, a_2, a_3, a_{n+1} we get

$$a_1 = y_1$$

$$a_2 = \frac{y_2 - a_1}{h} = \frac{y_2 - y_1}{h}$$

$$a_3 = \frac{y_3 - 2y_2 + y_1}{2!h^2}$$

$$a_{n+1} = \frac{y_{n+1} - ny_n + \ldots. + y_1}{n!h^n}$$

Netwon's Forward Difference Interpolation Formula

- Using forward difference table we get

$$a_1 = y_1$$

$$a_2 = \frac{\Delta y_1}{h}$$

$$a_3 = \frac{\Delta^2 y_1}{2!h^2}$$

.

.

.

$$a_{n+1} = \frac{\Delta^n y_1}{n!h^n}$$

Netwon's Forward Difference Interpolation Formula

- Substituting these values of a_1, a_2, a_3, a_{n+1} in equation [a] we get

$$y(x) = y_1 + \frac{\Delta y_1}{h}(x - x_1) +$$

$$\frac{\Delta^2 y_1}{2!h^2}(x - x_1)(x - x_2) +$$

$$\ldots\ldots + \frac{\Delta^n y_1}{n!h^n}(x - x_1)(x - x_2)\ldots(x - x_n)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [b]$$

Netwon's Forward Difference Interpolation Formula

- If we use the relation $(x-x_1) / h = u$ then $x - x_1 = hu$
  $x – x_2 = x – (x_1 + h) = h(u-1)$
  $x – x_3 = x – (x_2 + h) = h(u-2)$ …..
  $x – x_n = h(u-(n-1))$
- Substituting these values $(x-x_1)$, $(x-x_2)$, …., $(x – x_n)$ in equation [b] we get

Netwon's Forward Difference Interpolation Formula

$$y(x) = y_1 + \frac{\Delta y_1}{h} hu +$$

$$\frac{\Delta^2 y_1}{2!h^2} hu * h(u-1) +$$

$$\ldots + \frac{\Delta^n y_1}{n!h^n} hu * h(u-1) * \ldots * h(u-(n-1))$$

$$\ldots\ldots\ldots\ldots\ldots\ldots [c]$$

Netwon's Forward Difference Interpolation Formula

- Simplifying we get

$$y(x) = y_1 + \Delta y_1 u +$$

$$\frac{\Delta^2 y_1}{2!} u(u-1) +$$

$$\ldots + \frac{\Delta^n y_1}{n!} u(u-1) * \ldots * (u-(n-1))$$

$$\ldots\ldots\ldots\ldots\ldots [d]$$

- The above derivation assumes that $x_1 < x < x_2$
- This polynomial is called Netwon's Forward Difference Interpolation Formula

Netwon's Backward Difference Interpolation Formula

- Let us assume that the function y(x) is tabulated at (n+1) equally spaced (interval size h) points.
- To derive the formula for netwon's backward difference interpolation assume a polynomial of type $y(x) = a_1 + a_2(x-x_n) + a_3(x-x_n)(x-x_{n-1}) + \ldots + a_{n+1}(x-x_n)(x-x_{n-1}) \ldots (x-x_1)$ ……………………..[a]
- It is the n^{th} order polynomial in x.
- It is an interpolation polynomial for the table values x_i and y_i then the polynomial must pass through all points.

∀ ∴ we can obtain y_i by substituting the corresponding x_i for x

Netwon's Backward Difference Interpolation Formula

- At $x = x_n$ $y_n = a_1$
- $x = x_{n-1}$ $y_{n-1} = a_1 + a_2(x_{n-1} - x_n)$
- $x = x_{n-2}$ $y_{n-2} = a_1 + a_2(x_{n-2} - x_n) + a_3(x_{n-2}-x_n)(x_{n-2}-x_{n-1})$
- And so on.
- Since x_i's are equally spaced therefore we can write $x_i - x_{i-1} = h$ and $x_{i-1} - x_i = -h$ and $x_{i-m} - x_i = -mh$ then we can express the function values y_i's in terms of intervals values as

  $y_n = a_1$

  $y_{n-1} = a_1 + a_2(-h)$

  $y_{n-2} = a_1 + a_2(-2h) + a_3(-2h)(-h)$

  …….

$y_1 = a_1 + a_2(-nh) + a_3(-nh)(-(n-1)h)+\ldots+a_{n+1}(-nh)(-(n-1)h)\ldots(-h)$

Netwon's Backward Difference Interpolation Formula

- Now for a_1, a_2, a_3, a_{n+1} we get

$$a_1 = y_n$$

$$a_2 = \frac{y_{n-1} - a_1}{h} = \frac{y_n - y_{n-1}}{h}$$

$$a_3 = \frac{y_n - 2y_{n-1} + yn-2}{2!h^2}$$

$$a_{n+1} = \frac{y1 - ny_2 + \ldots + y_n}{n!h^n}$$

Netwon's Backward Difference Interpolation Formula

- Using backward difference table we get

$$a_1 = y_n$$

$$a_2 = \frac{\nabla y_n}{h}$$

$$a_3 = \frac{\nabla^2 y_n}{2!h^2}$$

.
.
.

$$a_{n+1} = \frac{\nabla^n y_n}{n!h^n}$$

Netwon's Backward Difference Interpolation Formula

- Substituting these values of a_1, a_2, a_3, a_{n+1} in equation [a] we get

$$y(x) = y_n + \frac{\nabla y_n}{h}(x - x_n) + \frac{\nabla^2 y_n}{2!h^2}(x - x_n)(x - x_{n-1}) + \dots + \frac{\nabla^n y_n}{n!h^n}(x - x_n)(x - x_{n-1})\dots(x - x_1) \quad \dots\dots\dots\dots\dots\dots [b]$$

Netwon’s Backward Difference Interpolation Formula

- If we use the relation $(x-x_n) / h = u$ then $x - x_n = hu$
  $x – x_{n-1} = x – (x_n - h) = h(u+1)$
  $x – x_{n-2} = x – (x_{n-1} - h) = h(u+2)$ …..
  $x – x_1 = h(u+(n-1))$
- Substituting these values $(x-x_n)$, $(x-x_{n-1})$, …., $(x – x_1)$ in equation [b] we get

Netwon’s Backward Difference Interpolation Formula

$$y(x) = y_n + \frac{\nabla y_n}{h} hu +$$

$$\frac{\nabla^2 y_n}{2!h^2} hu * h(u+1) +$$

$$\ldots + \frac{\nabla^n y_n}{n!h^n} hu * h(u+1) * \ldots * h(u+(n-1))$$

$$\ldots\ldots\ldots\ldots [c]$$

Netwon's Backward Difference Interpolation Formula

- Simplifying we get

$$y(x) = y_n + \nabla y_n u + \frac{\nabla^2 y_n}{2!} u(u+1) + \ldots + \frac{\nabla^n y_n}{n!} u(u+1) * \ldots * (u+(n-1)) \quad \ldots\ldots\ldots\ldots\ldots\ldots [d]$$

- The above derivation assumes that $x_{n-1} < x < x_n$
- This polynomial is called Netwon's Backward Difference Interpolation Formula

Netwon's Divided Difference Interpolation Formula

- Let us assume that the function y(x) is tabulated at (n+1) equally spaced (interval size h) points.
- To derive the formula for netwon's divided difference interpolation assume a polynomial of type
  $y(x) = a_1 + a_2(x-x_1) + a_3(x-x_1)(x-x_2) + \ldots + a_{n+1}(x-x_1)(x-x_2) \ldots (x-x_n)$ ……………………..[a]
- It is the n^{th} order polynomial in x.
- It is an interpolation polynomial for the table values x_i and y_i then the polynomial must pass through all points.

∀ ∴ we can obtain y_i by substituting the corresponding x_i for x

Netwon's Divided Difference Interpolation Formula

- At $x = x_1$ $y_1 = a_1$
- $x = x_2$ $y_2 = a_1 + a_2(x_2-x_1)$
- $x = x_3$ $y_3 = a_1 + a_2(x_3-x_1) + a_3(x_3-x_1)(x_3-x_2)$
- And so on.
- Solving for a_1, a_2, a_3, a_{n+1} we get

Netwon's Divided Difference Interpolation Formula

$$a_1 = y_1$$

$$a_2 = \frac{y_2 - a_1}{x_2 - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

$$a_3 = \frac{1}{(x_3 - x_2)} * \left[\frac{(y_3 - y_1)}{(x_3 - x_1)} - \frac{(y_2 - y_1)}{(x_2 - x_1)}\right]$$

$$and so on$$

Netwon’s Divided Difference Interpolation Formula

- Using divided difference table we get

$$a_1 = y_1$$

$$a_2 = \Delta_d y_1$$

$$a_3 = \Delta_d^2 y_1$$

$$\vdots$$

$$a_{n+1} = \Delta_d^n y_1$$

Netwon’s Divided Difference Interpolation Formula

- Substituting these values of a_1, a_2, a_3, a_{n+1} in equation [a] we get

$$y(x) = y_1 + \Delta_d y_1 (x - x_1) +$$

$$\Delta_d^2 y_1 (x - x_1)(x - x_2) +$$

$$\dots + \Delta_d^n y_1 (x - x_1)(x - x_2)\dots(x - x_n)$$

$$\dots\dots\dots\dots\dots\dots\dots [b]$$

- This polynomial is called Netwon’s Divided Difference Interpolation Formula

Lagrangian / Lagranges Interpolation Formula

- In order to derive a general formula for lagrangian interpolation, we consider a second order polynomial of type.
  $y(x) = a_1(x-x_2)(x-x_3) + a_2(x-x_1)(x-x_3) + a_3(x-x_1)(x-x_2)$ ……………………..[a]
  passing through the points (x_1,y_1), (x_2,y_2), and (x_3,y_3) where a_1, a_2, and a_3 are unknown constants whose values are determined as follows.
- At $x = x_1$ $y(x_1) = a_1(x_1 - x_2)(x_1-x_3)$

Lagrangian / Lagranges Interpolation Formula

$$At\ x = x_1 \quad y(x_1) = a_1(x_1 - x_2)(x_1 - x_3)$$

$$\Rightarrow a_1 = \frac{y_1}{(x_1-x_2)(x_1-x_3)}$$

$$At\ x = x_2 \quad y(x_2) = a_2(x_2 - x_1)(x_2 - x_3)$$

$$\Rightarrow a_2 = \frac{y_2}{(x_2-x_1)(x_2-x_3)}$$

$$At\ x = x_3 \quad y(x_3) = a_3(x_3 - x_1)(x_3 - x_2)$$

$$\Rightarrow a_3 = \frac{y_3}{(x_3-x_1)(x_3-x_2)}$$

Lagrangian / Lagranges Interpolation Formula

- Substituting these values of a_1 , a_2, a_3 in equation [a] we get

$$y(x) = y_1 * \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} +$$

$$y_2 * \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} +$$

$$y_3 * \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}$$

Lagrangian / Lagranges Interpolation Formula

$$f(x) = f(x_0) * \frac{(x-x_1)(x-x_2)\ldots(x-x_n)}{(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n)} +$$

$$f(x_1) * \frac{(x-x_0)(x-x_2)\ldots(x-x_n)}{(x_1-x_0)(x_1-x_2)\ldots(x_1-x_n)} + \ldots +$$

$$f(x_n) * \frac{(x-x_0)(x-x_1)\ldots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\ldots(x_n-x_{n-1})} +$$

- This polynomial is known as the Lagrange's polynomial

Minal Shah

Skewness

Computer Oriented Numerical and Statistical Methods

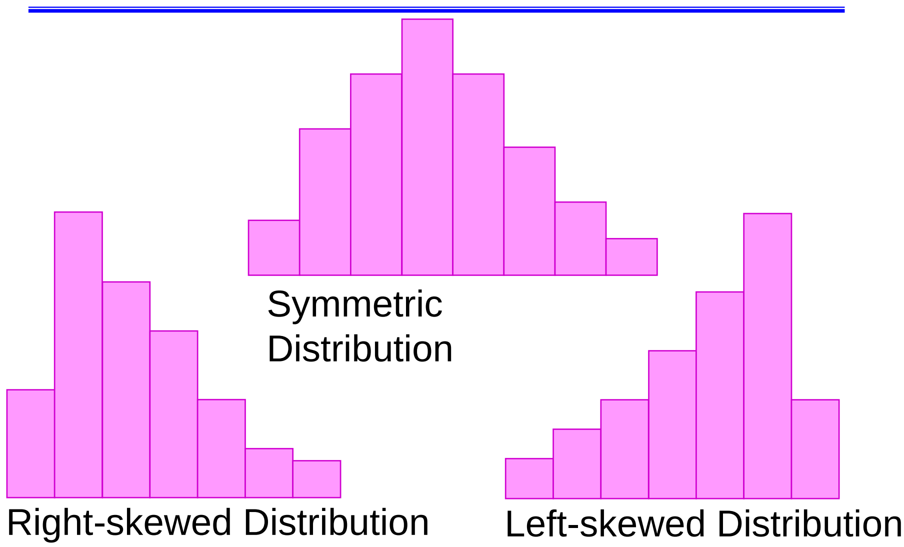
Outline

- Introduction
- Types of Skewness
- Measure of skewness.
- Karl Pearson's Measure
- Bowley's Measure

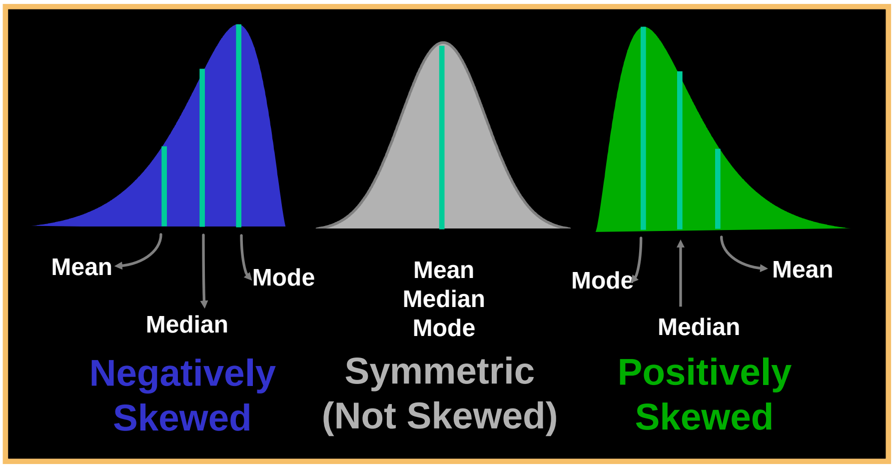
Introduction

- Measure of central tendency gives us an estimate of the representative value of a series, the measure of dispersion gives an indication of the extent to which the items cluster around or scatter away from the central value and the skewness is a measure that refers to the extent of symmetry or asymmetry in a distribution.
- It describes the shape of a distribution,

Skewness In The Form Of Histogram

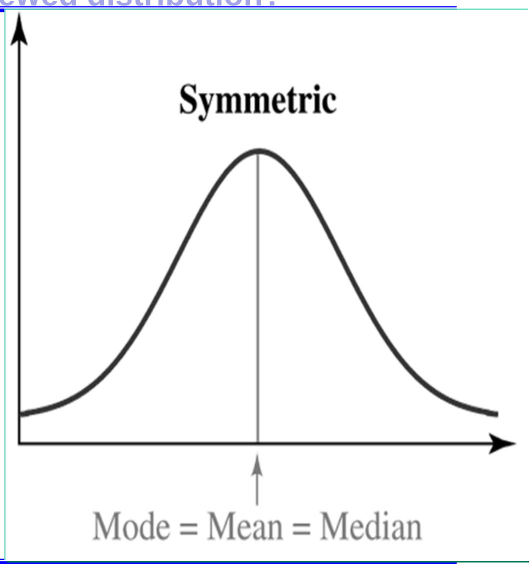


Skewness



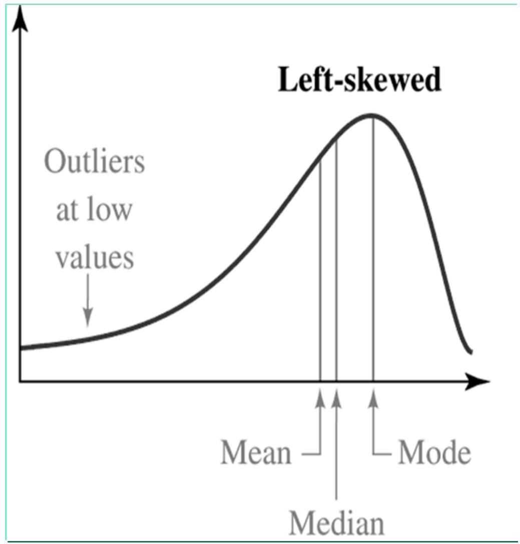
What is the relationship between mean, median and mode of a skewed distribution?

- Find the mean median and mode of: 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 6, 6, 7
- Mean is 4.
- Median is 4.
- Mode is 4.



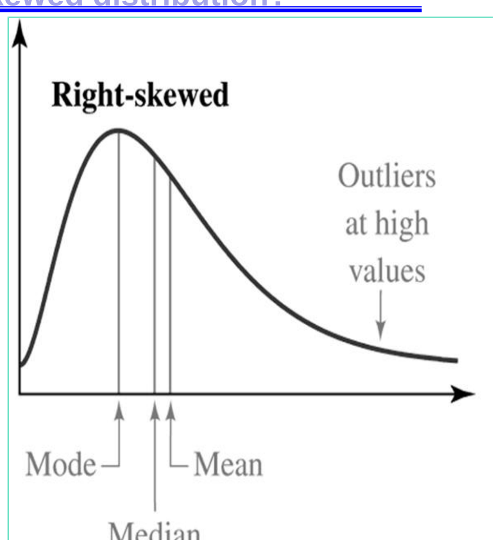
What is the relationship between mean, median and mode of a left-skewed distribution?

- Find the mean, median and mode of:
  0, 5, 10, 20, 40, 45, 45, 50, 50, 50, 60, 60, 60, 60, 60, 60, 70, 70, 70, 70, 70, 70, 70, 70
- The mean is 51.5.
- The median is 60.
- The mode is 70.



What is the relationship between mean, median and mode of a right-skewed distribution?

- Find the mean, median, and mode of:
  20, 20, 20, 20, 20, 20, 20, 20, 30, 30, 30, 30, 30, 30, 45, 45, 45, 50, 50, 60, 70, 90
- The mean is 36.1.
- The median is 30.
- The mode is 20.



Various Distribution And The Position Of Average

Size	Frequency (c)	Frequency (b)	Frequency (a)
0 – 5	10	10	10
5 – 10	90	30	20
10 – 15	50	50	30
15 – 20	40	70	40
20 – 25	30	50	50
25 – 30	20	30	90
30 – 35	10	10	10
skewness	positive	symmetry	Negative
average	$\bar{x} > M_d > M_0$	$\bar{x} = M_d = M_0$	$\bar{x} < M_d < M_0$
Quartiles	$Q_3 - M_d > M_d - Q_1$	$Q_3 - M_d = M_d - Q_1$	$Q_3 - M_d < M_d - Q_1$
curve	Skewed to the right	Normal	Skewed to the left

Objectives of Skewness

- It helps in finding out the nature and the degree of concentration whether it is in higher or the lower values.
- The empirical relations of mean and median and mode are based on a moderately skewed distribution. The measure of skewness will reveal to what extent such empirical relationship holds good.
- It helps in knowing if the distribution is normal. Many statistical measures, such as the error of the mean are based on the assumption of a normal distribution.

Measures of Skewness

- To find out the direction and the extent of asymmetry in a series statistical measures of skewness are employed.
- This measure can be absolute or relative.
- Absolute measure of skewness tell us the extent of asymmetry and whether it is positive or negative.
- The absolute skewness is based on the difference between mean and mode. Symbolically,

  absolute S_k = Mean – Mode.
- If the value of mean is greater than the mode, skewness will be positive.

Measures of Skewness

- If the value of mean is less than the mode, skewness will be negative.
- Absolute measure of skewness is not adequate because it cannot be used for comparison of skewness in two distributions if they are in different units, since difference between mean and mode will be in terms of units of distribution.
- For comparison purpose we use the relative measure of skewness known as coefficient of skewness.

Measures of Skewness

- There are four types of relative measure of skewness :
  1. The Karl Pearson's Coefficient of Skewness.
  2. The Bowley's Coefficient of Skewness.
  3. The Kelly's Coefficient of Skewness.
  4. Measure of Skewness based on Moments and Kurtosis.

Karl Pearson's Coefficient of Skewness

- Karl Pearson's Coefficient of Skewness or Pearsonian Coefficient of skewness is given by the formula:

$$S_k = \frac{\text{Mean} - \text{Mode}}{\text{Standard Deviation}}$$

- If in a particular frequency distribution, it is difficult to determine precisely the mode, or the mode is ill-defined, the coefficient of skewness can be determined by the following formula:

$$S_k = \frac{3(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

Karl Pearson’s Coefficient of Skewness

- Theoretically, skewness lies between the limits ± 3, but these limits are rarely attained in practice.

Bowley’s Coefficient of Skewness

- Bowley’s coefficient of skewness also known as Quartile coefficient of skewness and is especially useful :
  - When the mode is ill-defined and extreme observations are present in the data.
  - When the distribution has open-end classes or unequal class-interval.
- The quartile measure depends upon the fact that normally Q_3 and Q_1 are equidistance from the median, i.e. for symmetrical distribution $Q_3 - M_d = M_d - Q_1$.
- If a distribution is asymmetrical, then one quartile will be farther from the median than the other.

Bowley's Coefficient of Skewness

- In such a case skewness can be measured by the following formula given by Bowley :

  Skewness = $(Q_3 – M_d) – (M_d – Q_1)$

  Skewness = $Q_3 + Q_1 – 2M_d$
- If the first part is more than the second part, the skewness is positive and in the reverse situation it is negative.

Bowley's Coefficient of Skewness

- To make the measure a readily comparable, the coefficient of skewness is obtained by dividing it by quartile range viz $Q_3 – Q_1$

$$S_k = \frac{(Q_3 - M_d) - (M_d - Q_1)}{(Q_3 - M_d) + (M_d - Q_1)}$$

$$S_k = \frac{Q_3 + Q_1 - 2M_d}{Q_3 - Q_1}$$

Bowley's Coefficient of Skewness

- The range of variation under this method is ± 1.
- The main drawback of this measure is that it is based on the central 50% of the data and ignores the remaining 50% of the data towards the extremes.

Probability

Computer Oriented Numerical and Statistical Methods

Outline

- Introduction
- Meaning
- Basic Definitions
- Types
- Basic Probability Rules

Introduction

- So far we have studied the methods of collection, description and analysis of data.
- This does not cover the entire operational field of statistics.
- Modern statistics has also to provide for :
  1. Estimation of population ‘parameters’ on the basis of sample ‘statistics’
  2. Drawing inferences about the sample ‘statistics’ on the basis of population ‘parameters’
  3. Testing of hypothesis with regards to (1) and (2) above and

Introduction

4. Decision-making under risk and uncertainty be estimating the degree of risk and the likely effect on business objectives in terms of pay off and expected values of a decision.

- The theory of probability helps in all these areas.
- Probability helps a person to make ‘educated guesses’ on matters, where either full facts are not known or there is uncertainty about the outcome.
- The decision-makers always face some degree of risk while selecting a particular decision (course of action or strategy) to solve a decision problem.

Introduction

- It is because each strategy can lead to a number of different possible outcomes (or results).
- Thus it is necessary for the decision-makers to enhance their capability of grasping the probabilistic situation so as to gain a deeper understanding of the decision problem and base their decision on rational considerations.

Meaning

- Probability means the number of occasion that a particular event is likely to occur in a large population of events.
- The particular event may be expressed positively where the event is likely to happen or negatively where the event is not likely to happen.

Meaning

- Broadly, there are three possible states of expectations :
  - Certainty
  - Impossibility
  - Uncertainty
- The probability theory describe certainty by 1, impossibility by 0 and the various grades of uncertainties by coefficients ranging between 0 and 1

Probability in Everyday Life

- Possible, it will rain tonight.
- Probably you will catch the train.
- There is a high chance of your getting the job in August.
- This year's demand for the product is likely to exceed that of the last year's.

Definitions

- **Probability** : It means the number of occasion that a particular event is likely to occur in a large population of events.
  - The particular event may be expressed positively where the event is likely to happen or negatively where the event is not likely to happen.

Definitions

- **Random Experiment or Trial**: An experiment is said to be a random experiment (or trial or act or operation or process), if it's out-come can't be predicted with certainty.
- Example :
  - If a coin is tossed, we can't say, whether head or tail will appear. So it is a random experiment.
  - Drawing a card from a pack of cards

Definitions

- **Sample Space (Possible outcomes)**: The set of all possible out-comes of an experiment is called the sample space. It is denoted by 'S' or 'U' and its number of elements are n(s).
- **Example:**
  - In throwing a dice, the number that appears at top is any one of 1,2,3,4,5,6. So here: S ={1,2,3,4,5,6} and n(s) = 6
  - In the case of a coin, S={Head,Tail} or {H,T} and n(s)=2.
- The elements of the sample space are called sample point or event point.

Definitions

- **Event**: Every subset of a sample space is an event. It is denoted by 'E'.
  - The empty set ∅ is called impossible event and the sample space U is called certain event.
  - Clearly E is a sub set of S
- **Example:**
  - In throwing a dice S={1,2,3,4,5,6}, the appearance of an event number will be the event E={2,4,6}.
- **Simple event** : An event, consisting of a single sample point is called a simple event.
- **Example**:
  - In throwing a dice, S={1,2,3,4,5,6}, so each of {1},{2},{3},{4},{5} and {6} are simple events.

Definitions

- **Compound event:** A subset of the sample space, which has more than one element is called a mixed event (compound event).
- Example:
  - In throwing a dice, the event of appearing of odd numbers is a compound event, because E={1,3,5} which has '3' elements.

Definitions

- **Equally likely events**: Events are said to be equally likely, if we have no reason to believe that one is more likely to occur than the other. OR The outcomes are said to be equally likely or equally probable if none of them is expected to occur in preference to other.
- Example:
  - When a dice is thrown, all the six faces {1,2,3,4,5,6} are equally likely to come up.

Definitions

- **Exhaustive events**: When every possible out come of an experiment is considered.
- **Example:**
  - A dice is thrown, cases 1,2,3,4,5,6 form an exhaustive set of events.
- **Collectively Exhaustive events**: The total number of possible outcomes of a random experiment is called the collective exhaustive events.
- **Example:**
  - A dice is thrown, cases 1,2,3,4,5,6 form an exhaustive set of events and number of cases is 6.
  - In toss of a single coin exhaustive number of cases is 2

Definitions

- **Complementary events**: The set of all elements of the sample space U except the elements of event A is called the complementary event A. It is denoted by A' or $\bar{A}$ it means that 'the event A does not occur.
- **Union events**: Let A and B be two events. The set of elements which are either in A or B is called the union event of event A and B. it is denoted by $A \cup B$. $A \cup B$ means the event 'either A occurs or B occurs'.
- **Intersection events**: Let A and B be two events. The set of elements which are in A and in B is called the intersection event of event A and B. it is denoted by $A \cap B$. $A \cap B$ means the event in which A and B occurs simultaneously.

Definitions

- **Differnce events**: Let A and B be two events. The set of elements all elements which are in A but not in B is called the difference event of event A and B. it is denoted by A - B. A - B means the event in which 'the event A occurs but the event B does not occurs'.
- **Mutually exclusively events /disjoint event** : If the intersection event of two events is the impossible event then these two events are said to be mutually exclusive. Thus it is $A \cap B = \varnothing$ then A and B are mutually exclusive events.
- **Example:**
  - When a coin is tossed, the event of occurrence of a head and the event of occurrence of a tail are mutually exclusive events.

Definitions

- **Dependent or Independent events**: Two or more events are considered to be independent if the occurrence of one event is no way affects the occurrence of the other.
- **Example :**
  - Tossing a coin a trial is not affected by the result of the previous trial
- If the occurrence of one event influences the occurrence of the other event, the events are said to be dependent events.
- **Example :**
  - If a card is drawn from a pack of shuffled cards, and not replaced before drawing the second card, then the second card drawn is dependent on the first one.

Definitions

- **Probability Set Function**: Let U be a finite sample space and let S(U) be its power set. Let P: S(U) $\rightarrow$ R be a set function satisfying the following postulates.
  1. $P(A) \geq 0$ for every $A \in S(U)$
  2. $P(U) = 1$
  3. If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$

  Then function P is said to be probability set function on S(U) and the real number P(A) is said to be the probability of event A

Definitions

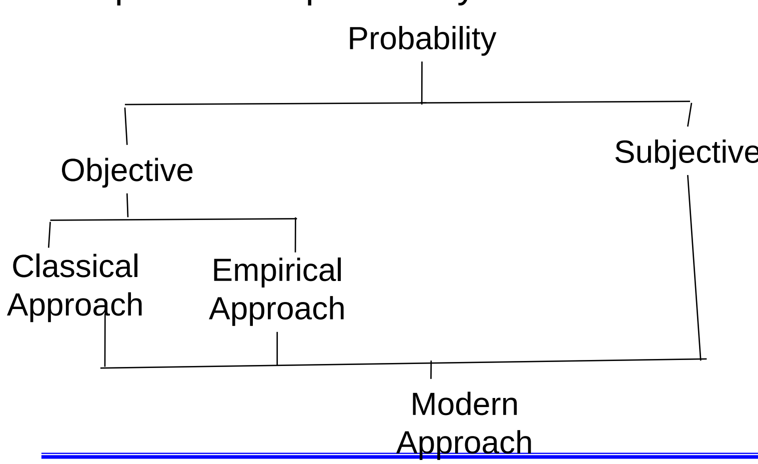
- **Elementary event**: If $U = \{x_1, x_2, \ldots, x_n\}$ is a sample space, then the single element subset $\{x_1\}, \{x_2\}, \ldots, \{x_n\}$ of U are called elementary events or primary events.
- If the probability of each primary event is same, then the primary event are called equi-probable
- If the above n primary events are equi-probable then the probability of each primary event is 1/n.

Definitions

- **Probability of an event / Classical definition of probability:** If 'S' be the sample space, then the probability of occurrence of an event 'E' is defined as:
- $P(E) = n(E)/N(S) = \frac{\text{(number of elements in 'E')}}{\text{(number of elements in sample space 'S')}}$
- **Example:**
  – Find the probability of getting a tail in tossing of a coin.
- Solution:
  – Sample space S = {H,T} and n(s) = 2
  – Event 'E' = {T} and n(E) = 1 therefore P(E) = n(E)/n(S) = ½
- Note: This definition is not true, if (a) The events are not equally likely. (b) The possible outcomes are infinite.

Types

- The following is the broad classification of the concepts used in probability



Classical Approach

- If a random experiment results in N exhaustive mutually exclusive and equally likely outcomes out of which m are favourable to the happening of an event A, then the probability of occurrence of A, usually denoted by P(A) is given by : $P(A) = m / N$
- Example : What is the chance of getting king in a draw from the pack of 52 cards?
- Solution : Total number of cases that can happen = 52
  No. of favourable case = 4

  $\therefore$ Probability of drawing a king = 4 / 52 = 1/ 13

Empirical Approach

- Empirical concept the probability of an event ordinary represents the proportion of times, under identical circumstances, the outcome can be expected to occur.
- The value refers to the event's long run frequency of occurrence.
- The main assumptions are :
  - The experiments or observations are random. As there is no bias in favour or any outcome all elements enjoy equal chance of selection.
  - There are a large number of observations.

Empirical Approach

- “If the experiment be repeated a large number of times under essentially identical conditions, the limiting value of the ratio of the number of times the event A happens to the total number of trails of the experiments as the number of trails increases indefinitely, is called the probability of the occurrence of A.”

Empirical Approach (Example)

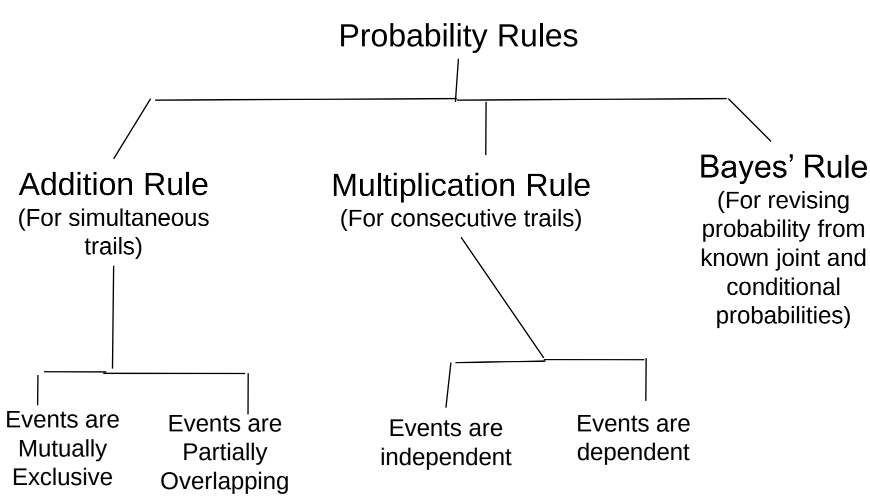
- A foreman in a factory examines the lots of 100 parts each after an interval of half hour during the day and records the number of defective parts. The day’s record of 16 lots reveals the following number of defective parts.

No. of Defective parts	No. of lots
0	1
1	4
2	5
3	3
4	2
5	1

Subjective Probability

- It measures the confidence that an individual has in the truth of a particular proposition. It is bound to vary with person to person and is therefore called subjective probability.

Basic Probability Rules



Addition Rule (Mutually Exclusive Events)

- If A and B are mutually exclusive events, then
  $P(A \cup B) = P(A) + P(B)$
  In other words : P(A or B) = P(A) + P(B)
- Example
  The probability that a company executive will travel by train in 2/3 and the he will travel by plane is 1/5. the probability of his travelling by train or plane is :
- Solution :
  P(T or P) = P(T) + P(P)
  = 2/3 + 1/5 = 13/15
  the probability of not travelling by either train or plane =
  1- P(T or P) = 1 – 13 / 15 = 2 /15.

Addition Rule (Not Mutually Exclusive Events)

- If A and B are any two events, then
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  In other words : P(A or B) = P(A) + P(B) - $P(A \cap B)$
- Generalization : it can be shown that for any 3 events A, B, C
  $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$
- In general, for any events $A_1, A_2, \ldots, A_n$ we have
  $P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$ – (sum of probabilities of all possible intersections taken two at a time) + (sum of probabilities of all possible intersections taken three at a time) + ….. + $(-1)^{n-1}P(A_1 \cap A_2 \cap \ldots \cap A_n)$

Addition Rule (Not Mutually Exclusive Events)

- Deduction : If the two events A and B are mutually exclusive i.e. disjoint, then the addition rule of probability reduces to $P(A \cup B) = P(A) + P(B)$
- In general, for any events $A_1, A_2, \ldots, A_n$ we have $P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots.. + P(A_n)$

Multiplication Rule (Events Are Independent)

- The probability that both independent events A and B will occur is $P(A \cap B) = P(A) * P(B)$
  In other words : P(A and B) = P(A) * P(B)

Multiplication Rule (Events Are Dependent)

- If event A and B are so related that the occurrence of B is affected by the occurrence of A, then A and B are called dependent events. The probability of event B depending on the occurrence of event A is called conditional probability and is written as P(B / A) which may be read, "the probability of B given A."
- The probability that both the dependent events A and B will occur is $P(A \cap B) = P(A) * P(B/A)$
  In other words : P(A and B) = P(A) * P(B/A)

Conditional Probability

- If it is given that a particular event B has already occurred then the probability of occurrence of event A is the conditional probability of A. It is denoted by the symbol P(A/B). If P(B) > 0 then

$$P(A/B) = \frac{P(A \cap B)}{P(B)}\,;\ P(B/A) = \frac{P(A \cap B)}{P(A)}$$

- If P(B) = 0 then B does not occur at all and the question of defining P(A/B) does not arise.
- If the events are independent, the happening of B shall not affect the probability of A and therefore

$$P(A/B) = P(A)$$

Joint And Marginal Probability

- Joint probability is the probability of the occurrence of two or more events.
- Normally the probabilities can be expressed in a 2x2 table or joint probabilities table.
- Example : Two events A and B, there are four joint probabilities possible as explained below :

Joint Event	Sample points belonging to the joint event	Probability
$A \cap B$	$n(A \cap B)$	$n(A \cap B) / n(s)$
$A \cap B'$	$n(A \cap B')$	$n(A \cap B') / n(s)$
$A' \cap B$	$n(A' \cap B)$	$n(A' \cap B) / n(s)$
$A' \cap B'$	$n(A' \cap B')$	$n(A' \cap B') / n(s)$

Bayes' Rule

- Joint and marginal probabilities can be used to revise the probability of a particular event in the light of additional information.
- Example we have two box containing defective and non-defective items.
- One item is picked at random from any one of the boxes and is found defective, and now we might like to known the probability that it came from box one.
- To answer questions of this sort, we use Bayes' Rule which may be considered as an application of conditional probabilities.

Bayes' Rule

- The popularity of the theorem has been mainly because of its usefulness in revising a set of old probabilities called prior probabilities (derived subjectively or objectively) in the light of additional information made available and derive a set of new probabilities called the posterior probabilities.
- Although Bayes' Rule may be applied to more than two mutually exclusive and exhaustive events, we shall for the sake of simplicity confine generally to the application of Bayes' rule for two mutually exclusive and exhaustive events.

Bayes' Rule

- We know that the marginal probabilities are the sum of the two relevant joint probabilities as indicated below.
  - $P(A) = P(A \cap B) + P(A \cap B')$
  - $P(B) = P(A \cap B) + P(A' \cap B)$
- We can restate them in terms of conditional and marginal probabilities as follows
  - $P(A) = P(A /B) * P(B) + P(A / B') * P(B')$
  - $P(B) = P(B / A) * P(A) + P(B / A') * P(A')$
- Now recollect the original formulation of conditional probabilities viz $P(A/B) = P(A \cap B) / P(B) = P(B \cap A) / P(B)$ which can be written as

Bayes' Rule

$$P(A/B)=\frac{P(B/A)*P(A)}{P(B/A)*P(A)+P(B/\overline{A})*P(\overline{A})}$$

- Similarly

$$P(B/A)=\frac{P(A/B)*P(B)}{P(A/B)*P(B)+P(A/\overline{B})*P(\overline{B})}$$

- Proceeding on the same lines, other conditional probabilities viz P(A'/B) and P(B/A') can be revised.

Bayes' Rule

- Generalization : If an event B can only occur in conjunction with one of the n mutually exclusive and exhaustive events A_1, A_2, ….., A_n and if B actually happens, then the probability that it was preceded by the particular event A_i (i = 1, 2, …., n) is given by
$P(A_i / B) = P(A_i \cap B) / P(B)$ i = 1, 2, …., n
and $P(B) = P(A_1 \cap B) + P(A_2 \cap B) +\ldots.. P(A_n \cap B)$

Minal Shah

Theoretical Distribution

Computer Oriented Numerical and Statistical Methods

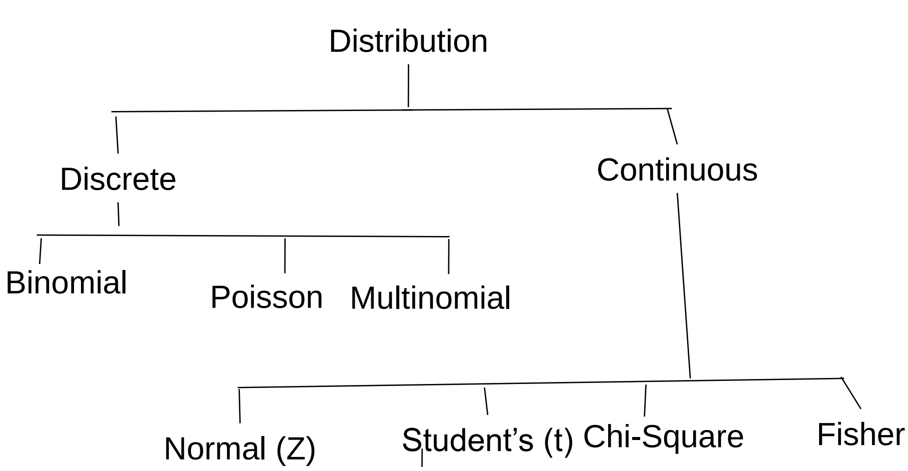
Outline

- Introduction
- Types of Theoretical Distribution

Introduction

- Theoretical Distribution refers to mathematical models of relative frequencies of a finite number of observations of a variable.
- It is a systematic arrangement of probabilities associated with mutually exclusive and collectively exhaustive elementary elements in an experiment.
- Where the relative frequency distributions are based on actual observations, the above distribution are based on mathematical functions.
- Important features of these distributions is that with some known parameters like mean and S.D. or the number of trials and the chances of success, the probabilities of various values of a variate can be found in the form of a complete distribution.

Types of Theoretical Distribution



Discrete Distribution

- The Discrete probability distributions are known as point functions defined over a sample space.
- The random variables in these takes only a finite integer value.
- These are normally represented by line graphs when not grouped and by histograms when grouped, each bar is raised on the mid-value of the class.
- The cumulative probabilities in this case are represented by a staircase type of histogram.

Binomial Distribution

- Binomial is also known as the “Bernoulli Distribution” after the Swiss mathematician James Bernoulli (1654-1705)
- The distribution can be used under the following conditions:
  - The random experiment is performed repeatedly a finite and fixed number of times.
  - The result of any trial can be classified only under two mutually exclusive categories called success (the occurrence of event) and failure (the non-occurrence of event).

Binomial Distribution

- The proportion of outcomes falling in the "success" category are denoted generally by p and the proportion of item falling in the category of "failure" by q = 1 - p
- The probability of success in each trial remains constant and does not change from one trail to another.
- The trails are independent, so that the result of any trial in ineffective by the result of previous trials.

Probability Function Of Binomial Distribution

- If a trial of an experiment can result in success with the probability p and failure with probability q = 1 – p, the probability of exactly x success in n trails is given by $P(X=x) = P(x) = {}^{n}C_{x}p^{x}q^{n-x}$; x = 0, 1, 2,…n
- The quantities n and p are called parameters of the binomial distribution and the notation b(x:n,p) reads "the binomial probability of x given n and p."
- The entire probabilities distribution of x = 0, 1, 2, …, n success with two types of expression, where p stands for success and q for non-success can be written as follows :

Probability Function Of Binomial Distribution

BINOMIAL PROBABILITY DISTRIBUTION			
No. of success x	$(p+q)^n$ Prob. p(x) p(x)	No. of success x	$(p+q)^n$ Prob. p(x) p(x)
n	${}^nC_n p^n q^0$	0	${}^nC_0 p^0 q^n$
n-1	${}^nC_{n-1} p^{n-1} q^1$	1	${}^nC_1 p^1 q^{n-1}$
n-2	${}^nC_{n-2} p^{n-2} q^2$	2	${}^nC_2 p^2 q^{n-2}$
…..	……	…..	……
0	${}^nC_0 p^0 q^n$	n	${}^nC_n p^n q^0$
Total	1		1

Obtaining Coefficient Of The Binomial

- To find the term of the expression of $(q+p)^n$
- The first term is q^n
- The second term is ${}^nC_1 q^{n-1} p$
- In each succeeding term the power of q is reduced by 1 and the power of p is increased by 1.
- The coefficient of any term is found by Pascal's triangle.
- The mean of binomial distribution is n*p
- The S.D. of binomial distribution is $\sqrt{npq}$

Using Binomial Tables

n = 10									
x	…	p=.20	p=.25	p=.30	**p=.35**	p=.40	p=.45	p=.50	
0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	10
1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098	9
2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439	8
3	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172	7
4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051	6
5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461	5
6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051	4
7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172	3
8	…	0.0001	**0.0004**	0.0014	0.0043	0.0106	0.0229	0.0439	**2**
9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098	1
10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0
	…	p=.80	**p=.75**	p=.70	p=.65	p=.60	p=.55	p=.50	**x**

Examples:

n = 10, p = 0.35, x = 3: P(x = 3|n =10, p = 0.35) = 0.2522

n = 10, p = 0.75, x = 2: P(x = 2|n =10, p = 0.75) = 0.0004

Poisson Distribution

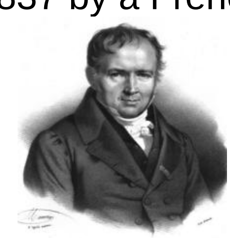
- In Binomial distribution it was found that there is a sample of a definite size so that it is possible to count the number of times an event is observed; in other words, n is precisely known.
- There are certain situations where this may not be possible.
- The basic reason for is that the events is the rare and causal.
- Successful events in the total events space are few e.g. the events like accidents on a road, defects in a product, goals scared at a football match, etc.

Poisson Distribution

- In these, we known the number of times an event occurs but not how many times it does not occur.
- The total number of trials in regard to a given experiment are not precisely known.
- The Poisson distribution is very suitable in case of such rare events.
- Only the average chance of occurrence based on past experience or a small sample extracted for the purpose will enable us to construct the whole distribution.

Poisson Distribution

- The binomial distribution requires only two parameters (p and n) this distribution requires only the value of m which is the mean of the occurrence of an event (np) based on existing knowledge on the matter.
- Poisson distribution was derived in 1837 by a French-Mathematician Simeon D. Poisson.
- Poisson distribution may be obtained as a limiting case of Binomial probability distribution under the following conditions.
  - n, the number of trails is indefinitely large i.e. $n \to \infty$



Poisson Distribution

- p, the constant probability of success for each trial is indefinitely small i.e. $p \to 0$
- np = m (say) is finite.

- The probability function of random variable X following Poisson Distribution is P(X=x) = p(x) = $(m^x / x!) * e^{-m}$ x = 0, 1, 2,…
  - where X = the number of successes (occurrence of the event)
  - e = 2.71828 [The base of the system of natural logarithm] and x! = x(x-1)(x-2)…3.2.1

Poisson Distribution

Values of variables (x)	**0**	**1**	**2**	**3**	**….**	**total**
Prob. P(X=x)= p(x)	e^{-m}	$e^{-m} * m$	$(e^{-m} * m^2)$ /2!	$(e^{-m} * m^3)$ / 3!		1

Constants Of Poisson Distribution

$$\text{Mean } \mu = \lambda$$

$$\text{Variance } \sigma^2 = \lambda$$

$$\text{S.D.} = \sqrt{\lambda}$$

$$p(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \quad k = 0,1,2,\ldots..$$

Minal Shah 17

Using Poisson Tables

X	λ								
	0.10	0.20	0.30	0.40	**0.50**	0.60	0.70	0.80	0.90
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
2	0.0045	0.0164	0.0333	0.0536	**0.0758**	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Example: Find P(X = 2) if λ = 0.50

$$P(X = 2) = \frac{e^{-\lambda}\lambda^X}{X!} = \frac{e^{-0.50}(0.50)^2}{2!} = 0.0758$$

Minal Shah

Multinomial Distribution

- The binomial distribution is generalized as follows. Suppose the sample space of an experiment is partitioned into say mutually exclusive events A_1, A_2, A_2, ….., A_s with respective probabilities P_1, P_2, P_3, …., P_s (hence $P_1 + P_2 + P_3 + \ldots + P_s = 1$).
- Theorem : In n repeated trials, the probability that A_1 occurs K_1 times, A_2 occurs K_2 times and A_s occurs K_s times is equal to

$$\frac{n!}{k_1!k_2!k_3!\ldots k_s!}\, p_1^{\,k_1} p_2^{\,k_2} p_3^{\,k_3} \ldots\ldots p_s^{\,k_s}$$

  where $k_1 + k_2 + \ldots\ldots + k_s = n$
- The above numbers from the so-called multinomial distribution since they are precisely the terms in the expansion of $(p_1+p_2+\ldots+p_s)$

Minal Shah

Continuous Distribution

- These distribution are associated with continuous variables.
- A continuous variable defined over a given range, may take any of the intermediate values. It is always written as an approximate values. For ex. Weight, height etc.
- The continuous variables are generally represented by a smooth curve.
- The cumulative distribution of a continuous variables is also a smooth curve.
- Normal distribution is one of the continuous distribution.

Minal Shah

Normal Distribution

- The most important continuous probability distribution used in the entire statistics is the normal distribution.
- Its graph, called the normal curve, is a bell shaped curve that extends indefinitely in both directions, coming closer and closer to the horizontal axis without even reaching it.
- The mathematical equation of normal curve was developed by De-Moivre in 1733.
- The normal distribution is often referred to as the Gaussian distribution in honour of Karl F. Gauss (1777-1855) who also derived the equation from the study of errors in repeated measurements of the same quantity.

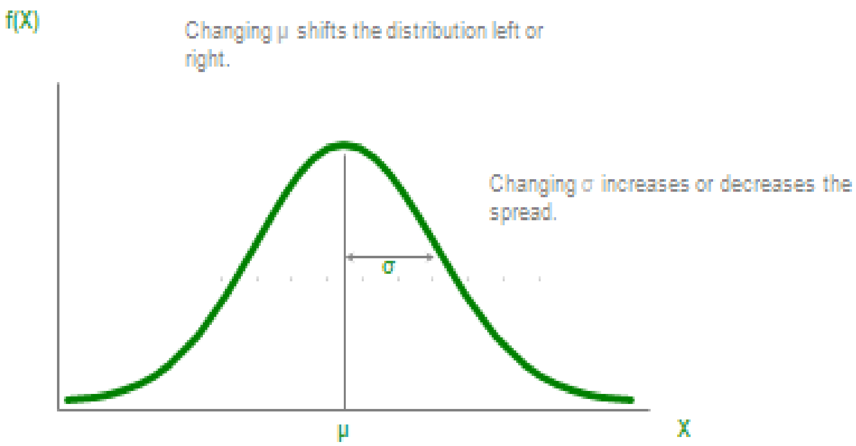
Normal Distribution

- Definition : A continuous random variable X is said to be normally distributed if it has the probability density function represented by the equation:

$$p(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(1/2)[(X-\mu)/\sigma]^2}$$

  - $-\infty \leq x \leq \infty$
  - Where μ and σ, the mean and the standard deviation are known as two parameters and Π = 3.14159 e = 2.7183 are two constant.

Normal Distribution



Normal Distribution

- Now, standard normal distribution or Z – distribution is

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

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

  - Where
  - e = the mathematical constant approximated by 2.71828
  - π = the mathematical constant approximated by 3.14159
  - Z = any value of the standardized normal distribution
  - The Z distribution always has mean = 0 and standard deviation = 1

Fitting A Normal Curve

- There are two main objects of fitting a normal curve to sample data.
  – To provide a visual device for judging whether or not the normal curve is a good fit to the sample data, and
  – To use the smoothed normal curve, instead of the irregular curve representing the sample data, to estimate the characteristics of the population.

Methods Of Fitting

- There are two methods of fitting a normal curve.
  – Method of ordinates,
  – Method of area.

Minal Shah

Statistical Inference

Computer Oriented Numerical and Statistical Methods

- **Statistical Inference** is branch of statistics which is concerned with using probability concept to deal with uncertainty in decision making
- Statistical inference treats two different classes of problems
  - **Hypothesis Testing :** To test some hypothesis about parent population from which the sample is drawn
  - **Estimation :** To use the statistics obtained from the sample as estimate of the unknown parameters of the population from which the sample is drawn

- Hypothesis Testing begins with an assumption, called a hypothesis
- A **hypothesis** in statistics is simply a quantitative statement about a population.
- In order to make statistical decisions, we make an certain assumptions about the population parameters to be tested.
- These ***assumptions*** are known as hypothesis

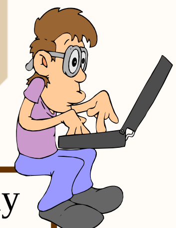
Hypothesis Testing

- There can be several types of hypotheses
- **For example** : The average marks of the 100 students of a class and may get the result as 65% we are now interested in testing the hypothesis that the sample has been drawn from a population with average marks 70%
- A coin may be tossed 100 times and we may get heads 75 time and tails 25 times, we are now interested in testing the hypothesis that the coin is unbiased

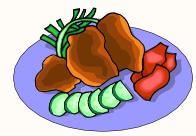
Hypothesis Testing

What is a Hypothesis?

A statement about the value of a population parameter developed for the purpose of testing.



The mean monthly income for systems analysts is $6,325.



Twenty percent of all customers at Bovine's Chop House return for another meal within a month.

What is a Hypothesis?

Hypothesis testing

Based on sample evidence and probability theory

Used to determine whether the hypothesis is a reasonable statement and should not be rejected, or is unreasonable and should be rejected

What is Hypothesis Testing?

Step 1: State null and alternate hypotheses

Step 2: Select a level of significance

Step 3: Identify the test statistic

Step 4: Formulate a decision rule

Step 5: Take a sample, arrive at a decision

Do not reject null | Reject null and accept alternate

Hypothesis Testing

Step One: State the null and alternate hypotheses

Null Hypothesis H_0

A statement about the value of a population parameter

Alternative Hypothesis H_1:

A statement that is accepted if the sample data provide evidence that the null hypothesis is false

Level of Significance

The probability of rejecting the null hypothesis when it is actually true; the level of risk in so doing.

Type I Error

Rejecting the null hypothesis when it is actually true (α).

Type II Error

Accepting the null hypothesis when it is actually false (β).

Step Two: Select a Level of Significance.

- **Defines the unlikely values of the sample statistic if the null hypothesis is true**
  - Defines rejection region of the sampling distribution
- Is designated by α , (level of significance)
  - Typical values are 0.01, 0.05, or 0.10
- Is selected by the researcher at the beginning
- Provides the critical value(s) of the test

Level of Significance, α

Step Two: Select a Level of Significance.

Null Hypothesis	Researcher Accepts H_o	Researcher Rejects H_o
H_o is true	Correct decision	Type I error (α)
H_o is false	Type II Error (β)	Correct decision

Risk table

Test statistic

A value, determined from sample information, used to determine whether or not to reject the null hypothesis.

Examples: z, t, F, χ^2

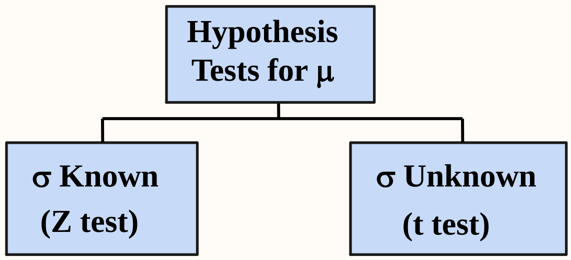
z Distribution as a test statistic

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

The *z* value is based on the sampling distribution of X, which is normally distributed when the sample is reasonably large (recall Central Limit Theorem).

Step Three: Select the test statistic.

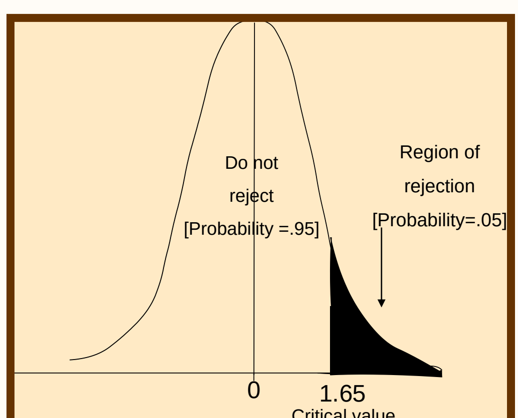
Hypothesis Tests for the Mean



Step Four: Formulate the decision rule.

Critical value: The dividing point between the region where the null hypothesis is rejected and the region where it is not rejected.

Sampling Distribution Of the Statistic *z*, a Right-Tailed Test, .05 Level of Significance



p-Value

The probability, assuming that the null hypothesis is true, of finding a value of the test statistic at least as extreme as the computed value for the test

Decision Rule

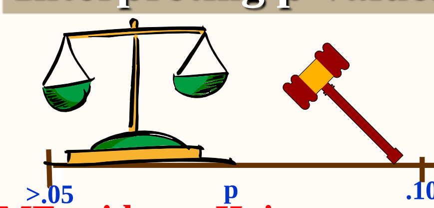
If the *p*-Value is larger than or equal to the significance level, α, H_0 is not rejected.

If the *p*-Value is smaller than the significance level, α, H_0 is rejected.

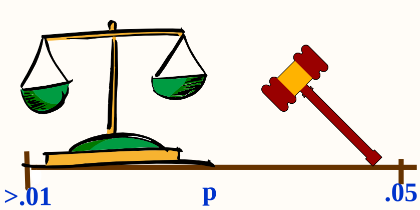
Calculated from the probability distribution function or by computer

Using the p-Value in Hypothesis Testing

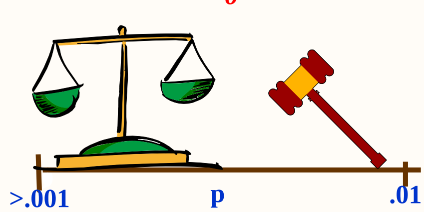
Interpreting p-values



SOME evidence H_o is not true

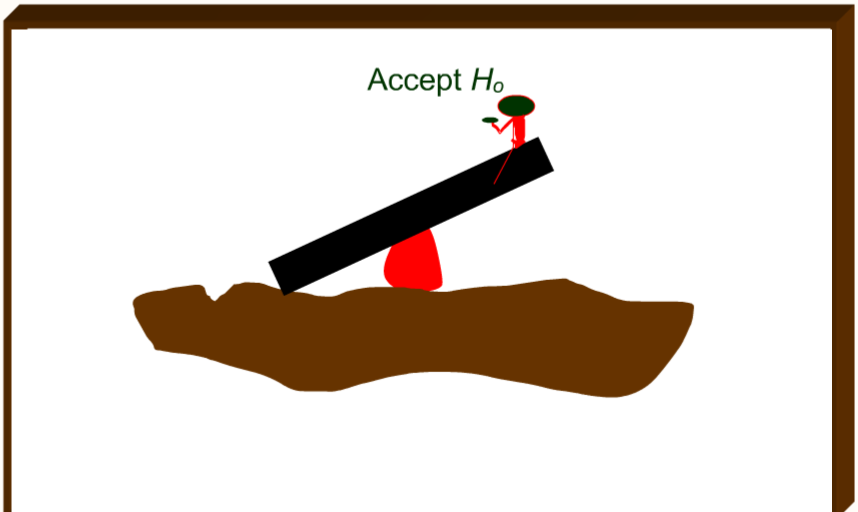


STRONG evidence H_o is not true



VERY STRONG evidence H_o is not true

Step Five: Make a decision.



Movie

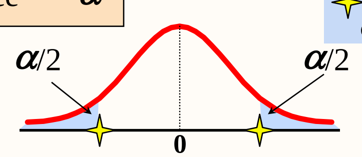
Level of Significance and the Rejection Region

Level of significance = α

Represents critical value

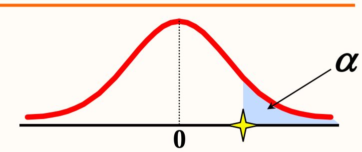
Rejection region is shaded

H_0: $\mu = 3$
H_1: $\mu \neq 3$



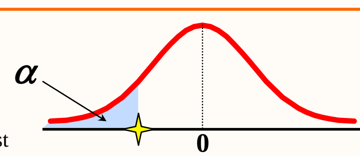
Two-tail test

H_0: $\mu \leq 3$ H_1: $\mu > 3$



Upper-tail test

H_0: $\mu \geq 3$
H_1: $\mu < 3$



Lower-tail test

Test for the population mean from a large sample with population standard deviation known

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Testing for the Population Mean: Large Sample, Population Standard Deviation Known

Hypothesis Testing Example

Test the claim that the true mean # of TV sets in US homes is equal to 3.
(Assume $\sigma = 0.8$)

1. State the appropriate null and alternative hypotheses
   - H_0: $\mu = 3$ H_1: $\mu \neq 3$ (This is a two-tail test)
2. Specify the desired level of significance and the sample size
   - Suppose that $\alpha = 0.05$ and $n = 100$ are chosen for this test

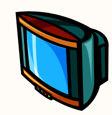
Hypothesis Testing Example

3. Determine the appropriate technique
   - σ is known so this is a Z test.
4. Determine the critical values
   - For α = 0.05 the critical Z values are ±1.96
5. Collect the data and compute the test statistic
   - Suppose the sample results are

     n = 100, $\bar{X}$ = 2.84 (σ = 0.8 is assumed known)

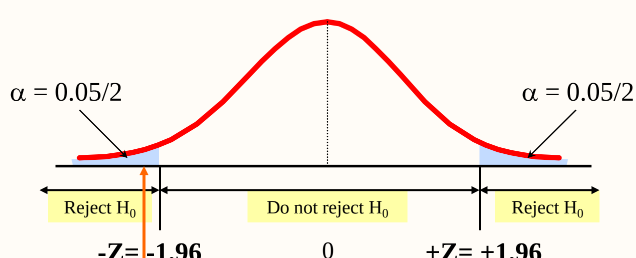
So the test statistic is:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.84 - 3}{\frac{0.8}{\sqrt{100}}} = \frac{-.16}{.08} = -2.0$$



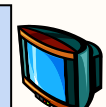
Hypothesis Testing Example

6. Is the test statistic in the rejection region?



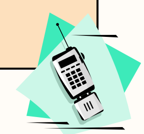
Reject H_0 if Z < -1.96 or Z > 1.96; otherwise do not reject H_0

Here, Z = -2.0 < -1.96, so the test statistic is in the rejection region and conclude that there is sufficient evidence that the mean number of TVs in US homes is not equal to 3



Example: Upper-Tail Z Test for Mean (σ Known)

A phone industry manager thinks that customer monthly cell phone bills have increased, and now average over \$52 per month. The company wishes to test this claim. (Assume $\sigma = 10$ is known)



Form hypothesis test:

H_0: $\mu \leq 52$ the average is not over \$52 per month

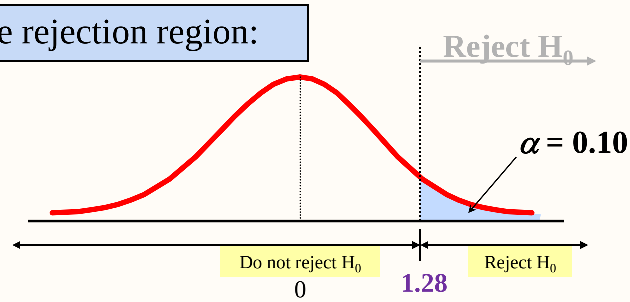
H_1: $\mu > 52$ the average is greater than \$52 per month (i.e., sufficient evidence exists to support the manager's claim)

Example: Find Rejection Region

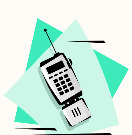
(continued)

- Suppose that $\alpha = 0.10$ is chosen for this test

Find the rejection region:



Reject H_0 if $Z > 1.28$

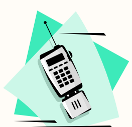


Example: Test Statistic

Obtain sample and compute the test statistic

Suppose a sample is taken with the following results:
$n = 64$, $\bar{X} = 53.1$ (σ=10 was assumed known)

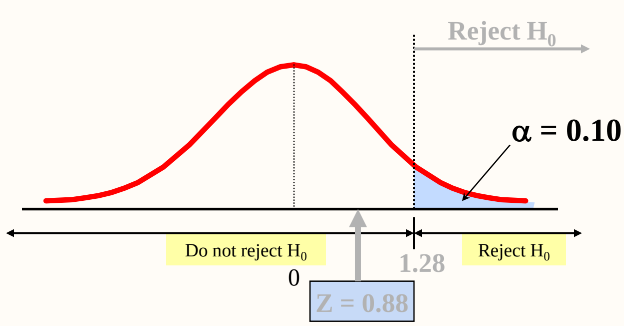
- Then the test statistic is:

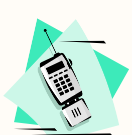


$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{64}}} = 0.88$$

Example: Decision

Reach a decision and interpret the result:





Do not reject H_0 since $Z = 0.88 \leq 1.28$

i.e.: there is not sufficient evidence that the mean bill is over $52

t Test of Hypothesis for the Mean (σ Unknown)

- Convert sample statistic ($\bar{X}$) to a t test statistic

Hypothesis Tests for μ

σ Known (Z test)

σ Unknown (t test)

The test statistic is:

$$t_{n-1} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

Example: Two-Tail Test (σ Unknown)

The average cost of a hotel room in New York is said to be \$168 per night. A random sample of 25 hotels resulted in $\bar{X} = \$172.50$ and

$S = \$15.40$. Test at the

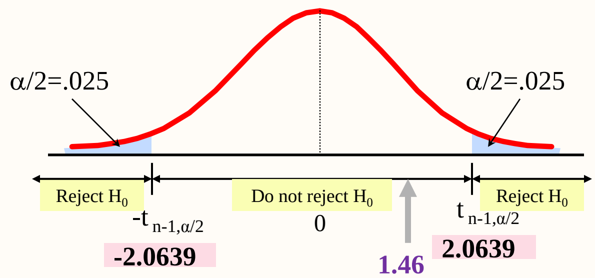
$\alpha = 0.05$ level.

(Assume the population distribution is normal)

H_0: $\mu = 168$
H_1: $\mu \neq 168$

Example Solution: Two-Tail Test

H_0: $\mu = 168$
H_1: $\mu \neq 168$



- $\alpha = 0.05$
- $n = 25$
- **σ is unknown, so use a t statistic**
- **Critical Value:**

$t_{24} = \pm 2.0639$

$$t_{n-1} = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \frac{172.50-168}{\frac{15.40}{\sqrt{25}}} = 1.46$$

Do not reject H_0: not sufficient evidence that true mean cost is different than $168