CHAPTER ONE

1.0 DIFFERENTIATION AND INTEGRAL CALCULUS

The *first derivative* of a function at the point x is the slope of the tangent line at *x*. All linear functions have a constant derivative because the tangent at every point is the line itself.

For instance, the linear function $f(x) = 3x+2$ has first derivative 3. But non-linear functions have derivatives whose value depends on the point x at which it is measured. For instance, the quadratic function $f(x) = 2x^2 + 4x + 1$ has a first derivative that is increasing with *x*. It has value 0 at the point $x=-1$, a positive value when $x>-1$, and a negative value when $x<-1$.

This section defines the derivatives of a function and states the basic rules that we use to Differentiate functions.

DEFINITIONS

We can define the derivative of a function f as:

$$f'(x) = \lim_{\Delta x \downarrow 0} \frac{(f(x+\Delta x) - f(x))}{\Delta x}$$

This is if we take the slope of the chord between two points, a distance Δx apart and see what happens as the two points get closer and closer. The derivative becomes the slope of the target/ slope of the chord as the distance between them becomes zero. The slope of the chord is:

$$\frac{f(x+\Delta x) - f(x)}{\Delta x}$$

If a tangent line is drawn and its slope is the derivative $f'(x)$. As the increment in *x*, Δx gets smaller, in the limit when $\Delta x = 0$, its slope coincides with the slope of the tangent.

The Second derivative is the derivative of the derivative and is given by:-

$$f''(x) = \lim_{\Delta x \downarrow 0} \left\{ \frac{f'(x) + \Delta x) - f'(x)}{\Delta x} \right\}$$

- If we differentiate a function m times, we denote the m^{th} derivative by $f^{(m)}(x)$

We can use the alternative notation df/dx for the first derivative $f'(x)$ and $\frac{d^m f}{dx^m}$ also stands for the m^{th} derivative

- The differential operator d is associated with the definition of the derivative. The differential or total derivative of a function $f(x)$ is given by:

For the derivative of $f(x)g(x)$

$$d/_{dx}\left(f(x)g(x)\right) = f'(x)g(x) + g'(x)f(x)$$

7. **Quotient**

The derivative of the reciprocal of *f (x):*

$$d/_{dx}\left(\frac{1}{f(x)}\right) = \frac{f'(x)}{f(x^2)}$$

More generally, the derivative of a quotient of functions is:-

$$\frac{d}{dx}\left(\frac{g(x)}{f(x)}\right) = \frac{f'(x)g(x) - g'(x)f(x)}{f(x^2)}$$

The above rules allow the derivatives of certain functions to be easily computed. E.g Rule 1 to find the successive derivatives of:-

$$f'(x) = 4x^3, \mathrm{f}''(\mathrm{x}) = 12\mathrm{x}^2, \mathrm{f}'''(\mathrm{x}) = \mathrm{f}^{(3)}(\mathrm{x}) = 24\mathrm{x},\ \mathrm{f}^{(4)}(\mathrm{x}) = 24 \text{ and } \mathrm{f}^{(5)}(\mathrm{x}) = 0$$

Rules 1 and 5 show the first derivative of $f(X) = 2x^2 + 4x + 1$ *is* $4x + 4$, so the first derivative is Zero when $x = -1$

EXERCISE:

Find the first derivative of the functions below

(i) $x^3 - 7x^2 + 14x - 8$

(ii) $10 - 0.5x^2 + \sqrt{\ln(1 + x^2)} + \frac{4\ln(x)}{\exp(x)}$

(iii) $2X^4 + -3X^3 + 5X^2$

(iv) $2X^3 + 3X^2 \ -2/X$

1.2 FIRST ORDER DIFFERENTIAL EQUATIONS AND ITS APPLICATIONS

The emphasis is the formation of differential equations from physical situations. A typical example is the problem of growth in which the rate of change of population of a certain species,

$\frac{dP}{dt}$ At any time t is proportional to the value of p at that instant. This is referred to as the growth constant which is a positive constant:- $\frac{dp}{dt}\, \alpha P$ *or* $\frac{dp}{dt} = kp$

There are many physical problems which can lead to first order differential equations of variables separable type. Some of them are:-

1. Population Growth

Procedure for Selecting the Sample

- Define the sampling frame from which the sample is to be drawn

  **A sampling frame is a list of all units/people included in the sample*
- Define the sample size
- Minimize the level of Bias through inadequacy of sampling frame, items of selected sample not all available and the interviewer or observer bias.

  NB: Bias can rarely be eliminated completely
- Decide on the Sampling method

3.3 METHODS OF SAMPLING

Probability and Non-probability Sampling

The choice of a sampling method will depend on the:-

- Aim of the survey
- Type of Population involved
- Time and funds at your disposal

In probability sampling, every item in the Population has a known chance of being selected as a sample member. In non probability sampling, the probability that any item in the population will be selected for the sample cannot be determined.

Probability Sampling:

1. Simple random sampling

This is a method of sampling in which every member of the population has an equal probability of being selected. The most convenient method for drawing a sample for a survey is to use a table of random numbers. The advantage of this method is that it always produces an unbiased sample while its disadvantage is that sampling units may be difficult or expensive to contact.

2. Systematic sampling

Sometimes called Quasi-random sampling. It involves the selection of a certain proportion of the population. First you decide the size of the sample and then divide it into the population to calculate the proportion of the population you require. This method reduces the amount of time the sample takes to draw. Hence one of its major advantages is the speed with which it can be selected. The major disadvantage is that the sampling frames can easily cause over or under representation of certain characteristics in the sample.

Non Sampling Error

These are the errors that may occur in all surveys during collection of actual information.

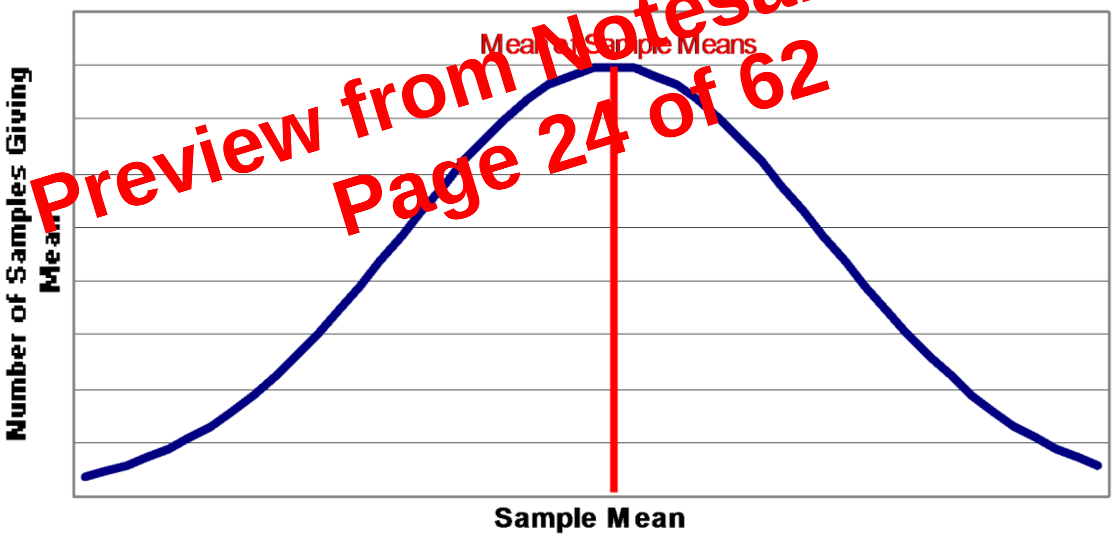
3.5 SAMPLING DISTRIBUTION

A sample distribution is a reflection of the population distribution. As the sample size increases, the sample relative frequency in any class interval gets closer to the true population relative frequency. Therefore the sample distribution looks more like population distribution. A sampling distribution shows how a statistic would vary in repeated data production. It is a probability distribution that determines probabilities of the possible values of a sample statistic (Agresti & Finlay 1997)

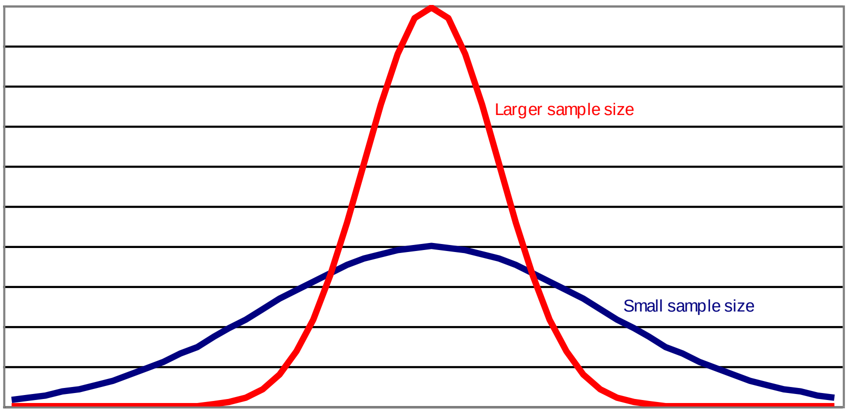
3.6 CENTRAL LIMIT THEORY

The distribution of a sample mean will tend to the normal distribution as sample size increases, regardless of the population distribution. The **central Limit theory** states that if *n* is *sufficiently large*, the **sample means** of random samples from **any** population with mean μ and standard deviation σ are **approximately normally distributed** with mean μ and standard deviation σ/√n

Sampling Distribution



The smaller the sample size the more “spread out” the sampling distribution will be:-



THEORY ESTIMATION:

- **Point Estimation:** When a single sample value (t) is used to estimate parameter (Θ), is called point estimation.
- **Confidence Level:** The confidence level is the probability value associated with a confidence interval. It is often expressed as a percentage. For example say a confidence level of 95% (1- 0.05).
- **Interval Estimation**

3.9 Interval Estimation:

Instead of estimating parameter Θ by a single value, an interval of values is defined. It specifies two values that contain unknown parameter.

I.e. $P(t' \leq \Theta \leq t'') = 1 - a$. Then (t', t'') is called confidence interval.

a is called level of significance e.g 5% or 1%

1-a is called confidence level e.g 95% or 99%

3.10 TEST OF HYPOTHESIS

Statistical inference refers to the process of selecting and using a sample statistic to draw conclusions about the population parameter. Statistical inference deals with two types of problems

1. Testing of hypothesis
2. Estimation

Hypothesis

Hypothesis is a statement subject to verification. More precisely, it is a quantitive statement about a population, the validity of which remains to be tested. In other words, it is an assumption made about a population parameter.

TESTING HTYPOTHESIS

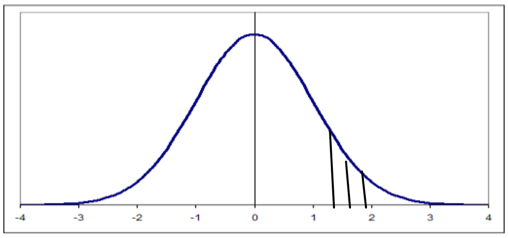
Testing of hypothesis is a process of examining whether the hypothesis formulated by the researcher is valid or not. The major objective is whether to accept or reject the hypothesis.

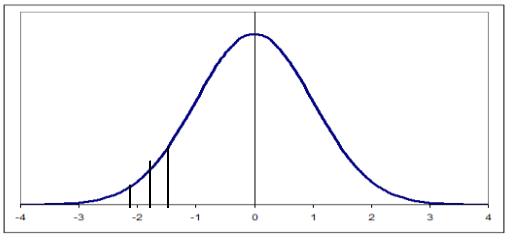
Type II Error: In a hypothesis test, a type II error occurs when the null hypothesis H_0, is not rejected when in fact it is false.

Type II error = (Accept H_0/ H_0 is not true)

3.10.3 One Tailed and Two Tailed test

One Tailed Test: Here the alternate hypothesis H_A is one sided and we test whether the test statistic falls in the critical region on only one side of the distribution





Two Tailed Tests: Here the alternate hypothesis H_A is formulated to test for difference in either direction

Common Test Statistic

One Sample Z-test

$$z = \frac{\overline{x} - \mu_0}{(\sigma / \sqrt{n})}$$

Two Sample Z-test

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\frac{\sigma_1^2}{n1} + \frac{\sigma_2^2}{n2}}}$$

CRITICAL VALUES:

The critical value(s) for a hypothesis test is a threshold to which the value of the test statistic in a sample is compared to determine whether or not the null hypothesis is rejected.

For normal tests:

$$Z = \frac{\text{Difference between X and } \mu}{SE}$$

$$SE = \frac{S}{\sqrt{n}} = \frac{120}{\sqrt{100}} = \frac{120}{10} = 12$$

$$Z = \frac{1600 - 1570}{12}$$

$$= \frac{30}{12} = 2.5$$

The table value at 5% level of significance and infinity d.f. is 1.96. As the calculated value is greater than the table value, we reject the H_0. There is significant difference between mean life of sample and the mean life of population.

2. A factory was producing electric bulbs of average length of 2000 hours. A new manufacturing process was introduced with the hope of increasing the length of life of bulbs. A sample of 25 bulbs produced by the new process were examined and the average length of life was found to be 2200 hours. Examine whether the average length of bulbs was increased assuming the length of lives of bulbs follow normal distribution with σ= (α 0.05)

   SOLUTION:

   H_0: $\mu = 2000$

   H_1: $\mu > 2000$

   Since sample is small, apply test

$$Z = \frac{\text{Difference}}{SE}$$

$$SE = \frac{\sigma}{\sqrt{25}} = \frac{300}{5} = 60$$

$$t = \frac{2200 - 2000}{60} = \frac{200}{60} = 3.33$$

Table value of 't' at 5% significance level and 24 d.f. = 1.711
Calculated value is greater than the table value.

; We reject the null hypothesis and accept alternative hypothesis. So we conclude that the new manufacturing process has increased the life of bulbs, i.e. μ = 200

The factorial symbol! denotes the product of decreasing positive whole numbers.
For example,

$$4! = 4 \bullet 3 \bullet 2 \bullet 1 = 24.$$

By special definition, 0! = 1.

FACTORIAL RULE

A collection of *n* different items can be arranged in order *n*! different ways. (This factorial rule reflects the fact that the first item may be selected in *n* different ways, the second item may be selected in *n – 1* ways, and so on.)

PERMUTATIONS RULE (WHEN ITEMS ARE ALL DIFFERENT)

Requirements:

1. There are *n* different items available. (This rule does not apply if some of the items are identical to others.)
2. We select *r* of the *n* items (without replacement).

We consider rearrangements of the same items to be different sequences. (The permutation of ABC is different from CBA and is counted separately.)
If the preceding requirements are satisfied, the number of permutations (or sequences) of *r* items selected from *n* available items (without replacement) is

$${}_nP_r = \frac{n!}{(n-r)!}$$

PERMUTATIONS RULE (WHEN SOME ITEMS ARE IDENTICAL TO OTHERS)

Requirements:

1. There are n items available, and some items are identical to others.
2. We select all of the n items (without replacement)
3. We consider rearrangements of distinct items to be different sequences.

If the preceding requirements are satisfied, and if there are n_1 alike, n_2 alike, . . . n_k alike, the number of permutations (or sequences) of all items selected without replacement is

$$\frac{n!}{n_1! \,.\, n_2! \,\ldots\ldots\ldots\, n_k!}$$

COMBINATIONS RULE

Requirements:

1. There are *n* different items available.
2. We select *r* of the *n* items (without replacement).
3. We consider rearrangements of the same items to be the same. (The combination of ABC is the same as CBA.)

If the preceding requirements are satisfied, the number of combinations of *r* items selected from *n* different items is

$${}_nC_r = \frac{n!}{(n-r)!\,r!}$$

PERMUTATIONS VERSUS COMBINATIONS

When different orderings of the same items are to be counted separately, we have a permutation problem, but when different orderings are not to be counted separately, we have a combination problem.

5.0 LINEAR PROGRAMMING

5.1 GENERALIZED LINEAR PROGRAMMING MODEL

A generalized linear model (or GLM) consists of three components:

1. **A random component**, specifying the conditional distribution of the response variable, Yi (for the ith of n independently sampled observations), given the values of the explanatory variables in the model. In the initial formulation of GLMs, the distribution of Yi was a member of an exponential family, such as the Gaussian, binomial, Poisson, gamma, or inverse-Gaussian families of distributions.

2. **A linear predictor**—that is a linear function of regressors, $\eta i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}$

3. **A smooth and invertible linearizing link function g(·)**, which transforms the expectation of the response variable, $\mu i = E(Yi)$, to the linear predictor: $g(\mu i) = \eta i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}$

5.2 THE FUNDAMENTAL ASSUMPTIONS OF LINEAR PROGRAMMING

A problem can be realistically represented as a linear program if the following assumptions hold:

1. The constraints and objective function are **<u>linear</u>.**

- This requires that the value of the objective function and the response of each resource expressed by the constraints is <u>proportional</u> to the level of each activity expressed in the variables.
- Linearity also requires that the effects of the value of each variable on the values of the objective function and the constraints are <u>additive</u>. In other words, there can be no interactions between the effects of different activities; i.e., the level of activity X_1 should not affect the costs or benefits associated with the level of activity X_2.

Step 4:

If the current basic feasible solution is not optimal, select a nonbasic variable that should become a basic variable and basic variable which should become a nonbasic variable to determine a new basic feasible solution with an improved objective function value.

The current solution is not optimal. There are negative coefficients in row 0. Since x_2 has the most negative coefficient in row 0 and s_2 has the lowest ratio, the entering and the leaving variables are x_2 and s_2, respectively.

Step 5:

Use elementary row operations to solve for the new basic feasible solution. Return to Step 3

The new basic feasible solution is shown in Table 4, which is the same as Table 2.

Table 4. The tableau for the new basic feasible solution in the first iteration

Basic	Z	x_1	x_2	s_1	s_2	s_3	RHS	Ratio
Z	1	-120	0	0	$160/3$	0	$1760/3$	
s_1	0	2	0	1	0	0	10	5
x_2	0	0	1	0	$1/3$	0	$11/3$	None
s_3	0	1	0	0	$-1/3$	1	$4/3$	$4/3$

Step 6:

Determine if the basic feasible solution is optimal.

The basic solution in Table 4 is still not optimal as the objective function value can be increased by increasing the value of x_1.

Step 7:

If the current basic feasible solution is not optimal, select a non basic variable that should become a basic variable and basic variable which should become a non basic variable to determine a new basic feasible solution with an improved objective function value.

In the second iteration, since x_1 has the most (and only) negative coefficient in row 0 and s_3 has the lowest ratio, the entering and leaving variables are x_1 and s_3, respectively.

Step 8:

Use elementary row operations to solve for the new basic feasible solution. Return to Step 3

The new basic feasible solution is shown in Table 5.