ODUNLAMI, A. A.

INTRODUCTION TO EDUCATIONAL STATISTICS

EDUCATIONAL STATISTICS AND MODEL

ODUNLAMI, A. A.

ISBN:- 978-978-50642-6-1

All correspondence in respect of this publication must be made with the author through the address below:

ODUNLAMI, A. A.

S. 52, ODEREMI ABATAN STREET,

ALAGBADO, LAGOS.

Published by

It's by God's Grace Publishing Coy.

Ilesa, Osun State.

08033035630, 09031568040, 08056459520

ACKNOWLEDGEMENTS

I wish to express my gratitude to God almighty for the opportunity granted me to complete this book.

Words are not enough to express gratitude to my amiable Head of Department of Educational Management Venerable L. A. Oni. Moreso let me use this medium to re-affirm my vow of undying love and affection to the lecturers in the Department. Mr. Ijila, S. O; Mr. Ogunniyi, J. O; Mrs. Opeloye and also Dean of Education, Mr. Kolawole, B. K.

I cannot but appreciate other lecturers in the College whose mentorship role has brought me to where I am. You are indeed a tutor, facilitator, initiator and mentor whose knowledge impartation cannot be quantifiable. I say a big thanks to you all.

My indefatigable appreciation to the following Professors in the Faculty of Education, Ekiti State University, Ado-Ekiti, Ekiti State: Omirin, M. S; Kolawole, E. B; Ajayi, I. A; Komolafe, C. O; Ayodele, J. B. whose role model and knowledge acquisition has been the brain behind the publication of the text. Also to other lecturers' within and outside the faculty whose names were not mentioned. I say thanks

To Ambrose, Emmanuel and Precious for type-setting the Manu-scripts of this text, I say thank you, and Corper Gift for editing, I equally appreciate your effort.

To my lovely siblings, I say kudos to you all. To my wife, and children whose immeasurable efforts play a predominant role in the completion of this book, I say a big thanks to you.

A million thanks to you all.

FOREWORD

As I went through the pages of this book, I could easily conceive a straight forward, but rigorous introduction to the subject of Educational Statistics and Models with details for beginner to understand the basic concepts and techniques of the subject.

The simplicity of the writing style, the day to day practiced examples and the graded exercises are the enticing features of the text.

The author is an experienced teacher and a scholar. He is currently my doctoral students in the Department of Educational Management, Faculty of Education.

I have no doubt in my mind that this text would be of significant interest to students in Secondary schools, Colleges of Education, Polytechnics and Universities.

Obviously, researcher and all users of statistical techniques will find the book handy.

Professor J. B. Ayodele.

Dean, Faculty of Education,

Ekiti State University, Ado-Ekiti, Ekiti State.

TABLE OF CONTENTS

CHAPTER ONE

MEANING OF STATISTICS

Numerical facts are called data and the study of data is called statistics.

Statistics deal with scientific and analytical method of collecting, organizing, analyzing, and presenting data in such a way that some meaning and conclusion could be made out of something that appears to be jungle of data.

Statistics can be defined as the study that involves in making unintelligible intelligible in the sense that, it is from the unintelligible mass of figure we derive intelligible decision which enables us to cope better with life.

Odunlami (2018), define statistics as a systematic and scientific process for gathering, organizing, analyzing, and interpreting numerical data before a meaningful conclusion is drawn.

Types of statistics

Statistics can be broadly classified into two categories viz descriptive and inferential statistics.

Descriptive statistics refer to the type of statistics which deals with collecting, organizing, summarizing, and describing quantitative data. For example, suppose a mathematics teacher finds the average score of his class. The average score is a descriptive statistics since (average score) describes the performance of that class but does not make any generalization about other classes. Example of descriptive statistics are graph chart (pie) chart columnar chart, bar histogram, Pictograph table etc.

Other example of descriptive statistic is central tendency (mode mean median) correlation coefficient (degree of relationship) kurtosis, skewness etc.

Inferential statistics deals with the method by which inferences are made to a larger sample on the basis of the observation made of the smaller sample e.g. suppose a mathematics teacher decide to use the average score of one class to estimate the average score of other two or more classes of the same mathematics course, the process of this estimation is problem of inferential. In short any procedure of making generalization that goes beyond the original data is called inferential statistics.

Inferential statistical provide a way to test the significance of result obtained when data are collected example of inferential statistical tools are student t-test, analysis of variance (ANOVA), analysis of covariance (ANCOVA), correlational analysis etc.

Data classification

Data can be categorized into the following parts.

- Classification by arrangement
- Classification by source
- Classification by measurement
- Classification by precision
- Classification by number of variance

Classification by arrangement include raw data and array

Raw data is the set of unorganized, unarranged and unprocessed information. Raw data can also be defined as any piece of information obtained before it is arranged, analyzed or processed. It is called raw data because it has not been processed by any statistical method.

Example

2	3	5	6	2	4	5	6	7	8
4	6	8	1	0	4	2	7	6	4
5	6	1	10	2	9	8	6	4	9

Array data are data arranged either in ascending or descending order of magnitude.

Example

0	1	1	2	2	2	2	3	4	4
4	4	4	5	5	5	6	6	6	6
6	6	7	7	8	8	8	9	9	10

Advantages of array data

- The smallest and highest scores in the data can be easily seen.
- We can easily obtain the range of the data
- We can see at a glance whether any score appear more than once in the data
- We can easily divide the score into classes
- We can easily see the trend of the data

- We can easily calculate the mode, median and mean of the data
- Classification by source: Source includes primary data and secondary data.

Primary data: The term primary data refers to the type of data originated by the investigator or researcher for the purpose of problem at hand i.e. data collected by the researcher himself for a given purpose is called primary data suppose a researcher collect a data for the purpose of finding out the relationship between school certificate and degree certificate of a given set of students, the set of grades of both school certificate and degree certificate collected by himself and used for the research work are primary data.

Finally, primary data are the data collected and used specially for the purpose for which such data have been collected.

Secondary data: secondary data are the ones which are not originated by the investigator or researcher himself but which he obtain from someone's record i.e. data taken from other data are secondary in the above instance. if the researcher does collect the information himself, he collect the grade from the examination / record office of the student, then the data collected in this case is called secondary data.

Primary Vs Secondary Data

The term primary and secondary data are relative in the sense that data that are primary to one person may be secondary to the other. For example data collected during the grade II examination by institute are primary to National Teacher Institute Department (NTI) but, to a person who uses those data for further research, it is termed as secondary data.

- Classification by measurement: It can be classified into qualitative and quantitative data.

Quantitative data: This data are recorded on a naturally occurring numerical scale. The following are example of quantitative data

A) The weight of 100 students in a class

B) The ages of a set of teacher

c) The height of certain trees in a forest reserve

d) The scores of a sample of 100 testees.

e) The volume of water in a swimming pool

f) The number of civil servant in a state

Qualitative data: Qualitative data are measurement that cannot be measured on a natural scale; they can only be classified into one of the following categories. Example of qualitative data is (i) beauty (ii) attractiveness (iii) the political parties in a particular country (v) the species of a plant.

- Classification by precision**:** variable is any quantity that may take more than one value within a given context. In other words, a values, called the domain of the variable.

Suppose a class teacher collects the ages of the students in his class, then the set of ages is a variable since their ages vary from one member to another.

A constant is any quantity which does not change as value within a given context, examples of constant are $\Pi = {}^{22}/_{7}$, e= 2.7182.

In a given equation y = 4x + 3,

Y and x are variables while 3 is a constant.

Data can also be classified by preciseness, as discrete data and continuous data.

Discrete data can be described precisely as one way of obtaining data by counting. A discrete data is the quantity that assumes an integral value. Discrete data are always expressed in whole numbers. Discrete data are those ones that can be counted. Example is:

(i) The number of lecture theatres in a university.

(ii) The number of theatres could be none, 1,10, 20 and so on but, it could neither be $1\frac{1}{2}$ nor$2^3/_4$

(iii) The number of houses in a given town which can assume any of the values 0,1,2,3,.... But cannot be 1.5 or $100^1/_2$the number of houses is a discrete data.

(iv) The number of motor cars in a car assembly can only take integral value such as 0,1,2,3,4,...10...15..100, thus the number of cars in an assembly is a discrete data. (v) the number of children in a family it could be 0,1,2,3....10 and so on. Discrete data can also be obtained from situation where counting is not involve. Example.

Shoe size of a set of people 10, 42, 6, $3^1/_2$8,4

Shirt size of a set of men 13 ,15,$^1/_2$,14 ,$6^1/_2$, 20

Bed size $3^1/_2$, 5, 5,$^1/_2$ 6 $^1/_2$

A particular characteristics of discrete data is the fact that possible data values progress in definite step like bed size are measured as $2^1/_2$,$3^1/_2$4,5,7,12and so on or there are 1,2,3,4,10……..cars in a motor park (not $1^1/_2$,$3^1/_2$, 10,27)

Continuous data is a data which can take an infinite number of values between any two points on the scale. In otherwords, they cannot be measured precisely their values can only be approximated. Continuous data is the type of data that does progress from point to the next without a break they involve numerical measurement such as weight, height, volume, pressure, temperature, age at times etc. Continuous data can be expressed as decimal or fraction or whole numbers.

Worked example

State which of the following represent continuous and discrete data

A number of children within a family.

B the life span of a fish.

C height of a set of student

D weight of iron sheets

E number of towns in a country.

Solution

(a) Discrete (b) continuous (c) continuous

(d) Continuous (e) discrete

Give the domain of each of the following variable and state whether the variable is discrete or continuous

Solution

Number N of children in a class.

Domain – any integer from o to the number N is in the class

(b)number V of litres of petrol in a can.

Solution

Domain – any number value starting from o to the capacity of the can

Variable –continuous

(C) state X in the old western region of Nigeria

Solution

Domain – Oyo, Ondo, Ogun, Lagos , Edo, Ekiti, Delta, and Osun variable – discrete

(D) radius r of a circle start from zero assuming a point to be a circle)to any value of the radius r.

(E)number of K of cars in assembly

Solution

Domain – K takes any values 01,2,3,…,…….

Variable – discrete

EXERCISE

Express each of the following in numerical form (domain) and state whether it is discrete or continuous

1. Number of lecturers in mathematic department in a college.
2. Number of headquarters in Isokan local government area in Ondo state Nigeria
3. Normal set of teeth of a man.
4. Life span of human beings
5. Number of pages of book
6. Duration of examination
7. Number of state in Nigeria
8. Volume of sphere
9. Ages of animals
10. Number of rivers all over the world

Observation is the value of a variable for a member of a population

Parameter is a characteristic of a population which helps to summaries information about the population with regard to the variable under study. Some of the common parameter is measures of location and measure of dispersion.

A statistic is a characteristic of a sample

Collecting data once you decide on the type of data appropriate (quantitative or qualitative) for the problem at hand there is a need to collect data you can obtain data in four different ways

1 data from a designed experiment

2 data collected observationally

3 data from a publish source

4 data from a survey.

Data from a designed experiment:- This collection involves conducting a designed experiment in which the researcher exert strict control over unit (object or thing or people). In the study, experimental design is made up of treatment /experimental group and control group.

Data collected observationally: in observational study, the researcher observes the experimental unit in their natural setting and records the variable(s) of interest.

Data from a published source: This is done by collecting data from a published source(s) such as books, journals or newspapers, sporting, news, statistics, abstract, annual abstract of statistic, monthly digit of statistics, financial statistics, economic trends and regional trends.

Data from survey: with a survey, the researcher samples a group of people, asks one or more question and records the responses.

Methods of collecting data: The method of collection statistical data may be classified as primary method and secondary method.

The primary methods include:

- Direct personal interview
- Indirect personal interview
- Information from correspondent
- Questionnaire to be filled by the enumerator

Direct personnel interview this is the method whereby the researcher or his agent collect the data through personal contact this is the best method and preferable method since he (researcher) has personal contact with every member of the population used for the research this method reduces the chances of incorrect data being recorded even though this method is the most reliable but most expensive and time consuming. It also involves a lot of experience on the part of a researcher to interact well with his clients in asking personally for his required information.

Indirect personnel interview:- This is whereby the researcher ask his agent to interview his sample on his behalf

The disadvantage of interviewing is that inaccurate or false data may be given to the interviewer the reason may be (1) forgetfulness (2) misunderstanding of concepts.

Questionnaire:- This is the quickest and easiest method of gathering from large and widely scattered members of the population. Questionnaire may be personally delivered to each members of the population under investigation or otherwise the questionnaire should be accomplished with self-addressed envelope with a stamp fixed to it because the respondent may not return it.

EXERCISE

Distinguish between

1. Primary and Secondary data
2. Descriptive and Inferential statistics
3. Observation and Parameters
4. Interview and Questionnaire
5. Give three examples each of raw and array data.

CHAPTER TWO

IMPORTANCE OF STATISTICS

Before any statistical work could be done data must be collected, the collection of data is a very important part of statistics any mistake, error and bias which arise in collection of data will automatically affect our conclusion and decision. The following are some of the importance of statistics:

- A knowledge of statistics is an essential aspect of the training all students of educational planning / educational management must be acquainted with. In Nigeria, Federal and State government have been devoting significant share of their budget to cater for the expansion in the education sector. The role of statistics in ensuring efficiency in the educational system towards the effective pursuance of educational goals is of utmost importance

- Statistics enable educational managers to predict intelligently how much of a thing do we have under a pre-determine conditions.

- It also enables educational managers to draw general conclusions and inference about the phenomenon they are dealing with and the extends to which such conclusion can be generalized.

- Besides, quantifiable (quantitative) data are required on the input and output of the educational system while non-quantifiable (qualitative) data are required in the process of the system.

- Students input: - To access educational goals and development, statistics are needed to take stock of student's population in schools at different levels of educational system and the flow of students to determine their progressive and the efficiency of the system.

- Staff input: - Teaching and non-teaching staff are needed in adequate number (quantity) and staff (quality) in the educational system. Educational managers must know what is in stock and what is required.

- Physical and materials resources input: - physical facilities in terms of land, building, equipment, machines and consumables materials are needed in the educational system. Statistics are needed to determine this for their procurement and for use adequately.

- Financial resources input: - Adequate financing of education demands that there should be adequate facts and figures about sources of income and aspect of educational expenditures.

INPUT---------------PROCESS------------------OUTPUT

Input: - Students input, staff input: - How many students are we taking care of, do we need teachers, how many do we need.

Process: statistics are required for formulating policy and standard operating control and monitoring services in the educational system.

Output: Statistics are needed to access the quantity and quality of the product of the educational system.

Employer's opinion about the product in different worlds of life is also statistics about the product.

Problems of data collection in Nigeria

Despite the enormous benefit of the knowledge of statistics to educational management and manager. It has been very difficult (if not impossible collecting adequate data for planning in Nigeria).

- Data collection at any level is very expensive
- There are problems emanating from government officials; executives or people in high income groups are not easily accessible or approachable under the pretense of official secret or burecracy. This will no doubt hinder accurate and correct data which will in-turn affect the expected outcome.
- In Nigeria the high rate of obsolete of data (outdated) how accurate is the data needed for national planning or historical perspective of Nigeria settings. How many of our museum or document in most Archie's are correct.
- Level of illiteracy is another impediment to data collection in the country. At times it is very difficult in getting information from illiterate people even the ones release may be incorrect
- At times researchers may seat in his office or at home and give report to a problem without recovered to data collection

CHAPTER THREE

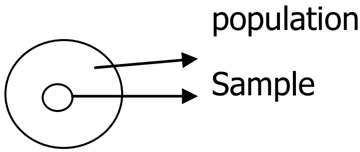
SAMPLING TECHNIQUES

In most cases it is very difficult to have contact with every member of the population so it is best to choose a good respective sample of the population. The method of selecting part of the population is called sample

What is Population?

This can be described as the entire element of subjects in an area of study. The elements may be in-animates or animates, provided they fall within the area of study.

Sample: This is an integral part of a subject of the population.



Characteristics of a good sample

- A good sample must be representatives of the population i.e. all segments in the population must be represented.

- A good sample must take mathematics analysis easy i.e. it must not be lopsided e.g. if you want to pick a sample of 100, we must not make it 90 male and 10 female. If we do that such sample is already lopsided.

- A good sample must be reasonably large i.e. the larger the population, the larger the sample.

- A good sample must not be biased. In otherwords selection must be objective.

- A good sample must be accessible.

Sampling Techniques

In selecting our sample for the population one or more of the following sampling techniques must be considered.

- Simple Random Sampling Techniques:- This is a method of choosing our sample from the population without being biased, whereby every member of the population has equal right of been chosen / picked i.e. there is no partiality and the selection is purely objectives.

This can be done in 3 ways

(a) By balloting

(b) By use of a computer

(c) Use of random number (Raffle draw)

This method has major disadvantages of being lopsided.

- Stratified Random Sampling Techniques:- To solve the problem of being loop sided or a section not been touched at all. This is a techniques whereby the population is divided into different heterogeneous group that are available and from each heterogeneous group, simple random sampling techniques is used (e.g. separate Islam, Christian, Pagan etc.)

Disadvantage:- There could be more elements in one group or the other e.g. in a class of 20. There are 18 Christians and 2 Muslim. This is lopsided already.

- Proportional Stratified Random Sampling Techniques:- The population is divided into different heterogeneous of group in which the sample to be choosing would depend on the number of subjects in that sub-group. It would be proportional e.g. ratio A to B – C. You say total ratio = no picked from each group.

- Multi-Stage Sampling Techniques:- This is a technique whereby choosing our subject will involve more than one stage e.g. we may select some states from the Federation and from those selected states. We may also select some local government and from those local government selected some schools may be selected. It takes more than one stage before you reach your subject.

- Cluster Sampling Techniques:- This is similar to stratified random sampling techniques but the subjects are homogenous i.e. they have the same characteristics. The first thing we do is to divide the population into a number of strata (sub-groups) e.g. all the students here are B.Ed students. I may divide you into group since I am teaching the same group. Once I ask one student where do we stop? I do not need to doubt him because he falls within the same group I am teaching while in stratified they do not have anything in common. If they are

20 sub-group. You must visit all the groups because they are heterogeneous groups.

- Snow-balling Sampling Techniques:- This is a technique where you start with a small group or even an individual who is a volunteer and he would be bringing other members. If a revival is organized and the first day there are 2 to 3 people and wonders happen the next day. The population increases on daily basis. If a student is caught for examination mal-practice and the student caught say he is not the only one. It is called snow-baling. Take a stick of matches and strike it, it can burn the whole house. It starts from a very small point before expanding gradually.

- Systematic Sampling Techniques:- This is a technique whereby the sample is selected at a regular and constant interval. It could be a multiple of N. For instance if your no is multiple of 5 e.g. 5, 10, 15, 20, 25, 30 etc. It must be at regular interval.

- Quota Sampling Techniques:- Here, every segment is represented either qualify or not qualify. The New P. D. P. chairman if not for quota, he would not have been choosing.

- Purposeful / qualify Sampling Techniques:- This is a type of techniques whereby any subject selected must have a reason and be for a particular purpose. If you put anybody there, you must be able to give reasons why you have choosing such a person. For instance you want to go and pray for a classmate who is in hospital and he is a Muslim. You must pick one person who is a Muslim among those that are going for such visitation.

EXERCISE

1a Define data.

1b List and discuss various classification of data.

1c. Mention and discuss method of collecting data.

2a. Explain how population and sample differ.

2b.Mention and discuss various method of sampling techniques.

CHAPTER FOUR

CENTRAL TENDENCY

An average or measure of central tendency is a typical or representative of a set of data, since such representative score tends to lie centrally within the array of data.

Types of Central Tendency

The commonest types of average are (a) mode (b) median (c) mean

Find the range, mean, median and mode of the following ungrouped data.

1,3,5,3,3,6,7,8,1,3,9

Range = Highest – Lowest number

Since 9 is the highest number and 1 is the lowest number. Therefore, range = 9 – 1 = 8

To find mean which is otherwise called average

Addition of all the number divided by the number of appearance.

$$\frac{1+3+5+3+3+6+7+8+1+3+9}{11}$$

Mean =49/11 = 4.45

Properties of Mean

1. The value of the mean is determined by every score in the series.
2. It is greatly affected by extreme values.
3. The sum of the deviations about the mean is zero.
4. The sum of the squares of deviations from the mean is less than those computed from any other score.
5. The mean reflects every score in the distribution.
6. The mean is shifting towards where ever there is a change.

Median= to find the median of any number, it must be re-arrange either through ascending or descending order.

Properties of Median

1. Median is a positional average. It is not influenced by the size of items but by the position of the items.
2. If the median is less than the mean, the distribution is skewed towards right (positively skewed). But if the median is greater than the mean, the distribution is skewed towards the left (negatively skewed, if the mean =median = mode, then the distribution is symmetrical.
3. The sum of the absolute values of the deviations is at least from the deviation are measured from the median.

1,1,3,3,3/,3,/5,6,7,8,9

Median = 3 since 3 is the half way of 11 and when it involves even number the median would be two e.g. 2,2,3,/5,6,/7,8,9. The median is 5 and 6 divided by 2.

Mode on the otherhand would be the number that occurs most

Mode = 3 since three occur 4 times.

Importance of Mode

1. The mode is the mode descriptive average since it signifies the most typical value in the given set, and indicates the precise value of an important part of it.

2. The mode is not affected by the extreme; hence, it will be a more representative average of many purposes.

GROUPED DATA

Class Interval	F	X	FX
11-15	2	13	26
16-20	1	18	18
21-25	4	23	92
26-30	2	28	56
31-35	1	33	33
	10		225

X ∑FX/∑F =225/10 Mean = 225

Let assume mean be preferable to the one around the middle. Anytime you see d Assume mean must be there.

Class Interval	F	X	d	Fd
11-15	2	13	-5	-10
16-20	1	18	0	0
21-25	4	23	5	20
26-30	2	28	10	20
31-35	1	33	15	15
	10			45

Let assume mean = 18

Let: A.M. + ∑Fd/∑F

18 + 45/10

18+4.5

= 22.5

ALTER MEAN OR ALTERNATIVE MEAN

Class Interval	F	X	X/i	FXi
11-15	2	13	6.5	13
16-20	1	18	9	9
21-25	4	23	11.5	46
26-30	2	28	14	28
31-35	1	33	16.5	16.5
	10			112.5

I = 2 x =i

Mean = what you would have use to divide use it to multiply.

X iΣFX/ΣF

= 2 (112.5)/10

X = 22.5

Let's change our alter to 5

Class Interval	F	X	X/i	FXi
11-15	2	13	2.6	5.2
16-20	1	18	3.6	3.6
21-25	4	23	4.6	18.4
26-30	2	28	5.6	11.2
31-35	1	33	6.6	6.6
	10			45

I = 5 x =i

X iΣFX/ΣF

= 5 (45)/10

X = 22.5

MEDIAN

Class Interval	F	FX
11-15	2	2
16-20	1	3

21-25	4	7
26-30	2	9
31-35	1	10
	10	

Size of median is obtained by N/2

Median: It means middle or half way

Where L_1 =Lower boundary of the median class

N =Total frequency

C =median class size

Fw =median class frequency

Cfb =cumulative frequency of all the classes lower than the median class.

$$\text{Med. } L_1 + \frac{(N/2 - Cfb)}{Fw} \text{ Class interval}$$

20.5 + (10/2 -2) x 5

20.5 + (2/4) x 5

20.5 + 2.5

Median = 23

ALTER OR ALTERNATIVELY WAY OF GETTING MEDIAN

Class Interval	F	CF
11-15	2	2

16-20	1	3
21-25	4	7
26-30	2	9
31-35	1	10
	10	

3 and 7 which one is greater 7. That is why we pick 7 for Median.

Mean = N/2 = 10/2 = 5

Inferimo, lower boundaries, lower limit, superimo, higher limit, higher boundaries.

ALTER ALTERNATIVE WAY OF FINDING MEDIAN

$$\text{Median} = L_1 + \frac{L_2 - L_1}{Fw}(M - Cfb)$$

$$\frac{20.5 + 25.5 - 20.5}{4} (5 - 3)$$

20.5 + 5/4 (2)

= 20.5 + 2.5 = 23

MODE

Class Interval	F	CF
11-15	2	2
16-20	1	3
21-25	4	7

26-30	2	9
31-35	1	10
	10	

$$\text{Mode} = L_1 = \frac{(D_1)}{(D_{1} + D_2)} \times i$$

$$= 20.5 + \frac{(4-1)}{(4-1) + (4-2)} \times 5$$

$$20.5 + \frac{(3)}{(3+2)} \times 5$$

$20.5 + (3/5) \times 5$

$20.5 + 3$

23.5

ALTER OR ALTERNATIVE WAY OF GETTING MODE

Class Interval	F	CF
11-15	2	2
16-20	1	3
21-25	4	7
26-30	2	9
31-35	1	10
	10	

$$\text{Mode} = L_1 + \frac{F_1 - F_0}{(2F_1 - F_0 - F_2)} \times i$$

$$= 20.5 + \frac{(4-1)}{2(4) - 1 - 2}$$

$$20.5 + \frac{(3)}{(8-3)}$$

20.5 + (3/5)5

20.5+ 3

= 23.5

Where

L_2 =upper boundaries of the modal class.

L_1 =lower boundaries of the modal class.

C =class size of the modal $L_2 - L_1$.

F_0 =frequency of the immediate lower class to the modal class.

F_1=frequency of the modal class.

F_2=frequency of the immediate higher class is the modal class.

Estimate the mode of the distribution

Class interval	0-4	5-9	10-14	15-19	20-24	20-24	25-29	30-34
Frequency	1	2	3	5	4	2	2	1

Solution

Class interval	frequency
0-4	1
5-9	2
10-14	3 f_0
15-19	5 f_3
20-24	4 f_2
25-29	2
30-34	1

Modal class = 15-19

L_1 = 14.5 (2) L_2 = 19.5 (3) C=L2-L1=19.5-14.5 = 5 (4) f_0 = 3 (5) f_1 = 5 (6) f_2 = 4

$$Mode = L1 + \frac{(f_1 - f_0)}{F1 - f2 + f1 - f0} C$$

$$= 14.5 + \frac{(5-3)5}{(5-3)+(5-4)}$$

$$= 14.5 + \frac{10}{2+1}$$

= 14.5 + 10/3 =14.5 + 3.333 = 17.833

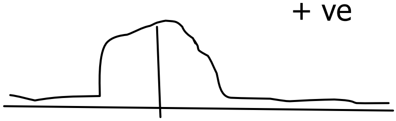
Skewness

Skewness is the degree of asymmetry i.e. the existent to which a distribution deviate from the normal.

Types of Skewness

There are two types of skewness positive skewness and the negative skewness. A distribution is said to be positively skewed when the tail of the distribution is towards the right. For example if you are given any distribution in which the mode is smaller. You have positive skewness and the answer is bad.

+ ve



Mode 20 mean 60 median 70

It shows that the performance is poor.

Positive skewness occurs whereby a test is too difficult. The scores may look like the figure on the right with little difference among the poorer students. While Negative skewness is a type of skewness whereby the tail of the distribution is towards the left.

-ve

mean 30 median 40 mode 70

If the answer for mode is bigger than mean, median. The answer is very good. It shows that the performance is good.

Negatively skewness occurs when a test is too easy. The score may look like figure on the left with no appreciative difference among the best students.

Two levels of Skewness

$$Sk = \frac{Mean - Mode}{SD} \text{ ---------------------(1)}$$

Sk = 3 (mean – median) -----------------(2)

Complete (1) the coefficient of skewness for each of the following and give literary interpretation to your answer.

Mode = 70, Mean =80 standard deviation = 20

Median = 35, Mean =35, standard deviation = 10

Mode = 17, Mean =12, Median =15

Median =13, Mean = 20, Mode = 15

Solution

Coefficient of skewness = mean – mode

$$= \frac{80-70}{20} = \frac{10}{20} = 0.5$$

Coefficient of skewness = 3 (mean – median)

$$= \frac{3(35-45)}{10} = \frac{3(-10)}{10} = -3$$

Standard Deviation

Class interval	F	x	Fx	$(x-\bar{x})^2$	$F(x-x)^2$
11-15	2	13	26	90.25	180.5
16-20	1	18	18	20.25	20.25
21-25	4	23	92	0.25	1
26-30	2	28	56	30.25	60.25
31-35	1	33	33	110.25	110.25
	10		225		372.5

$$\bar{x} = \frac{\Sigma fx}{\Sigma f}$$

$\bar{X} = 225$

10

$\bar{X}$ =22.5

(x - $\bar{x}$) =

13 – 225

-9.5^2

90.25

Variance = $\Sigma f \underline{(x - \bar{x})^2}$

Σf

= $\frac{372.5}{10}$

Variance = 372.5

S.D = $\sqrt{\frac{\Sigma f (x-\bar{x})^2}{\Sigma f}}$

= $\frac{372.5}{10}$

= 37.25

S.D = 6.1032

S.D = 6.1

EXAMPLE

1. In a moderately asymmetrical distribution the mode and mean are 25.6 and 36.5 respectively. What is the median?

2. In a moderately skewed distribution, the mean and median re respectively 21 and 20.6. Calculate the mean.

3. If the mean and median of an asymmetrical distribution are 23.5 cm and 24 cm respectively. Estimate the mode.

(1)Given mode = 25.6 mean = 36.5, median?

Mean – Median = 1/3 (mean – mode)

36.5 – median = 1/3 (36.5-25.6)

= 1/3 (10.9) = 3.63

Median = 36.5 -3.63 = 32.87

(2) Mode = 21, median = 20.6 and mean?

Mean = ½ (3 median-mode)

=1/2 (2 x 20.6) – 21

½ (61.8 -21)

½ (40.8)

=20.4

(3) Mean = 23.5, median = 24.0, mode?

Mode = 3 median – 2 mean

3 (24,0)- 2(23.5)

72 – 47 = 25cm

Kurtosis

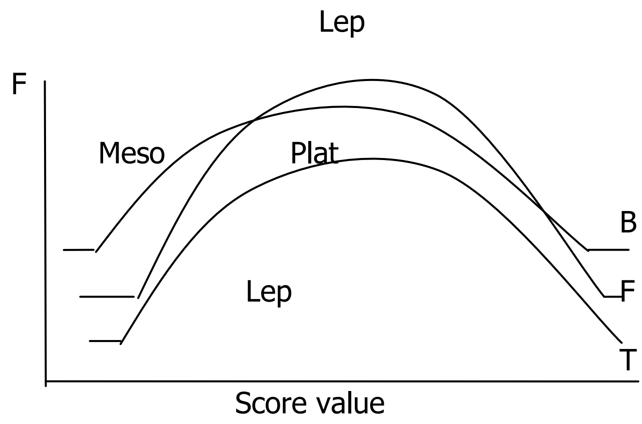
Kurtosis of a distribution is the peakness or otherwise of a distribution.

There are three main type of kurtosis vis – lepokurtic and platkurtic.

Leptokurtic – highly peaked (lepo-thin)

Mesokurtics – moderately peaked

Platykurtic – flattened (plat – flat)



In the fig. above represents two curves with similar central tendencies but different kurtosis Curve T is more peaked that F and the change in height of T is more rapid than that of F as the score value increases. Distribution T is more leptokurtic (thinner) than B. On the otherhand, F is said to be more platykurtic (flatter) than T. Percentile coefficient of kurtosis =

$$\frac{1}{2}\frac{(Q_3 - Q_1)}{P_{90} - P_{10}}$$

Where Q_1 = Lower quartile

Q_3 = Upper quartile

P_{90} = 90th percentile

p_{10} = 10th percentile

EXERCISE

Class Interval	F
10000 - 11999	12
12000 - 13999	14
14000 – 15999	24
16000 – 17999	15
18000 – 19999	13
20000 - 21999	7
22000 – 23999	6
24000 – 25999	4
26000 – 27999	3
28000 – 29999	2

Find the mean, assume mean, median, alter median, mode, alter mode, standard deviation and skewness.

CHAPTER FIVE

QUANTILES

Quantiles are a point value which divides the set of observation into two groups with known proportions in each group. Examples of quantiles are deciles, percentiles etc.

Quartiles are the three point values (Q_1, Q_2, Q_3) which divide an array data into four equal parts.

Q_1 known as the first quartile is a point value which divides the distribution into ratio 1:3 (25% or ¼) of the items are below and 75% = ¾ of the items are above).

Q_2 known as the median is a point value that divides the distribution based on the frequency into two equal parts.

Q_3 known as the upper quartile is a point value that divides the distribution based on the frequency into ratio 3:1 (75% = ¾ of the items are below the point and 25%= ¼ are above it).

Example 1

Compute the first middle and 3rd quartiles of 2, 3, 5, 8, 9, 8, 15, 10, 7

Solution

Array	2	3	5	7	8	8	9	10	15
Rank	1st	2nd	3nd	4th	5th	6th	7th	8th	9th

1st quartile = lower quartile = $\frac{(N+1)^{th}}{(4)} = \frac{(9+1)^{th}}{(4)}$

$= 2.5^{th} = \frac{3 + 5}{2} = 4$

2nd quartile = median = $\frac{(N+1)^{th}}{(2)} = \frac{(9+1)^{th}}{(2)}$

$= 5^{th} \qquad = 8$

3rd quartile = upper quartile = ¾(N+1)th = ¾(10)

$= 5^{th} = 8$

2rd quartile = upper qua tile = ¾ (N + 1)th = $^3/_4(10)$

$= 7.5^{th} = \frac{9 + 10}{2} = 9.5$

Example

Compute the lower quartile, median and the upper quartile of the following distribution.

Score	11	12	13	14	15	16	17	18
Frequency	1	2	3	4	4	3	2	1

Solution

Score	frequency	Cumulative frequency	Ranks
11	1	1	1
12	2	3	2-3
13	3	6	4-6
14	4	10	7-10
15	4	14	11-14
16	3	17	15-17
17	2	19	18-19
18	1	20	20

N=20

Lower quartile $= Q_1 = {}^1/_4(N+1)^{th}$ score

$= {}^1/_4\,(20+1) = 5\,{}^1/_4{}^{th}$ score = 13

Median $= Q_2 = {}^1/_2(N+1)^{th}$ score

$= {}^1/_2\,(20+1) = 10\,{}^1/_2{}^{th}$ score = 14.5

Upper quartile $= Q_3 = {}^3/_4\,(N+1)^{th}$ score

$= {}^3/_4(21) = 15\,{}^3/_4{}^{th}$ score = 16

Quartiles for grouped data

$$Q_1 = L_1 + \left(\frac{\frac{N}{4} - CF_Q}{fQ_1}\right) C$$

Where I= 1, 2, 3

L_1 = Lower boundary of the lower quartile class.

CF_{Qi} = Cumulative frequency up to the lower class next to the Q class

F_{Qi} = frequency of the Q_1 class.

C = class size of the Q_1 class.

Example

Find Q_1, Q_3 semi-interquartile, inter-quartile, Decile 7 and 77^{th} percentile.

Class interval	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
Frequency	2	4	7	10	6	5	3	2	1

Class interval	frequency	CF	Ranks
10-20	2	2	1^{st} -2^{nd}
20-30	4	6cfQ	3nd- 6th
30-40	7	13	7^{th}- 13^{th} Q_1=10^{th}
40-50	10	23	14^{th} -23^{th} Q2 = 29^{th}
50-60	6	29	24^{th}-29^{th}
60-70	5	34	30^{th}-34^{th} Q^3 =30^{th}
70-80	3	37	35^{th}-37^{th}
80-90	2	39	38^{th}-39^{th}
90-100	1	40	40^{th}

Lower quartile

$$Q_1 = L_1 + \left(\frac{N/4 - Cfb}{Fw}\right) \times 10$$

$$30.5 + \left(\frac{40/4 - 6}{7}\right) \times 10$$

$$30.5 + \left(\frac{10 - 6}{7}\right) \times 10$$

$$30.5 + \left(\frac{4}{7}\right) \times 10$$

$$30.5 + 5.71 = 36.2\ q_1$$

$$Q_3 = L_1 + \left(\frac{3N/4 - Cfb}{Fw}\right) \times 10$$

$$60.5 + \left(\frac{3x40 - 29}{5}\right) \times 10$$

$$60.5 + \left(\frac{30 - 29}{5}\right) \times 10$$

$$60.5 + \left(\frac{1}{5}\right) \times 10$$

$$60.5 + 10/5$$

$$60.5 + 2 = 62.5\ q_3$$

Inter-quartile range = $Q_3 - Q_1$ = 62.5 – 36.2 Interquartile = 26.3

Semi-interquartile range = $\frac{Q_3 - Q_1}{2} = \frac{62.5 - 36.2}{2}\ \frac{26.3}{2}$

Semi-interquartile range =13.15

$D_7 = L_1 + \left(\frac{7N/10 - Cfb}{Fw}\right) \times 10$

$L_1 + \left(\frac{7 \times 40/10 - Cfb}{6}\right) \times 10$

$50.5 + \left(\frac{28 - 23}{6}\right) \times 10$

$50.5 + \left(\frac{5}{6}\right) \times 10$

50.5 + 50/6

50.5 + 8.333

D_7 = 58.8

$P_{77} = L_1 + \left(\frac{77N/100 - Cfb}{Fw}\right) \times 10$

$L_1 + \left(\frac{7 \times 40/100 - Cfb}{6}\right) \times 10$

$60.5 + \left(\frac{34 - 29}{5}\right) \times 10$

$60.5 + (\underline{5}) \times 10$

5

60.5+ 50/5

60.5 + 10

P_{77}=70.5

(1) Inter-quartile range is the positive difference between the upper and lower quartiles.

(2) semi inter-quartile range or quartile deviation is the half the difference between the lower and upper quartiles.

EXERCISE

The frequency distribution of marks in an examination was as follows:

Class interval	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-51
Frequency	2	3	5	10	12	13	9	6	3	1

Find q_1, q_3, inter-quartile range, semi-interquartile range, D_9 and 85^{th} percentile.

Deciles

Deciles are the nine point value which divides the array set of item into ten equal parts. The deciles are D_1,D_2,……..D_9

D_1 is a point value which divides the distribution into ratio 1:9, 10% of

the distribution lies below it and 90% lie above the point D_1

N.B.

$Q_2 = D_5$ = median

Percentiles are the ninety point value which divide the distribution into 100 equal parts.($p_1, p_2 \ldots\ldots p_{99}$).

Quartiles are the four point values which divide the distribution to 5 equal parts.

Qctiles are the seven point values which divide the distribution into eight equal parts.

Quartiles formula for ungrouped data

It should be noted that the location is based on the distribution frequency

$Q_1 = \frac{(N+1)}{4}^{th}$ score =Lower quartile

$D_2 = \frac{(N+1)}{2}^{th}$ score =median

$Q_3 = \underline{(N+1)}^{th}$ score =upper quartile

$D_1 = \frac{1}{10}(N+1)^{th}$ score

$D_7 = \frac{7}{10}(N+1)^{th}$ score

$D_{51} = \frac{51}{100}(N+1)^{th}$ 14.7^{th} =14.41

= 14 + (0.41) = 14.41

Example

Find the first deciles, 7th deciles, 21st percentile 3rd octile of the following

distribution.
12,13,14,16,18,17,19,11,23,20,22,24,25,26,28,27,29,30,15.

Solution

Mark 11 12 13 14 15 16 17 18 19 20 21 22 23

Rank 1^{th}, 2^{th}, 3^{th}, 4^{th}, 5^{th}, 6^{th}, 7^{th}, 8^{th},9^{th} ,10^{th} ,11^{th} ,12^{th} ,13^{th}

Mark 24 25 26 27 28 29 30

Rank 14^{th} 15^{th} 16^{th} 17^{th} 18^{th} 19^{th} 20^{th}

$D_1 = \frac{1}{10}(N+1)^{th}$ = 21th

$= 12 + 1 \frac{(13-12)}{10}$ =12.1

$D_7 = \frac{7}{10}(N+1)^{th} = 14.7^{th}$

$= 24 + \frac{7}{10}(25 - 24)$ =24.7

$P_{21} = \frac{21}{100}(N+1)^{th} = 14.7^{th}$ =14.41

= 14 + (0.41) 1 +14.41

Example

The below was the age distribution of United Kingdom population in hundred thousand at June 30, 1962.

10-14	119
15-29	99
30-44	108
45-59	102
60-74	62
75-89	21

Calculate the D_1,D_3,P_3,5_{th} Octile 3^{rd} Quintile.

Solution

Class interval	**Frequency**	**Cf**	**Rank**
10-14	119	119	1 – 119
15-29	99	218	120 -218
30-44	108	326	219 – 326
45-59	102	428	327 – 428
60-75	64	492	429 – 492
75-89	21	513	493 – 513
D_1=1st Decile,	N = 513,		N = 513/10 = 51.3

D^1 class =10-14, f_{D1} = 119, Cf_b=O

L_1=0.5,C=15

$$D_1 = L_1 + \frac{(N/10 - cf_{D1})C}{f_{D3}}$$

$$= -0.5 + \frac{(51.3-0)15}{119}$$

$= -0.5+6.47$

$=5.97$

N = 513, D_3 = 3N = 3(513) = 153.9

10 10

D_3 class = 15-29

L_1 = 14.5, 3N = 143.9, cfD_3 = 119, C =

10

fD_3 = 99

D3 = L1 + $(3N/10-cfD_3)C$
fD_3

=1 4.5 + (153.9-119)15
99

=14.5 + (34.9)15
99

=145 +5.29 =19.79

3rd percentile

N = 513, 3Nth = 3 x 513

100 100

P_3 class = 0 -14, L_1 = -0.5,cf_{P3} = 0, C = 15

P_3 =0.5 + (15.39)15

119

= 0.5 + 1.94 = 2.44

5^{th} Octile

N = 513, $5N^{th}$ = $5(513)^{th}$

8 8

=320.625^{th}

5^{th} Octile class = 320.625^{th} = 30 – 44

L1 = 29.5, 5N = 320.625. cfQ3 = 218, C = 15

8

5^{th} Octile = L1 + $\underline{(5N/8 - cf_{D3})}$15

180

= 29.5 + $\underline{(320.625 - 218)}$15

180

= 29.5 + 14.25 = 43.75

3^{rd} Quintile

N = 513, $\underline{3N^{th}}$ = 3 x $\underline{512^{th}}$

5 5

=307.8th

3^{rd} Quintile Class 307.8^{th} = 30-44

3^{rd} Quintile = L1 + $\underline{(3N/5 - CFQ)}$15

180

= 29.5 + $\underline{(307.8-218)}$15

180

=29.5 + 12.47 + 41.97

CHAPTER SIX

CHI-SQUARE

It is a degree of relationship which involve frequency counts

If calculate or critical value is greater than the table value. It shows that the result is significant which contradiction to the null-hypothesis is. The null-hypothesis would be rejected.

What is hypothesis?

Hypothesis can be defined as a tentative statement or a congestional statement connecting 2 or more variables. Some call it a guess, or a wise guess. There are 2 types

(1) Null hypothesis

(2) Alternative hypothesis or directional hypothesis

The null hypothesis is denoted by H_o it is an hypothesis of equality i.e. it is not partial e.g. there is No significant difference between the performance of male and female student in mathematics.

The Alternate hypothesis is denoted by H_i which says, there is a significant difference between the performance of male and female students.

I drink beverage, if you don't see 7up then you can bring coke.

While alternate hypothesis If I want ponded yam and you are bringing rice. Then you are a fool.

Language of Hypothesis

The language of hypothesis would help us to know the data and the statically tools to use, e.g. is it chi-square, regression e.t.c. And it would enable us to know the type of conclusion to draw.

When we are looking for relationship between 2 variables, we can either use chi-square or correlation but when it involves frequency count. It must be chi-square e.g. sex and smoking habit.

But if X^2 chi-square calculated is less than X^2 square table. It shows that the result is not significant which is in agreement with your null hypothesis will be accepted.

X_c^2 = if X_c^2 calculate value is greater than X_t^C it shows that the result is significant result which is against null hypothesis: reject null hypothesis

The change there is that we don't recognize negative. We discard it and accept positive n<u>o</u>

e.g $t_c = -3.4$

$t_t = 1.5$

Disregard -3.4 and take 3.4.

Smoking Habit

Sex	Smoking	Non-smoking	Total
M	3(3.3)	2(1.7)	5
F	7(6.7)	3(3.3)	10
Total	10	5	15

$\frac{RW \times CT}{GT} = \frac{5 \times 10}{15} = \frac{50}{15}$

$Fe = \frac{RT \times GT}{GT} = \frac{5 \times 5}{15} = \frac{25}{15}$

$\frac{10 \times 10}{15} = \frac{100}{15}$

Calculate

$X_c^2 = \Sigma \frac{(fo - fe)}{Fe}$

$= \frac{(3-3.3)^2}{3.3} + \frac{(2-1.7)^2}{1.7} + \frac{(7-6)^2}{6.7} + \frac{(3-3.3)}{3.3}$

$= \frac{(0.3)^2}{3.3} + \frac{(0.3)^2}{1.7} + \frac{(0.3)^2}{6.7} + \frac{(-0.3)^2}{3.3}$

$= \frac{0.09}{3.3} + \frac{0.09}{1.7} + \frac{0.09}{6.7} + \frac{0.09}{3.3}$

$= 0.027 + 0.053 + 0.013 + 0.027$

$X_c^2 = \underline{0.12}$ ans

Degree of freedom = raw-1 and column -1 (r - 1) x (c - 1)

Some limitations of chi-square

1. We can only use chi-square neither in frequency data nor for scale score.
2. No expected value in each cell must be less than 5 (in this case we use yate's correlation).
3. The sum of the expected frequency must be equal to the sum of the observed frequencies.
4. The frequencies must be independent (mutually exclusive).
5. When df is 1, we apply the correlation continuity.

EXERCISE

Assume that we want to find out whether an anxiety and social interaction are related.

Anxiety

	No	Low	Medium	High	Total
Interaction	3	4	8	10	25
No interaction	9	11	8	3	31
Total	12	15	16	13	56

(1) state the hypothesis

(2) Compute the chi-square the increase in anxiety leads to increase in social interaction. That is, there significant different between anxiety and social interaction?

Student T-test

Student's t-distribution is the statistic in calculating the probability associated with Ho. The t is a statistic generally applicable to a normally distributed Radom variable where the mean is known (or as we shall see, assumed to be known) and the population variance is estimated from a sample.

X	Y	x^2	y^2
5	3	25	9
1	4	1	16
2	3	4	9
4	2	16	4
	1		1
12	13	46	39

$\sum x = 12$ $\sum y = 13$

$\sum x^2 = 46$ $\sum y^2 = 39$

$\bar{X} = 12/4 = 3$, $\bar{Y} = 13/5 = 2.5$

$$t_C = \frac{x - y}{\sqrt{\left(\frac{\sum x^2 + \sum y^2}{N_1 + N_2 - 2}\right)\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

Where $\sum x^2 = \sum x^2 - (\sum x)^2$

$$= 46 - \frac{12^2}{4} \text{ N1}$$

$$= 46 - \frac{144}{4}$$

$$\sum x^2 = 46 - 36$$

$$= 10$$

$$\sum y^2 = \sum y^2 - \frac{(\sum y)^2}{N_2}$$

$$= 39 - \frac{13^2}{5}$$

$$= 39 - \frac{169}{5}$$

$$= 39 - 33.8$$

$$\sum y_2 = \underline{5.2}$$

$$t_c = \frac{3 - 2.6}{\sqrt{\left(\frac{10 + 5.2}{4 + 5.2}\right)\left(\frac{1}{4} + \frac{1}{5}\right)}}$$

$$= \frac{0.4}{\sqrt{\left(\frac{15.2}{7}\right)\left(\frac{9}{20}\right)}}$$

$$= \frac{0.4}{\sqrt{(2.17)(0.45)}}$$

$$= \frac{0.4}{\sqrt{0.9765}}$$

$$= \frac{0.4}{0.98818}$$

$$tc = 0.4048$$

$$tc = 0.41$$

When table value is greater than calculated value.
It shows the result is significant.
You accept null hypothesis. If table value is less reject the null hypothesis.

Summation Notation (Σ)

$\underline{7}$ = 2+3+4+5+6+7 = 27
Σ^{r}
2

$\underline{7}$ = 2^n+3^n + ------$+7^n$
Σr^n
r =2

$\underline{5}$ $\left(\Sigma^3 K^n\right)$ = $1^3+2^4+3^5$
Σ = 1 + 16 + 2 +3
n=3 k=1 = 260 ans.

EXERCISE

$\underline{10}$ (K^2+5)
Σ_7

$\underline{6}$ f_1 x f_1 t
Σ
i=1

$\underline{6}$ $(x+y)^k$
Σ
K = 4

When dealing with difference then you either use t test, analysis of variance (ANOVA) F- ratio, Analysis of co-variance (ANCOVA) etc. if the difference is between 2 variables. It must be t-test but when it is between 2 or more variables then it must be ANOVA is more preferable.

However, it should be noted that ANOVA can also handle 2 variables but t-test cannot handle more than 2. e.g. any Lecturer teaching post-graduate students should not be lower than Snr-lecturer. But if you have any Lecturer lower than the prescribed cadre can only teach Pre-degree, undergraduates. A Professor can teach Pre-degree, under-graduates, and Post-graduates students. But, if you are doing experimental work that involves pre-test and post-test. Then it must be ANCOVA, with the pre-test acting as co-varies. If you are dealing with prediction then it must be regression analysis.

Outlines

1. The statistics and its distribution.

2. Establishing confidence interval.

3. Test of difference of mean.

   a. Uncorrelated } Simple

   b. correlated }

4. Test of difference of variance correlated.

Notations and their meanings

M = population mean

σ^2 = variance

M = group mean

S.σ = standard deviation

Sm = standard deviation of the simple mean = standard Error $\frac{M-Ч}{\Sigma}$ is a constant

t = m-u varies. This was developed by W.S. Gusset. (Who used the Sm pseudonym "student") t is tabulated for various degrees of freedom, usually from 1 to 30. For degrees of freedom larger than 30.S^2 (the sample variance) is a sufficiently reliable estimate of σ^2 (population variance) so that the distribution of t is almost identical to that of Z normal distribution.

$Sm = \frac{\sigma}{\sqrt{n}}$ = standard error

$t = \frac{M-u}{\sigma} \times (\pm\sqrt{n})$

$tsm = (M - u)(\pm\sqrt{n})$

$u = M \pm tsm$

Example 1

Suppose M = 32, Sm = 5, n = 7, calculate the population mean.

Solution

$u = m \pm tsm$

t = 2.447 at 6 df with 95% confident interval

= 32 $\pm$ (2.447) 5 = (32 – 12,235) or (32 + 12.235)

= 19.77 or 44.24 = (19.765, 44.235) = (19.37, 44.24)

F-ratio

X_1	x_2	x_3	x_1^2	x_2^2	x_3^2
2	3	1	4	9	1
4	1	3	16	1	9
1	4	2	1	16	4
5	2	4	25	4	16
Σ_{x1}12	Σ_{x2}10	Σ_{x3}10	Σ_{x1}46	Σ_{x2}30	Σ_{x3}30

$4X_t^2$ = total of x =32

$4X_t^2$ = 46 + 30 + 30 = 106

Step I = sum of squares total.

$$ss_t = \Sigma x_t^2 - \frac{(\Sigma x_t)^2}{N}$$

N = N_1 =4

N_2 =4

=4+4+4+4=12

N_3 =4 N =N_1+N_2+N_3

b = between $106 - \frac{32}{12}$

t = total

$$106 - \frac{1024}{12}$$

$$106 - 15.33$$

SS_t = 20.67 = 20.7

Step II = Sum of squares between groups

$$SS_b = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - \frac{(\Sigma x)^2}{N}$$

$$\frac{12^2}{4} + \frac{10^2}{4} + \frac{10^2}{4} - \frac{(32)^2}{12}$$

$$\frac{144}{4} + \frac{100}{4} + \frac{100}{4}$$

$$36 + 25 + 25 - \frac{1024}{12}$$

$$86 - 85.33$$

$$SS_b = 0.67$$

Step III:- Sum of squares within groups

$$SS_w + SS_b = SS_t$$

$$SS_w + 0.67 = 29.67$$

$$SS_w = 20.67 - 0.67$$

$$SS_w = 20$$

$$SS_w = \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} + \sum X_2^2 - \frac{(\sum x_2)^2}{n_2} + \sum x_3^2 - \frac{(\sum x3)2}{n^3}$$

$$= \left(46 - \frac{12^2}{4}\right) + \left(30 - \frac{10^2}{4}\right) + \left(30 - \frac{10^2}{4}\right)$$

$$= \left(46 - \frac{144}{4}\right) + \left(30 - \frac{100}{4}\right) + \left(30 - \frac{100}{4}\right)$$

$$= 10 + 5 + 5$$

$$= 20$$

Step4:- Mean squares of between groups

$$msb = \frac{SSb}{dfb}$$ degree of freedom = 3 – 1 = 2

$$msb = \frac{0.67}{2}$$

$$msb = 0.335$$

Step5 = Mean square of within groups

$$msb = \frac{ssw}{d.fw}$$ = no of case - groups

= 12 – 3 = 9

$$msw = \frac{20}{9}$$

$$msb = 2.222$$

Step 6:- Between / within

$$\frac{msb}{msw} = \frac{0.335}{2.222}$$

F - calculated = 0.15

Source of variance

Source of variance	ssb	df	msw	fc	ft
Between groups ssb	0.67	2	0.335	0.15	
4.26					
Within groups ssw	20	9	2.222		
Total	20.67	11			

You accept the null-hypothesis since f- table value is higher than f –calculated value.

0.05 – mean the degrees at which you say are not perfect. But you can still make mistake.

Regression Analysis

Y = a + bx

Dependent intersect independent

Y = a x bx -------- (1)

Multiply by x

$Xy = ax + bx^2$ ------- (2)

Sum (1) x (2)

$\sum y = an + b\sum x$ ------- (3)

$\sum xy = a\sum + a\sum x^2$ ------- (4)

Re – arrange

$an + b\sum x = \sum y$

$a\sum x + b\sum x^2 = \sum x$

Using crammer's method

$$D = \begin{pmatrix} n^a & \sum x^b \\ \sum x & \sum x^2 \end{pmatrix} = n\sum x^{2-} - (\sum x)^2$$

$$D^a = \begin{pmatrix} \sum y & \sum x \\ \sum xY & \sum x^2 \end{pmatrix} = \sum y \ \sum x^2 - \sum x \sum xy$$

$$:- a = \frac{D^a}{D} = \frac{\sum x^2 \ \sum x \ \sum xy}{n\sum x^2 - (\sum x)^2}$$

$$D^b = \begin{pmatrix} n & \sum y \\ \sum x & \sum xy \end{pmatrix} = n\sum xy - \sum x\sum y$$

$b = \frac{D^b}{D} = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$

Substitutes in equation (1)

Y = a + bx

Correlation

Derived method

x	y	x=x – x	y = y – y	x y	x^2 -	y^2
2	3	-1	1	-1	1	1
4	1	-1	-1	-1	1	1
3	2	0	0	0	0	0
5	1	2	-1	-2	4	1
1	3	-2	1	-2	4	1
15	10			-6	10	4

$\overline{X} = \frac{15}{5} \frac{10}{5}$

$\overline{X} = 3 \quad \overline{x} = 2$ rxy = $\frac{\sum xy}{\sqrt{Ex^2 . Ey^2}}$

Maximum = $\frac{-6}{\sqrt{10 \times 4}}$

You can get is 1

$= \frac{-6}{\sqrt{40}}$

$= \frac{-6}{6.325}$

= - 0.949

Rxy = -0.95

Correlation spearman Brown's method

X	y	Rx	Ry
2	3	4	1.5
4	1	2	4.5
3	2	3	3
5	1	1	4.5
1	3	5	1.5

The highest figure takes 1st position and like that not when you have 2 numbers and you have only 1 chair

$$\frac{1+2}{2} = \frac{3}{2}$$

$$= 1.5$$

And the other forfeits the 2nd position so we talk about 3 now.

= 1, 2, 3, 4+5

$$= \frac{4+5}{2} = \frac{9}{2}$$

= 4.5

You now say d = deviation

$1.5 - 4 = 2.5^2$

$= 6.25$

$4.5 - 2 = 2.5^2$

$= 6.25$

6 = is constant

d^2

6.28

6.25

0

12.25

$\frac{12.25}{37}$

$$rxy = 1 - \frac{6\sum d^2}{N(N^2-1)}$$

$$= \frac{1-6\ (37)}{5\ (5^2 - 1)}$$

$$= 1 - \frac{222}{5(24)}$$

$$1 - \frac{222}{120}$$

$= 1 - 1.85$

Rxy = -0.85 ans

Pearson product Moment correlation method

X	y	xy	x^2	y^2
2	3	6	4	9
4	1	4	16	1
3	2	6	9	4
5	1	5	25	1
1	3	3	1	9
15	10	24	55	24

$$rxy = \frac{N\sum xy - \sum x \sum y}{\sqrt{N\sum N^2 - (\sum x)^2 - N(\sum y^2) - (\sum y)^2}}$$

$$= \frac{5\,(24) - (15)\,(10)}{\sqrt{5(55) - 15^2)\,(24) - 10^2)}}$$

$$= \frac{120 - 150}{\sqrt{(275 - 225)\,(124 - 100)}}$$

$$= \frac{-30}{\sqrt{(50)(20)}}$$

$= - 0.238$

$$\frac{-30}{\sqrt{1000}}$$

$$= \frac{-30}{31.62}$$

$= -0.948$

rxy = -0.95 ans

EXERCISE

1. You have written a computer programme to generate random number in the range 0 to 9. The programme was successfully run to produce 500

digits and the following distribution of the digit resulted.

Digit	0	1	2	3	4	5	6	7	8	9
Observed freq.	40	36	28	62	58	60	34	70	40	72

Are you satisfied that your method of generating random number is satisfactory?

2. Two horse, X and Y were tested according to the time (in seconds) taken to run a particular track with the following results.

Horse X 33 29 30 27 28 33 29 and 30

Horse Y 26 29 30 25 26 and 28

Analyses the above data and report whether or not you can decimate between the running times of the horses.

3. In a math's test, the average score for 100 boys taking the test was 72 with a standard deviation of 10, for 120 girls the average score was 60 and the standard 12. Test the hypothesis that boys are better in Math's than girls are.

4. The following are the JAMB scores and POST-JAMB examination of 10 JAMB candidates in mathematics.

S/N	JAMB	POST
1	60	58
2	72	69
3	50	56
4	80	78
5	60	59
6	20	23
7	68	65

8	75	72
9	48	50
10	40	43

Test whether there is a significant difference in the performance of these ten candidates in JAMB and post JAMB

5. In an aptitude test was conducted for Administrative and clerical officers, the result is as follows:

	Mean	sample standard Deviation	sample size
Administrative Officers	62	3	5
Clerical Officers	56	4	10

Is there any evidence of significant difference in the means of the two groups.

6. In an aptitude test, was conducted for administrative and clerical officers, the result is as follows:

	Mean	Sample standard Variation	sample size
Administrative Officers	62	3	5
Clerical Officers	56	4	10

Is there any evidence of significant in the means of the two groups.

7. Solve the following using derived correlation method.

X	15	10	20	14	11	9	10	7	6	1
y	12	13	10	15	10	15	13	9	7	8

8. Solve the following using Spearman Brown's correlation method

X	15	10	20	14	11	9	10	7	6	1
Y	2	4	3	7	14	13	12	10	20	11

9. Solve the following using Pearson product moment correlation

X	12	13	10	15	10	15	13	9	7	8
Y	2	4	3	7	14	13	12	10	20	11

10. Solve the following using student t-test

X	5	1	2	4	
Y	3	4	3	2	1

11. Solve the following using F -ratio

X	2	4	1	5
Y	3	1	4	2

Let say that there is a survey in mathematic criterion approach non-reference approach is used. Hypothesis: the approach does not produce any difference.

	Criterion	Non-Ref.	Total
0_1	20	10	30
0_2	4	9	13
Total	24	19	43

CHAPTER SEVEN

ANALYSIS OF SCHOOL ENROLMENT.

In diagnosis school environment. There things that are of vital importance.

- Enrolment trend.
- Enrolment ratio.
- Enrolment rate.

Enrolment trend: It enables us to know absolute increase or decrease in enrolment as well as growth rate of enrolment. This can be seen from two angles. Absolute increase over a given period of time or through the growth rate of enrolment over a given period of time. However growth rate of enrolment is found to be more useful in calculating the enrolment trend than the other method:

2001 / 2002 session: Table A

Schools	Boys	Girls	Total
A	1820	1600	3420
B	1680	1220	2900
C	875	-	875
D	1320	1480	2800

Total	5695	4300	9995

2002 / 2003 session Table B

Schools	Boys	Girls	Total
A	2040	1820	3860
B	1850	1325	3175
C	1250	-	1250
D	1680	1840	3480
Total	6780	4988	11,765

2002 / 2003 - 2040
2001 / 2002 - 1820
220

Increase = Enrolment 1 – Enrolment 2
= 2040 - 1820 = 220

(2) Growth rate of enrolment

$$G = \frac{Et + 1 - Et}{et} \times \frac{100}{1}$$

$$GT = \frac{2040 - 1820}{1820} \times \frac{100}{1}$$

$$= \frac{220}{1820} \times \frac{100}{1}$$

$= 12.087\%$

$= \underline{12.09\%}$

$$\frac{1820 - 1600}{160} \times \frac{100}{1}$$

$= 13.75 \qquad = \underline{13.8\%}$

Enrolment ratio: Is define as the ratio between the number of pupil's enroll of a given age of at a given level of education and the size of the population in that given age. The enrolment ratio can be calculated for a given sex weather boy or girl for private public schools for different region in the country.

There are 3 ways of measuring enrolment ratio.

1. Overall enrolment ratio : This is otherwise known as general enrolment ratio or crude enrolment ratio and it is the ratio between total school age pupils enroll in the education sector and total population of age group within the society

$$Oer = \frac{E(t)}{Pt} \times \frac{100}{1}$$

Overall enrolment ratio is adequate for distant study of enrolment development in a country. But it is weak because it does not inculcate the no of pupils or students enroll at each level of educational system. It is

- Et = total school enrolment at all levels
- Oer = overall environment
- Pt = total population of school age in year t

Therefor crude for example if enrolment figure at all levels of educational system in Nigeria is 212, 168, 240, as total school age 392, 407, 322 pupils

$$Oer = \frac{212,\ 168\ 240}{392\ 407\ 322} \times 100$$

$= 54.068$

$= \underline{54.07}$

Only 54-07% are enrolling in school.

b. level specific enrolment ratio: this is also known as level enrolment ratio. It is the most commonly used indicator of development is two types.

1. Gross level enrolment ratio

2. Net level enrolment ratio

1. Gross level: It relates total student enrolment a specified educational level (regardless of those enroll) to the students' population that most enroll for that school level. It is denoted by gross enrolment.

$$GLER = \frac{Emt}{Pgt} \times \frac{100}{1}$$

Emt enrolment at school level weather primary secondary and tertiary.

Pgt = population of those that supposed to be in secondary school.

2. Net level enrolment ratio: This includes the enrolment in school at a particular year and the number of student who is of the same age group and the level. Discard under age and ignore over age 11 – 16 are to be in secondary schools.

c. Enrolment rate: the term rate and ratio are binomials size of one number to rate is a special ratio often used in the analysis of floors. It indicates the relative frequency of the occurency of a population.

Thus we talk of birth, death, promotion/ transition rates and so no white enrolment rate are vital policy variable that can affect that determination of future enrolment. Enrolment ratio is indicator of countries educational growth enrolment rates are very essential ingredient in the flow model of enrolment projection.

The non-schooling gap.

It is the difference between the estimated population of appropriate age group and the number enroll in the education sector carries ponding to that group. It represents the number of student who should be benefiting from education but are not actually there.

PROMOTION RATE.

This is also known as progression rate. It is the ratio of the total number of students that are promoted to the next higher level in a given academic year.

To the total number of students that enroll in the same class in the previous academic session. This is symbolically describe as thus pg + 1 to power + equal no of student promoted to next class g + 1 class year + pupils one.

Grade class I

Year 2000.

Egt = total number of student that are enroll in year t = 2000

Pgt + 1 = total no of student enroll in the former class g in the previous year t.
Pt+1
g + t = class 2 in years 1

Repetition rate: This is define as the ratio of number of student that are repeating a specific class in a given academic year to the total number of students enroll in the same class levels in the previous academic years. It is express as:

$$R_g^t = \frac{R_g^t + 1}{Egt} \times \frac{100}{1}$$

Rgt + 1 = No of student that are repeating the same class level in year t + 1 = Egt = enrollment in year t class one in year 2000

4. Dropout rate: It is the ratio of student that are not promoted or those repeated from a specific class level in a given year to the total number of student enroll in the same class level in the same academic session.

A child that is first in a class still dropout in a school what could be the factor

- Religious crisis
- Death.
- Physical disabilities
- Financial problem
- Parent separation (devours)
- Family problem
- Mobility of labour
- Sickness.

$$E_g^t - P_g^t = \frac{(P^{t+1}_{g+1} \times R_g^t = R_g^{t+1} \times 1)}{Et} \times 100$$

$$Egt = \frac{d_g^t}{Egt} \times \frac{100}{1} \qquad (2)$$

6. Wastage rate: this is the proportion of total enrollment accounted for by stagnation (repetition and dropout) express as a percentage.

Add dropout + repetition
= Wastage rate:

Remove promotion from 1

$$W_t = \frac{1-p_{g+1}}{Egt} \times \frac{100}{1}$$

Promotion = 78% = 0.78

Repetition = 12% = 0.12

Dropout = 10% = 0.10

Assuming that in Ekiti state in 2000/11 session

The number that got promoted to J.S.S. (2) 216/ 200

The total student enrollment was 282 400 student. The number that repeated J.S.S. 1 in 2011/12 56,120 and the number that dropout (withdraw) from the system 10,080. Calculate the Promotion, Repetition, Dropout and wastage rate.

(7) Transition rate: it is the rate at which student move from one level of education in a new academic session: e.g. primary - secondary - tertiary. It should be noted that while students are transiting from one level of education to another. The following might happen

1. Some may be repeating the final class

2. Some might have withdrawn from the level of education without completing it.

3. Some might have completed the circle and move to the next higher level

4. Some might even complete the system and join the labour force.

Mathematically transition rate is given as:

$$T_t^i = \frac{E^{i+1}_{t+1}}{E_t^i} \times \frac{100}{1}$$

i = represent final class in the lower level of education

t+1 = represent first class in the next higher level of education.

E_t^i = represent enrollment in the first class in the lower level of education

E^{i+1}_{t+1} = represent no of new entrants in the first class of the next higher level of education in year t + 1.

How many year in 2000 and how many transited to J.S.S. 1 in 2001.

Repeater $\frac{56.120}{282,400} \times \frac{100}{1}$

=19,8725

Dropout $\frac{10,080}{282,400} \times \frac{100}{1}$

=3.5694

Wastage = 19.8 + 3.5

= 23.5

Promotion $\frac{226,280}{282,400} \times \frac{100}{1}$

= 80.127

Promotion Rate.

$Pgt = \frac{P^{t+1}_{g+1}}{Egt} \times \frac{100}{1}$

$\frac{216,200}{282,400} \times \frac{100}{1}$

= 76.58%

Repetition rate.

$\frac{R_g^{t+1}.}{Egt} \quad \frac{56.120}{282,400} \times \frac{100}{1}$

= 19.87%

Dropout rate.

$Egt - (pgt = p_g^{t+1} + R_g^{t} = R_g^{t+1}) \times \frac{100}{1}$

$100 - \frac{(216,200+ 56,120)}{282,400} \times \frac{100}{1}$

$\frac{272,320}{282,400} \times \frac{100}{1}$

=96.43

100 – 96.43 =3.57%

$$Ed_gt = \frac{d_gt}{Egt} \times \frac{100}{1}$$

$$\frac{10{,}080}{282{,}400} \times \frac{100}{1}$$

$$=19{,}8$$

Wastage rate.

$$W_t \frac{(1 - p_{g+1}t + 1)}{Egt} \times \frac{100}{1}$$

Wastage = repetition + dropout

(19.87 + 3.57)%

23.44%

100 – Promotion

100 – 76.56%

= 23.44%

Cohort Analysis

It shows the extent to which educational system is able to use his import in the production of set of students in a particular level to the time they left that level of education. Cohort analysis is used to determine the internal efficiency of an educational system. When the educational system is able to see the pupils (student) through the system in the shortest possible period then the system is efficient. Given the information below prepare a cohort analysis showing the movement of pupils from J.S.S.1 to J.S.S 3 and calculate the wastage ratio of the school system.

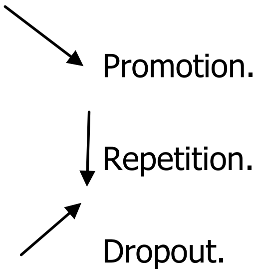
Entrant =1000 student

Promotion rate = 80%

Repetition rate = 15%

Dropout rate = 5%

- You don't admit student again.
- Students are to repeat twice before total withdrawer from the system.



	J.S.S.1	J.S.S.2	J. S.S.3	Output
2001	1000	800	640	512
2002	150	240	288	231
2003	23	54	86	69
2004				
2005				
Inputs	1,173	1094	1014	

$\frac{86 \times 80}{100} = 68.8$

$\frac{1000 \times 80}{100} = 800$ promotion $= \frac{69}{812}$

$\frac{1000 \times 15}{100} = 150$

$\frac{800 \times 80}{100} = 640$

$\frac{15 \times 100}{100} = 23$

Student input – years (input)

1173 + 1094 + 1014 = 3281

Input = 3281

Wastage ratio = Actual

Input – output ratio

Idea input – output ratio

Input – output = $\frac{\text{input}}{\text{Output}}$

Actual input – output ratio =

$\frac{\text{Actual input}}{\text{Actual output}}$

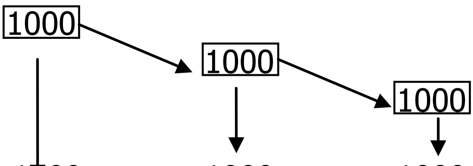
$= \frac{3281}{812}$

Actual input = 4.40 – output ratio

Ideal ratio.

J.S.S.1 J.S.S.2 J.S.S.3

2001



1000 + 1000 + 1000

Input = $\frac{3000}{}$

Output = 1000 = 3

Wastage ratio = $\frac{4.04}{3.00}$

Wastage = 1.35

$\underline{1} \leq Wr \leq 2$

Wastage ratio is greater than or equal to 1 or less than or equal to 2.

The closer the wastage ratio to 1 the move efficient is it. The fat away from 1 the less efficient the system is e.g. 1.77 is fending toward to 2 which mean they would use times two of the resources to train the children it is less efficient. While 1.35 is fairly efficient.

If the wastage ratio is equal to 1 that means the education system is perfectly efficient which is not attainable anywhere in the world.

Problem of wastage

1. Poor attitude of leaners toward learning e.g technological problem.
2. Death of personnel core subject teachers e.g physics, English, mathematics, or lack of teachers.
3. Library problem : lack of adequate text – books
4. Inadequate supervision both internal and external
5. Desire to make quick money
6. Parental influence- parents encouraging their children in examination mal-practice by buying questions for them.
7. Inadequate funding
8. Poor instructional materials
9. Incessant strike action e.g. Waec is a yearly exam if there is a strike that lasted for 6 months. Obviously, it would have the performance of the students.

Co-efficiency of efficiency equal to wastage ratio multiply by $\frac{100}{1}$

$$= \frac{1}{1.35} \times \frac{100}{1}$$

$$= \frac{100}{1.35}$$

Co-efficient of = 74.07%

Efficient is internally efficient.

CHAPTER EIGHT
ENROLMENT PROJECTION

A projection is a conditionally statement about the future. It is the elaboration of the effect in the future of making a set of assumption about trend in the parameters characterizing the educational system.

A projection does not necessary offers the most probable (in some person's judgment) outcome rather its main function is to demonstrate to the decision maker the result which follows from carrying some of the parameters (or from leaning them unchanged). Depending on the desirability of the decision maker of me projected outcome he may intervene with policy changes to affect the underline trends.

One major way of looking at enrolment projection in educational management is compounding and discounting.

Compounding and discounting are technics used in comparing the size of variable at different pout in times.

Compounding deals with finding the future worth of present resources and other variable such as population, enrolment e.t.c. growing by geometrical progression. Discounting is concern with the calculation of present worth of a future amount.

It is the opposite of compounding as it looks from the future to the present. Compounding is very useful in education planning as it helps to solve problems relating targets and projection of enrolment, teacher's demand and supply, cost, finance e.t.c.

For this purpose educational planners must have the mastery of this technique.

Assuming a school has an enrolment of 1,000 in year 2000. And the school is expected to grow at the rate of 6% for the next 10 years.

Years	Increase during year t	Enrolment in yrs t – 1	Enrolment in year t
2000	-	-	1000
2001	1000x0.06	1000	1060
2002	1060x0.06	64 1060	1124
2003	1124x0.06	67 1124	1191
2004	1191x0.06	71.46 1121	1262
2005	1262x0.06	76 1262	1338
2006	1517x0.06	80 1517	1418
2007	1001x0.06	85 1001	1503
2008	1001x0.06	96 1925	1593
2009	100x0.06	101 1936	1689
2010	100x0.06	205 2052	1790
			1791

To 2 s.f. If the last is up to five then approximate.

The computation of the figures as shown in the table above is time watching, time consuming and as well as it is difficult to calculate. Since we know that the annual growth rate is 6% and the time interval is 10 years, the enrolment for each year can be obtained directly by multiplying the enrolment of the preceding year by a factor which is always the same, 1 + rate of increase = 1 + 0.06

$= 6\% \quad \frac{6}{1000} = 0.06$

This is the characteristics of a geometric progression to this end. The information presented in table 1 can be calculated more directly as presented in table 2.

Year	Enrolment (t-1)	Calculate	Enrolment
2000	-	-	1000
2001	1000	1000x1.06	1060
2002	1060	1060x0.06	1124
2003	1124	1124x0.16	119
2004	1191	1191x0.06	1262
2005	1262	1262x0.06	1338
2006	1338	1338x0.06	1419
2007	1419	1419x0.06	1504
2008	1504	1504x0.06	1594
2009	1594	1594x0.06	1690
2010	1690	1690x0.06	1791

In the example given above the experiment of 1.06 between year 2000 and 2010 is 10. Therefore in mathematical terms, this can be stated as $1 + r^{10}$ rate of increase $(1 + 0.06)^{10}$ = 6% = 0.06

It thus, implies that to obtained enrolment in 2010. The enrolment of 2000 multiply by $(1.06)^{10}$ to this end the for mincers emerge

$$En = Eo \times (1 \times r)^n$$

Where En = enrolment in the final year 2010.

Eo = enrolment in the initial year = 2000

r = rate or increase 6%

n = no of years 10

The interval section 2000 and 2010 = 10 years

Computer method

Inserting the figure from our example

En = $1000 \times (1.06)^{10}$

1000×1.790847

En = $1791\ (1.06)^{10}$

out of the four variables in the above equation.

En, Eo, r, n. we only know 3. The fourth can be worked out using the logarithms table.

$$En = Eo \times (1 + r)^n \text{ ------- } (1)$$

$$\frac{\text{Log En}}{\text{Log Eo}} = n \log (1 + r) \text{ ------- } (2)$$

Log En = log Eo + n log (1 + r) ------- (3)

Using equation 3. We have log En = log 1000+ 10 log (1 + 0.06) log 100 (3)

Log En = log 1000 + 3 + 10 x 0.0252

Antilog = 3.253

=1,7910

=1791

Supposing we know that at the initial stage that enrolment in year 2000 was 1,000 and in 2010, it was 1,791 and ask 1,791 and we are ask to find rate of growth.

$$\frac{\text{Log En}}{\text{Log Eo}} = n \log (1+r)$$

$$n \log (1+r) = \frac{\text{log En}}{\text{log En}}$$

$$10 \log (1+r) = \frac{\text{log 1791}}{\text{Log 1000}}$$

Divide both side by 10

$$3 \log 1006 = \frac{\text{log 1,791}}{\text{Log 1000}}$$

$$\text{Log } (1+r) = \frac{1.791}{10}$$

$$\text{Log } (1+r) = \frac{0.2531}{10}$$

Anti–log 0.2531

$\text{Log}(1+r) = 1.060$

$1+r = 1.060$

$r = 1.060 - 1$

$r = 0.060 \times 100$

$r = 6\%$

Supposing we want to know how long it will take the enrolment of 1,000 to reach 1,791 if it grows at the rate of 6% per annum.

$$\frac{\text{Log En}}{\text{Log Eo}} = n \log(1+r) \text{ ----- } (2)$$

$$n \log(1+r) \frac{\log \text{En}}{\text{Log En}}$$

$$10 \log(1+r) \frac{1791}{1000}$$

$10 \log(1+0.06)\ 1.791$

$n \log(1+0.06) \log 1.791$

$n \log 1.06 = \log 1.791$

$n = (0.02531) = 0.2531$

$$n = \frac{0.2531}{0.02531}$$

$n = 10$

Discounting factor

It is the reciprocal of compounding factor

$$\frac{1}{(1+r)^n}$$

$$E_o = En \times \frac{1}{(1+r)^n}$$

It wants to look at project want of future amount.

$$E_o = 1791 \times \frac{1}{(1+0.06)^{10}}$$

$$1791 \times \frac{1}{(1.06)^{10}}$$

$$1791 \times \frac{1}{1.791}$$

1791×0.558

= 1000

Q_4 = 2002 / 2003

902, 200

2,200 – over age.

4800 – under age.

1,322,000

(a) 902, 200 – (2200 + 4800)

902,200 – 7000

=895,200

$\frac{895,200}{1322000} \times \frac{100}{1}$

= 67.72%

(b) $\frac{902,200}{1322000} \times \frac{100}{1}$

=67.72%

(16^B) 2005 / 2006

J.S.S. 2 = 326,463

Promoted to J.S.S 3 247, 403

Repeated of J.S.S 2

2006 / 2007 = 69.060

(1) Repetition $= \frac{69060}{326463} \times \frac{100}{1}$

= 12.15%

Dropout = promotion + repetition

= total enrolment

Compounding

$En = Eo \times (1+r)^n$

$En = 1000 \times (1+0.05)^{10}$

$En = 10,000 \times (1.05)^{10}$

10,000 x 1.63889

=16,300

CHAPTER NINE
MODEL

A model can come in many shapes, sizes, and styles. It is important to emphasize that a model is not the real world but merely a human construct to help us better understand real world systems. In general all models have an information input, an information processor, and an output of expected results. Simplifying assumptions must be made;

- boundary conditions or initial conditions must be identified;
- the range of applicability of the model should be understood.

Descriptive Models

A descriptive model describes logical relationships, such as the system's whole-part relationship that defines its parts tree, the interconnection between its parts, the functions that its components perform, or the test cases that are used to verify the system requirements. Typical descriptive models may include those that describe the functional or physical architecture of a system, or the three dimensional geometric representation of a system.

Analytical Models

An analytical model describes mathematical relationships, such as differential equations that support quantifiable analysis about the system parameters. Analytical models can be further classified into dynamic and static models. Dynamic models describe the time-varying state of a system, whereas static models perform computations that do not represent the time-varying state of a system. A dynamic model may represent the performance of a system, such as the aircraft position, velocity, acceleration, and fuel consumption over time. A static model may represent the mass properties estimate or reliability prediction of a system or component.

Hybrid Descriptive and Analytical Models

A particular model may include descriptive and analytical aspects as described above, but models may favor one aspect or the other. The logical relationships of a descriptive model can also be analyzed, and inferences can be made to reason about the system. Nevertheless, logical analysis provides different insights than a quantitative analysis of system parameters.

Domain-specific Models

Both descriptive and analytical models can be further classified according to the domain that they represent. The following classifications are partially derived from the presentation on *OWL, Ontologies and SysML Profiles: Knowledge Representation and Modeling* (Web Ontology Language (OWL) & Systems Modeling Language (SysML)) (Jenkins 2010):

- properties of the system, such as performance, reliability, mass properties, power, structural, or thermal models;
- design and technology implementations, such as electrical, mechanical, and software design models;

- subsystems and products, such as communications, fault management, or power distribution models; and
- system applications, such as information systems, automotive systems, aerospace systems, or medical device models.

The model classification, terminology and approach is often adapted to a particular application domain. For example, when modeling organization or business, the behavioral model may be referred to as workflow or process model, and the performance modeling may refer to the cost and schedule performance associated with the organization or business process.

A single model may include multiple domain categories from the above list. For example, a reliability, thermal, and/or power model may be defined for an electrical design of a communications subsystem for an aerospace system, such as an aircraft or satellite.

What is an Interactive Lecture Demonstration?

Interactive Lecture Demonstrations introduce a carefully scripted activity, creating a "time for telling" in a traditional lecture format. Because the activity causes students to confront their prior understanding of a core concept, students are ready to learn in a follow-up lecture. Interactive Lecture Demonstrations use three steps in which students:

1. **Predict** the outcome of the demonstration. Individually, and then with a partner, students

explain to each other which of a set of possible outcomes is most likely to occur.

2. **Experience** the demonstration. Working in small groups, students conduct an experiment, take a survey, or work with data to determine whether their initial beliefs were confirmed (or not).
3. **Reflect** on the outcome. Students think about why they held their initial belief and in what ways the demonstration confirmed or contradicted this belief. After comparing these thoughts with other students, students individually prepare a written product on what was learned.

Why Use Interactive Lecture Demonstrations

Research shows that students acquire significantly greater understanding of course material when traditional lectures are combined with interactive demonstrations. Each step in Interactive Demonstrations--Predict, Experience, Reflect--contributes to student learning. Prediction links new learning to prior understanding. The experience engages the student with compelling evidence. Reflection helps students identify and consolidate that they have learned.

How to Use Interactive Lecture Demonstrations in Class

Effective interactive lecture demonstrations require that instructors:

- Identify a core concept that students will learn.

- Chose a demonstration that will illustrate the core concept, ideally with an outcome different from student expectations.
- Prepare written materials so that students can easily follow the prediction, experience and reflection steps.

In simplest terms, a mathematical model is an abstraction or simplification that allows us to summarize (describe) a system. Once you have a mathematical model you have a list of inputs and a list of outputs and some sort of definite algorithm that tells you what the outputs will be given the inputs. Once we agree on that definition of what a mathematical model IS then we can talk about your question and its answer or answers.

Some of the benefits of building and using mathematical models:

- Ability to predict system behavior
- A clear idea of the important inputs and outputs
- Ability to analyze anomalous behavior by comparing it to the model-predicted behavior

Some of the disadvantages of using mathematical models:

- The model may eliminate important predictive power by being too simple
- The model may be capable in certain circumstances, but not in others and the assumed conditions may not be obvious or understood by later users

CHAPTER TEN

EFFICIENCY IN EDUCATION

Educators often feel ambivalent about the pursuit of efficiency in education. On the one hand, there is a basic belief that efficiency is good and worthy goal; on the other hand, there is sense of worry that efforts to improve efficiency will ultimately undermine what lies at the heart of high-quality education. Part of the difficulty stems from a misunderstanding about the meaning of efficiency as well as from the legacy of past, sometimes misguided, efforts to improve the efficiency of educational systems. It is therefore useful to begin with a basic discussion of the efficiency concept.

The notion of efficiency applies to a remarkably large number of fields, including education. It is a disarmingly simple idea that presupposes a transformation of some kind. One can think in terms of what was in hand before the transformation, what was in hand after the transformation, and one can also think about the transformation process itself. The *before* elements are commonly referred to as ingredients, inputs, or resources while the *after* elements are called results, outputs, or outcomes. The transformation process is sometimes less obvious and can become confused with ingredients. For example, in an educational setting, a teacher can be thought of as an ingredient while teaching is an important part of the actual transformation process.

The concept of efficiency is often connected to a moral imperative to obtain more desired results from fewer resources. Efficiency needs to be thought of as a matter of degree. Efficiency is not a "yes/no" kind of phenomenon. It is instead better thought of in relative or comparative terms. One operation may be more efficient than another. This said, the more efficient of the two operations could become even more efficient. The quest for greater efficiency is never over, and this sense of a perennially unfinished agenda is one source of the generalized sense of anxiety that tends to surround the efficiency concept.

The Choice of Outcomes

If the goal is to obtain more desired results from fewer resources, then it is important to be clear about what is being sought. Society might have a very efficient system because a large amount of outcome is being obtained relative to the resources being spent or invested, but if the outcomes are out of sync with what is truly desired, there is a real sense in which the system is not very efficient. Of course, this invites important questions about who gets to decide what counts as a desirable outcome, and in education there are longstanding and ongoing debates over what the educational system ought to be accomplishing.

In addition to reaching agreement about the mix of outcomes to pursue, there are important measurement issues to consider. An interest in efficiency is frequently accompanied by an interest in measuring magnitudes. If one is seeking more out of less, one frequently wants to know "how much more," and the result has been a boom in the efforts by educational psychologists and others to develop valid and reliable measures of the learning gains of students. Critics of efficiency analysis in education worry that ease of measurement can unduly influence the selection of the outcomes that the system will be structured to achieve. In other words, the worry is that the drive for efficiency will lead, perhaps inadvertently, toward the use of educational outcomes that are chosen more because they are easy to measure than because of their intrinsic long-term value for either individual students or the larger society. Standardized tests of various kinds have been relied upon as measures of the outcomes of schooling and have been criticized on these grounds.

Sometimes there is interest in the economic consequences of schooling, and this interest has prompted analysts to use earnings as a measure of schooling outcomes. A rich literature has developed in the economics of education where efforts have been made to estimate the economic rate of return to different levels and types of schooling. This is a challenging area of research because earnings are influenced by many factors and it is difficult to isolate the effects of schooling. The goal of this research is to capture the value added by schooling activities.

The relevance of the *value-added* concept is not limited to economists' studies of rates of return. Even in cases where the focus is on learning outcomes as measured by tests or other psychometric instruments, there are questions to answer about the effects of schooling activities relative to the effects of other potentially quite significant influences on gains in students' capabilities. Serious studies of the efficiency of educational systems measure educational outcomes in value-added terms.

Measurement issues also arise from the collective nature of schooling. The results gained from schooling experiences are likely to vary among individual students and this prompts questions about how best to examine the result for the group in contrast to an individual student. Is one primarily interested in, say, the average performance level, or is there a parallel and perhaps even more important concern with what is happening to the level of variation that exists across all of the students within the unit, be it a classroom, grade level within a school, a school, a district, a state, or a nation? The early research on educational efficiency in the 1960s placed a heavy emphasis on average test score results for relatively large units like school districts. More recent work demonstrates greater interest in measures of inequality among students. The standards-driven reform movement includes a considerable amount of rhetoric about all students reaching high standards; the analysis of efficiency presupposes an ability to move beyond the easy rhetoric to make clear decisions about how uniform performance expectations are for students.

In addition, there is an important distinction to maintain between the levels at which a system operates and the rate at which inputs are being transformed into outcomes. One can "get the outputs right" so that the desired items are being taught/learned in the correct proportion to one another. In such a case, gains in the understanding of mathematics are occurring in the correct proportion to, say, gains in language capabilities. But this says nothing about the absolute level at which the system is operating.

The naive view might be that the system should operate at 100 percent of its capacity, but this overlooks the fact that scarce resources are needed to operate at this level and that education is not the only worthy use of these precious resources. Policy-makers must make often difficult trade-off decisions about the level at which the educational system will operate relative to the level of other competing social services. The early twenty-first century is witnessing a considerable amount of debate over the proper level at which to set the educational system, often as part of an effort to define what counts as an "adequate" education.

With respect to outcomes, the goal is to reach agreement about (1) the relative mix of performance outcomes to realize; (2) the degree of uniformity of performance across students; and (3) the level of capacity at which the system should operate. In addition, there needs to be an ability to measure what is being accomplished.

INDEX

APPENDIX TABLE 2A

Areas in table combined for student's t distribution.*

EXAMPLE: To find the value of t which corresponds to and area of. 10 in both tails of the distribution combined, when there are 19 degree of freedom, look under the 10 column and proceed down to the 19 degree of freedom row, the appropriate f value is 1.729.

Area in both tails combined

Degrees of freedom	**.10**	**.05**	**.025**	**.01**
1.	6.314	12.706	31.821	63.657
2.	2.920	4.303	6.965	9.925
3.	2.353	3.183	4.541	5.841
4.	2.132	2.776	3.747	4.604
5.	2.015	2.571	3.365	4.032
6.	1.943	2.447	3.143	3.707
7.	1.895	2.365	2.998	3.449
8.	1.860	2.306	2.896	3.355
9.	1.833	2.262	2.821	3.250
10.	1.712	2.228	2.764	3.169
11.	1.796	2.201	2.718	3.106
12.	1.782	2.179	2.681	3.055
13.	1.771	2.160	2.650	3.012
14.	1.761	2.145	2.624	2.977
15.	1.753	2.131	2.602	2.947
16.	1.746	2.120	2.583	2.921
17..	1.740	2.110	2.567	2.898
18.	1.734	2.101	2.552	2.878
19.	1.729	2.093	2.539	2.861
20.	1.725	2.086	2.528	2.845
21.	1.721	2.080	2.518	2.831
22.	1.717	2.074	2.508	2.819
23.	1.714	2.069	2.500	2.807
24.	1.711	2.064	2.492	2.797
25.	1.708	2.060	2.485	2.787
26.	1.706	2.056	2.479	2.779
27.	1.703	2.052	2.473	2.771
28.	1.701	2.048	2.467	2.763
29.	1.699	2.045	2.462	2.756
30.	1.697	2.042	2.457	2.750
40.	1.684	2.021	2.423	2.604
60.	1.671	2.000	2.390	2.660
120.	1.658	1.980	2.358	2.617
Normal distribution	1.645	1.960	2.326	2.576

*Taken from Table III of Fisher and Yates, Statistical Table for Biological, Agricultural and Medical Research, published by Longman Group Ltd, London (previous published Oliver & Boyd, Edinburgh) and by permission of the authors and publishers.

APPENDIX 2B

TABLE 2 Critical values of t*

Level of significance for one tailed test
10. .05 .25 .01 .005 .0005

Level of significant for two-tailed test

Degrees of freedom	**.20**	**.10**	**.5**	**.02**	**.01**	**.011**
1.	3.078	6.314	12.706	31.821	63.657	636.619
2.	1.886	2.920	4.303	6.965	9.925	31.598
3.	1.638	2.353	3.183	4.541	5.841	12.941
4.	1.533	2.132	2.776	3.747	4.604	8.610
5.	1.476	2.015	2.571	3.365	4.032	6.859
6.	1.440	1.943	2.447	3.143	3.707	5.959
7.	1.415	1.865	2.365	2.998	3.449	5.405
8.	1.397	1.860	2.306	2.896	3.355	5.041
9.	1.383	1.833	2.262	2.821	3.250	4.781
10.	1.372	1.812	2.228	2.764	3.169	4.587
11.	1.363	1.796	2.201	2.718	3.106	4.437
12.	1.356	1.782	2.179	2.681	3.055	4.318
13.	1.350	1.771	2.160	2.650	3.012	4.221
14.	1.345	1.761	2.145	2.624	2.977	4.140
15.	1.341	1.753	2.131	2.602	2.947	4.073
16.	1.337	1.746	2.120	2.583	2.921	4.015
17..	1.333	1.740	2.110	2.567	2.898	3.965
18.	1.330	1.734	2.101	2.552	2.878	3.922
19.	1.328	1.729	2.093	2.539	2.861	3.883
20.	1.325	1.725	2.086	2.528	2.845	3.850
21.	1.323	1.721	2.080	2.518	2.831	3.819
22.	1.321	1.717	2.074	2.508	2.819	3.792
23.	1.319	1.714	2.069	2.500	2.807	3.767
24.	1.318	1.711	2.064	2.492	2.797	3.745
25.	1.316	1.708	2.060	2.485	2.787	3.725
26.	1.315	1.706	2.056	2.479	2.779	3.707
27.	1.314	1.703	2.052	2.473	2.771	3.690
28.	1.313	1.701	2.048	2.467	2.763	3.674
29.	1.311	1.699	2.045	2.462	2.756	3.659
30.	1.310	1.697	2.042	2.457	2.750	3.646
40.	1.303	1.684	2.021	2.423	2.704	3.551
60.	1.296	1.671	2.000	2.390	2.660	3.460
120.	1.289	1.658	1.980	2.358	2.617	3.373
A	1.282	1.645	1.960	2.326	2.576	3.291

APPENDIX 3

TABLE 3 Critical values of person r*

Level of significance for one-tailed test				
df	.05	.025	.01	.005
Level of significance for two-tailed test				
(=N-2; N= number of pairs)	**.10**	**.05**	**.02**	**.01**
1.	.988	.997	.9995	.9995
2.	.900	.950	.980	.990
3.	.805	.878	.934	.959
4.	.729	.811	.882	.917
5.	.669	.754	.833	.974
6.	.522	.707	.789	.834
7.	.532	.666	.750	.898
8.	.549	.632	.716	.765
9.	.521	.602	.685	.735
10.	.497	.576	.658	.708
11.	.476	.553	.634	.784
12.	.458	.532	.612	.661
13.	.441	.514	.592	.641
14.	.426	.497	.574	.623
15.	.412	.482	.558	.606
16.	.400	.468	.542	.690
17..	.389	.456	.528	.575
18.	.378	.444	.516	.516
19.	.369	.433	.503	.549
20.	.360	.423	.492	.537
21.	.352	.413	.482	.526
22.	.344	.404	.472	.515
23.	.337	.396	.462	.505
24.	.330	.388	.453	.496
25.	.323	.381	.445	.487
26.	.317	.374	.437	.479
27.	.311	.367	.430	.471
28.	.306	.361	.423	.463
29.	.301	.355	.416	.456
30.	.296	.349	.409	.449
35.	.275	.325	.381	.418
40.	.257	.304	.358	.393
45.	.243	.288	.338	.372
50.	.231	.273	.322	.354
60	.211	.250	.265	.325
70	.165	.232	.274	.302
80	.183	.217	.256	.283
90	.173	.205	.242	.267
100	.164	.195	.230	.254

From R.A. Fisher and F. Yates, "Statistical Table for biological, agricultural and Medical Research," 6th ed. Oliver and Boyd, Edinburg, 1963. Reproduced by permission of authors and publisher.

APPENDIX 4

TABLE 4 Critical values of chi-square*

Level of significance

df	**.20**	**.10**	**.05**	**.02**	**.01**	**.001**
1.	1.64	2.71	3.84	5.41	6.63	10.83
2.	3.22	4.61	5.99	7.82	9.21	13.82
3.	4.64	6.25	7.82	9.84	11.34	16.27
4.	5.99	7.78	9.49	11.67	13.28	18.46
5.	7.29	9.24	11.07	13.39	15.09	20.52
6.	8.56	10.64	12.59	15.03	16.81	22.46
7.	9.80	12.02	14.07	16.62	18.48	24.32
8.	14.03	13.36	15.51	18.17	20.09	26.12
9.	12.24	14.68	16.92	19.68	21.67	27.88
10.	13.44	15.99	18.31	21.16	23.21	29.59
11.	14.63	17.28	19.68	22.62	24.72	31.26
12.	15.81	18.55	21.03	24.05	26.22	32.96
13.	16.98	19.81	22.36	25.47	27.69	34.53
14.	18.15	21.06	23.68	26.87	29.14	36.12
15.	19.31	22.31	25.00	28.26	30.58	37.70
16.	20.46	23.54	26.30	29.63	32.00	39.25
17..	21.62	24.77	27.59	31.00	33.41	40.79
18.	22.76	25.99	28.87	32.35	34.81	42.31
19.	23.90	27.20	30.14	33.69	36.19	43.82
20.	25.04	28.41	31.41	35.02	37.57	45.32
21.	26.17	29.62	32.67	36.34	38.93	46.80
22.	27.30	30.81	33.92	37.66	40.29	48.27
23.	28.43	32.01	35.17	38.97	41.64	49.78
24.	39.55	33.20	36.42	40.27	42.98	51.18
25.	30.98	34.38	37.65	41.57	44.31	52.62
26.	31.80	35.56	38.89	42.68	45.64	54.05
27.	32.91	36.74	40.11	44.14	46.96	55.48
28.	34.03	37.92	41.34	45.42	48.28	56.89
29.	35.14	39.09	42.56	46.69	49.59	58.30
30.	36.25	40.26	43.77	47.96	50.89	59.70

*From R.A, Fisher and F Yates, "Statistical Table for Biological, Agricultural and medical Research," 6th ed. Oliver and Boyd, Edinburg, 1963. Reproduced by permission of authors and publisher.

*For df greater than 30, the value obtained from the expression may be used as t ratio
2X - 2df -1

APPENDIX 5

TABLE 5 Distribution of F

P = 0.5

n_1 n_2	1	2	3	4	5	6	8	12	14	a
1	161. 4	199. 5	215. 7	224. 6	230. 2	234.0	238. 9	243. 9	249. 0	254. 3
2	18.5 1	19.0 0	19.1 6	19.2 5	19.3 0	119.3 3	19.3 7	19.4 1	19.4 5	19.5 0
3	10.1 3	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	6.91	6.77	5.63
5	6.61	5.79	5.41	4.53	5.05	4.95	4.82	4.68	4.53	4.36
6	5.99	5.14	4.76	4.12	4.39	4.28	4.15	4.00	4.84	3.67
7	5.59	4.74	4.35	3.84	3.67	3.87	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	4.35	3.69	3.58	3.44	3.28	3.12	3.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	3.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	3.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.95	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	2.00	2.85	2.69	2.50	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.42	2.31
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.83	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.94	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.06	1.84
21	4.32	3.47	3.07	2.84	2.66	2.57	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.39	2.18	1.98	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.15	1.95	1.69
27	4.21	3.35	2.95	2.73	2.57	2.46	2.30	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39
12 0	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.63	1.61	1.25
A	3.84	3.99	2.60	2.73	2.21	2.09	2.94	1.75	1.52	1.00

Value of n1 and n2 represent the degrees of freedom associated with the larger and smaller estimates of variance respectively.