MODULE

TEACHING OF STATISTICS

CLASSES XI-XII

FOR

TRAINERS/TEACHERS

OF

INSERVCE TEACHER TRAINING PROGRAMME

DIRECTORATE OF CURRICULUM & TEACHER EDUCATION NWFP
ABBOTTABAD

Jan-Feb, 2003

Patron in Chief........................ UMAR FAROOQ
DIRECTOR

Guidance & Facilitated by........... Miss Shamim Sarfraz.
Deputy Director (Training)

Written by.............................. Mr. Abdus Salam,
SS (Statistics) GHSS, No.1,
Mansehra

Publisher............................... Directorate of Curriculum &
Teacher Education NWFP,
Abbottabad

Date of Publication.................... Jan-Feb, 2003

Printing.............................. Govt. Printing Press
NWFP, Peshawar

INDEX

FOREWORD

Directorate of Curriculum & Teacher Education, NWFP, Abbottabad is launching a comprehensive programme of in-service through out the province for all subjects/categories for the classes 6th to 12th under the title "Teacher Training Programme" scheme Improvement of Learning Environment For Quality Improvement for the year 2002-2004 as per policy of the Govt of NWFP, School & Literacy Department, Peshawar. The prime focus of this manual is training delivery effectively. There are two approaches to teacher's professional development, the carporate approach and the individual one, but in this guide book attempts are made to link the both practically.

To make the INSET Programme more effective and successful a "Survey Study" has been conducted to collect the feed back, needs of the learners, requirements of the teaching staff and desires of the concerned managers through, interview/questionnaires, survey form and classroom observation forms. Sample for the study was selected a few middle and secondary/Higher Secondary schools (Girls boys urban & rural).

The study was conducted by the Deputy Director (Training) and Subject Specialists of this Directorate.

In the light of above information & facts training strategy and instructional material has been developed to improve the learning environment for quality improvement through the innovative methodology and pedagogical techniques.

Instructional material consists on training manual for lead trainers & field trainers for delivery of training effectively and modules for each subject (VI – XII/Science/Arts) to facilitate the field Trainers as well as trainees of all categories (SS, SET (Science/Arts), CT, AT, TT).

The training manual comprises two parts, one for Subject Specialists training imparted by PITE and the other one for SET/CT/AT/TT training imparted by RITEs NWFP.

Umar Farooq
Director
Curriculum & Teacher Education
NWFP, Abbottabad

INTRODUCTION:

Now –a-days the science of statistics has shown its worth to such an extent that there is hardly any field in which its need is not felt. The educationists, the biologist, the astronomers, the economists, the businessmen and many others make use of statistical methods in their own sphere of activity. As a matter of fact there are millions of all over the world who don't know any thing about statistics, yet they make profuse use of statistics in their daily life. For example, when a man wishes to buy a scooter, he goes through the price list of various companies before taking the final decision, what he really aims at, is to know the average prices and the range of prices, without knowing any thing about these terms. Again we sometimes hear people saying, "As you sow, so shall you reap". In fact when they say this, they hint that there is a positive correlation between one's action and the consequences.

This clearly shows how important and universal the science of statistics is. Keeping in view, the importance of the subject, there is dire need of training the statistics teacher to teach statistics according to modern of methods of teaching. Keeping in view, the weaknesses of teacher education this module has been developed to reform the teacher through better in-service training program. It is hoped that that the use of module will increase the quality of instructional competence and will equip the teacher with background knowledge to reform his activities for effective teaching.

OBJECTIVES

After studying this module as a teacher you are expected to teach effectively by enabling students to:

Use various methods for calculating important descriptive values of the data.

Draw conclusion from sample estimate about population parameter.

Count the field in which important statistical parameters are used.

CONTENTS:

Arithmetic Mean

Median

Quartiles

Topic: Arithmetic Mean

Objectives of the topic: after studying this module you will be able to

1. Define and explain arithmetic mean
2. Calculate arithmetic mean by different methods
3. Know use/applications of arithmean.

CONCEPT AND CONTENT:

Arithmetic mean is a well-known measure of central tendency, which is generally called average.

Arithmetic mean is defined as the sum of values divided the number of values. Arithmetic mean is denoted by $\overline{X}$ and the sum of values by Σx and number of values by n. Thus the formula of arithmetic mean is:

$$\overline{X} = \Sigma x/n$$

For the frequency distribution the formula is:

$$\overline{X} = \Sigma fx/n$$

Where f is the frequency of a value.

When the values are large some special methods are used to reduce the computational labour.

One of these methods is called "short cut method". In short cut method, some middle value or in case of frequency distribution the value with the largest frequency is subtracted from all values. Then the deviations so obtained are averaged. The constant value is again added to the average of deviation to obtain the arithmetic mean of value.

If Dx denote deviation from a constant value say a from value x i.e.

$$Dx = x - a$$

And the sum of deviations is denoted by ΣDx then arithmetic is given by.

$$\overline{X} = \frac{\Sigma Dx}{n} + a$$

For frequency distribution

$$\overline{X} = \frac{\Sigma fDx}{n} + a$$

In case of uniform class interval of a frequency distribution, computation can be further simplified by dividing the deviation in a short cut method by the class interval The arithmetic mean is obtained by multiplying the average of deviations by the class interval and then adding the constant this is called step deviation method. If $D'x = \frac{Dx}{h}$

Where h is common class interval then arithmetic mean i.e.

$$\overline{X} = h \times \frac{\Sigma D'x}{n} + a$$

For frequency distribution

$$\overline{X} = h \times \frac{\Sigma f D'x}{(n)} + a$$

Properties of arithmetic mean:

Following are the properties of arithmetic mean

1) Arithmetic mean of a constant is constant.
   If x=a where a is constant then

   $$\bar{x} = \frac{a + a + \ldots + a}{n} = \frac{na}{n} = a$$

2) If a constant is added to all values then mean will also increase equal to the constant
3) If all values are multiplied by a constant then their mean will increase by constant time.
4) The sum of all deviations of values from the mean is zero.

   $$\Sigma(x - \bar{x}) = 0$$

Uses of arithmetic mean:

Arithmetic mean is used in various fields. some of which are discussed here.

In manufacturing the quality of products are compared with the average life of these products

In agriculture various variety of crops are compared with the average production of these varieties

Effectiveness of various methods is compared with their mean effects. Similarly the performance of student. players and many others are compared with their average score.

In laboratory observations, the exact value is found by taking the arithmetic of observation.

In business arithmetic mean is used for averaging cost. sale. profit. investment etc.

In meteorology arithmetic mean is used for averaging temperature. humidity and atmospheric pressure.

Advantages and disadvantages of arithmetic mean:

Arithmetic mean is easy to calculate and simple to understood. It uses all the observations. Arithmetic mean is mathematical term. The values of arithmetic mean remain relatively stable if calculated from different samples.

Its disadvantages are, firstly, it is affected by extreme value, secondly, it cannot be calculated from "open end" class' data without assuming the ends of classes.

Examples how to calculate Arithmetic mean:

Find the arithmetic mean of marks by (1) simple (2) shortcut method marks s (x): 70. 90. 120. 122. 130. 124.

(1) **By simple method**

$$\bar{X} = \frac{\Sigma x}{n} = \frac{70 + 90 + 120 + 122 + 130 + 140}{6} = \frac{672}{6} = 112 \text{ marks}$$

Marks (x)	Dx = x-a
70	-50
90	-30
120	0
122	2
130	10
140	20
Σ	-48

$$\bar{X} = \frac{\Sigma Dx}{(n)} - a \qquad a = 120$$

n = 6

(2) **Find the arithmetic of data in example 1. By step deviation method: -**

Solution:

Marks (x)	Dx = x – a	$Dx = \frac{Dx}{h}$
70	-50	-25
90	-30	-15
120	0	0
122	2	1
130	10	5
140	20	10
Σ	-	-24

$$\overline{x} = \frac{\sum Dx}{n} \times h + a$$

$\sum Dx = -24$
$h = 2$
$a = 120$

$$\overline{x} = \frac{\sum Dx}{n} \times h + a$$

$$\overline{x} = \frac{-24}{6}(2) + 120$$

$$\overline{x} = \frac{-48}{6} + 120$$

$\overline{x} = -8 + 120 = 112$ marks.

(3). Find the arithmetic mean of ages given below by different methods.

CLASS Age in years	10 - 19	20 - 29	30 - 39	40 – 49	50 – 59
No of persons (f)	10	20	30	10	5

Solution: - By simple methods.

Class	x	F	fx
10 – 19	14.5	10	145.0
20 – 29	25.5	20	490
30 – 39	34.5	30	1035
40 – 49	44.5	10	445.0
50 - 59	54.5	5	272.5
Σ	-	75	2387.5

$$\overline{X} = \frac{\sum fx}{(n)} = \frac{14.5}{75}$$

$$\overline{X} = \frac{2387.5}{75} = 31.83 \text{ years}$$

By short cut method:

Class	x	F	Dx = x – a	FDx
10 – 19	14.5	10	-20	-200
20 – 29	24.5	20	-10	-200
30 – 39	34.5	30	0	0
40 – 49	44.5	10	10	100
50 – 59	54.5	5	20	100
Σ	-	75	-	-200

$$\overline{X} = \frac{\Sigma fDx}{n} + n \qquad a = 34.5$$

$$\overline{x} = \frac{-200}{75} + 34.5\ 31.83 \text{ years}$$

By step deviation method:

Class	x	F	Dx= x-a	Dx = $\frac{Dx}{H}$	fDx
10 – 19	14.5	10	-2	-2	-20
20 – 29	24.5	20	-1	-1	-20
30 – 39	34.5	30	0	0	0
40 – 49	44.5	10	10	1	10
50 – 59	54.5	5	20	2	10
Σ		75			-20

$$\overline{X} = \frac{\Sigma fDx}{n} \times h + a \qquad h= 10 \text{ . } a = 34.5$$

$$\overline{X} = \frac{-20\ (\ 10\ )}{75} + 34.5 = 31.83 \text{ years}$$

METHODOLOGY

Activity No. 1

(Initiative)

1) ask student to tell the prices of their books
2) write the prices on the blackboard
3) ask the student to sum these prices and write the sum on blackboard.
4) Ask what is the number of books, whose prices have been written on the table.
5) Write the number of books on the board.
6) Ask the student to divide the sum of prices on the number of books and write the answer on the board.
7) Ask student what has been calculated? If their answer is "average ". Then tell them that this average in statistic is called "Arithmetic mean".

Now tell simple definition of arithmetic mean as. "the sum of values of divided by the number of values".

Tell the student that arithmetic mean in symbol is written as $\overline{x}$ (read as x bar) and the sum of the values is denoted by Σx (read as summation x) and the number if values by n written on board

Arithmetic mean = $\frac{\text{sum of values}}{\text{No of values}}$ $\overline{x} = \frac{\Sigma x}{n}$

Activity No. 2

Divide the class into suitable groups. As each group to find the arithmetic mean of the marks of group leader in SSC.

Guide and monitor the groups. In the last write on blackboard the arithmetic mean of all groups. Ask student to compare whose arithmetic mean is high.

Activity No 3

Write these on blackboard. Represent prices by x
X(prices in Rs)=20, 30, 35, 50, 60, 80, 90
Ask the students to note these prices on notebook.
Divide the class in suitable groups
Ask them to find arithmetic by

$$\overline{X} = \frac{\Sigma x}{(n)}$$

then ask them to make the following table. Also make the table on black board.

X (1)	Dx=x – a (2)

Write the values in column head x. Select a middle value say 50 and denote it by a. Tell the student to subtract this middle from all the values and write in the column (2) under column head Dx = x- a
Tell them sum column 2 and write this sum in sum box. in the bottom of the column 2.
Tell the student to represent the sum of column (2) by ΣDx
Write the formulae

$$x = \frac{\Sigma Dx}{n} + a$$

on the table ask the student to put the values in the formulae. and calculate the answer. Ask them to compare the answer with that of simple formulae.
Tell them that is shortcut formulae of arithmetic mean.
Tell the student that this method reduces the computation labor.

Activity No. 4

Ask the student to make the following 3-column table. Write the values of activity # in column (1) with column head x

X (1)	Dx = x – a (2)	$Dx = \frac{x-a}{h}$ (3)

Find the values in column (2) as in previous activity. Then divide the values of column of (2) by the highest common divisor (if common divisor is possible other wise do not use this method) and write values in 3rd column under column head $Dx = \frac{x-a}{h}$.

Sum column (3). Write in the sum box write the following formulae on board.

$$\overline{X} = \frac{\Sigma Dx}{N} \times h \quad a$$

Ask the student to represent the sum of column (2) by ΣDx.
Write the formulae.

$$X = \frac{\Sigma Dx}{N} + a$$

On the table.
Ask the student to put the values in the formulae and calculate answer. And compare the answer of simple and shortcut formulae.
Tell them this method is called step deviation method for the calculation of arithmetic mean.

Activity No 5

Write the following data on the blackboard.

Age group class	4 – 9	10 – 15	16 – 21	22 – 27	28 – 33
f persons	10	12	18	13	9

Make the following table:

Class	x	f	fx

Write the group in column (1).
Find the mid point by dividing the sum of lower and upper limit of each class by 2 and write in column 2 under column head x.
Write the frequency of each class in column (3) under f. multiply each value under x by the opposite value under f, and write in column (4) under head fx and
Sum column (3) and denote it by Σfx:
use the formulae

$$x = \frac{\Sigma fx}{n}$$

Write answer on blackboard.

Activity No 6

Writer the data of activity # 4 on blackboard make following table.

Class	x	f	Dx = x – a	$Dx = \frac{Dx}{h}$	fDx

Write classes in column (1) mid point in column (2) take a value that has greatest frequency denote it by a. subtract this value from all values of column (2) and write in column (4).
Divide the values of column (4) by class interval and denote interval by h. write the divided values in column (5).
Multiply column (3) and column (5) and write in column (6).
Add column 3 and column 6
So the sum of column 3 is $\Sigma f = n$
And the sum of column (6) is ΣfDx
Write the following formulae and put the values.

$$\overline{x} = \frac{\Sigma fDx}{n} \times h + a$$

this is step deviation method of arithmetic mean for data with frequency.

SELF ASSESSMENT

Q. No. 1: Give short answer:

1) Define arithmetic mean.
2) What are the properties of arithmetic mean?
3) Count four ideas in which arithmetic mean is used?

Q. No. 2: Mark true or false.

1) Arithmetic mean is simple to understand? (T, F)
2) Arithmetic mean is suitable average for data with extreme values? (T, F)
3) Arithmetic mean is un mathematical? (T, F)
4) Arithmetic mean change if constant is added to value? (T, F)
5) Arithmetic mean of a constant is zero? (T, F)

Would you use arithmetic mean fo

1) x = 2, 4, 5, 7, 50.
   Give reason for your decision?
2) x = 2, 8, 10, 12, and more

Find the arithmetic mean of the following marks:

Marks	10 – 15	16 – 21	22 – 27	28 – 33	34 – 39
F	10	12	13	15	19

Answer of Q. No. 2:

1. T (True)
2. F (False)
3. F (False)
4. T (True)
5. T (True)

MEDIAN

Objectives:

1) define median.
2) calculate median by different methods.
3) use / apply median in practical life.

Contents:

The median is defined as a value. which divides a set of data that has been arranged in the ascending order. in to two equal parts. One part is greater than it and the other part is less than it.

The median of n observations x1. x2. x3…xn when arranged in order of smallest to the largest is the middle value if n is odd. and the average of two middle values of n is even. It means that when n is odd then $\frac{n+1}{2}$th observation is median and when n is even then the median is the average of n/2th and (n/2 + 1)th observation.

In case of discrete frequency distribution. median can be found by the cumulative frequency distribution.

For data grouped onto frequency distribution. the median is the value on the horizontal axis. through which a vertical line divides the histogram of th distribution into two parts of equal area. thus median is that value on the horizontal axis. which correspond to a cumulative frequency n/2.

This value lies in a group called the median group. A single value from the median is calculated by the formula of linear interpolation. which is

Median $= l + h/f\,(n/2 - c)$

Median can also be read from the group of ogive curve.

Median has the following advantages:

It is easily calculated and understood. it is located even in qualitative data, not affected by extreme values and can be calculated in " open end " classes as in income and price data.

In highly skewed distribution median is a suitable average. However there are few disadvantages of median. It is not well defined, it cannot be further statistically treated. As median requires the arranging of values which is tedious and time consuming for a large set of data.

➔ **Ask to calculate the median for the two given series of value.**

➔ **Guide monitor the group work.**

➔ **Ask each group to present their work.**

➔ **Make necessary corrections if needed, encourage the groups.**

<u>Activity No 1:</u>

Median from discrete frequency distribution

- Divide the class into suitable groups
- Write the following marks or frequencies on the blackboard.

B.B = activity

X = 50, 55, 60, 62, 63, 64, 70
f = 4, 3, 5, 2, 7, 3, 4

- Tell them to write marks in column (1) with column head x and the frequencies in column (2) with column head f.
- Tell them to write cumulative frequency in column (3) explain the cumulation of frequency on board.
- Guide / tell that cumulative frequency of value 50 is 4 i.e. its own frequency because no value is less than 50 in the given data frequency of 55 is equal to the sum of previous values and its own i.e. 4 +3 = 7.
- Ask student what will be the cumulative frequency of the other values.
- Check and guide
- Ask from each group to tell the c.f of the last value 70 and ask them whether it is equal to total frequency or not.

Activity No 2:

- Divide the class into suitable groups.
- Provide each group with odd number of sticks of different sizes.
- Tell them to arrange these sticks according to their sizes in ascending or descending order.
- Ask to select the stick, which occur in the middle according to the length.
- Tell them that the height of this stick represents the median height of the sticks.
- Ask the students that what will be the median if one stick is increased to make the number of sticks even.
- Tell them that in even case the two middle sticks, average height will be the median of these sticks.
- Tell them now that median is defined as the middle value of the data arranged in ascending or descending order.
- Tell them that mathematically median for n values, when n is odd is given by

$$\text{Median} = \frac{n+1}{2}\text{th value}$$

In case when n is even

$$\text{Median} = \frac{1}{2}\left[\frac{n}{2}\text{th} + \left(\frac{n}{2}+1\right)\text{th}\right]$$

Activity No 3:

- Divide the class into suitable groups.
- Provide each group with (1) odd number of values (2) even number of values.
- Tell the students that cumulative frequency of a value shows the positional ranks on which that value falls.

Write b.b that of 50 is 4, which means at 1st, 2nd, 3rd, and 4th position is value 50 then c.f of 55 is 7

Activity No 4 from book: calculation of median for grouped data.

- ✓ Divide the class into suitable groups.
- ✓ Ask to open the books and write the question on page no. ____
- ✓ Ask to write the classes in column 1.
- ✓ Tell the students that write the class boundaries in column 2

Using the following steps:

a. Tell students how to find class boundaries.
b. Write on the board for demonstration.
c. Ask students to find the successive class boundaries. and write on blackboard.
d. Ask to write the frequencies in column (3)
e. Ask to find the cumulative frequency and write in column (4)
f. Write the formulae for finding median in grouped data on board. explain each term of it. i.e.

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c\right)$$

Where l = lower class boundary of median class
h = class-interval of the median class
f = frequency of the median class.
c =cumulative frequency of the class preceding the median class.

SUMMARY:

Median is the middle value when the data is arranged in ascending or descending order
Median is simple to calculate and easy to understand
Median is not affected by extreme value
Median can be calculated for qualitative data
Median is not mathematical in nature

Self-Assessment:

Q.1) Define median how it is calculated from odd and even values?
Q.2) how the median is calculated for grouped data?
Q.3) what are the situations in which median is used?
Q.4) mark the true and false.

i. Median is affected by the extreme values in the data
ii. Median is used to describe the qualitative data
iii. Median cannot be calculated for open-end classes.
iv. Median is used when there is great variation on the data.

Ans Q.4) (i) F. (ii) T. (iii) F. (iv) T.

Q.5) find the median from the given data:

Class	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64	65 – 69
F	10	16	21	15	11	8

TOPIC QUANTILES

Objectives:

After studying this unit students will be able to

1. define quintiles and knows its various kinds.
2. calculate quintiles.
3. apply quintiles in daily life.

Introduction:

This unit include

1. definition of quintiles
2. various kinds .
3. relationship between quintiles.

Content:

When the number of observations is very large, the principle according to which a distribution or an ordered data divided in to two equal parts may be extended to any number of divisions.
The three values, which divide the distribution into four equal parts. are called quartiles. The values are denoted by Q1, Q2. and Q3 respectively. Q1 is called first or lower quartile and Q3 is known as the third or upper quartile.25 % values are less than Q1. 50 % less than Q2 and 75 % less than Q3 in a data set. Similarly the nine values. which divide the distribution into ten equal parts. are called deciles and are denoted by D1. D2. and D3...D9. While the ninety-nine values dividing the data into one hundred equal parts a called percentiles and are denoted by P1, P2...P99.
The second quartile or the fifth decile or the fiftieth percentiles are identical with median.
Quartiles, deciles and percentiles are collectively called quartiles.
All the quartiles are percentiles. e.g. the third quartile is the seventy fifth percentile and the sixth decile corresponds to the sixtieth percentile therefore all the quartiles are calculated on the pattern of percentile. Thus any jth percentile denoted by Pj for an ordered set of n values is

1. $Pj = \left[\frac{jn}{100}\right] + 1$ when $\frac{jn}{100}$ is not an integer.

2. $P_j = \frac{1}{2}\left\{\left(\frac{jn}{100}\right) + \left(\frac{jn}{100} + 1\right)\right\}$ when $\frac{jn}{100}$ is integer.

thus in the same manner

$$Q1 = \left(\left[\frac{n}{4}\right] + 1\right)\text{th value for } \frac{n}{4} \text{ is not an integer}$$

$$Q1 = \frac{1}{2}\left\{\frac{n}{4}\text{th} + \left(\frac{n}{4} + 1\right)\text{th}\right\} \text{ for } \frac{n}{4} \text{ is integer}$$

In grouped data to find a particular quartile first the group containing that quartile is found and then a single value is calculated by using the formulae of linear interpolation for example Q3 is found by first finding the class corresponding to cumulative frequency 3n/4and then single value of Q3 is found as

$$Q3 = l + \frac{h}{f}\left(\frac{3n}{4} - c\right)$$

Activity No.1

X=13.12.21.14.9.8.25.3 find Q1

Arrange the values in ascending order and find ranks

Ranks	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
X	3	8	9	12	13	14	17	21	25

As n = 9

Note the value. which has 1/4th less than it. and 3/4th greater than it is 3rd value. which is 9.

So Q1 = 9.

In fact $\left(\left[\frac{9}{4}\right] + 1\right)$th = (2+ 1)th =3rd

where [9/4] means retain integral part i.e. 2 and omit fractional part i.e. 0.25

ıus when n/4 is not integer then

Q1 = ([n/4] + 1)th

HIRD QUARTILE Q3 :

A value which has 3/4th of values less than and i/4th of values greater than it is called third quartile, denoted Q3

In above case 7th value i.e. 17 is Q3. This 7th is in fact obtained from n. using following relation:

$([\frac{3(9)}{4}] + 1)$th = (6 + 1)th = 7th

thus when 3n/4 is non-integer then

$$Q3 = ([\frac{3n}{4}] + 1)\text{ Th}$$

DECILE:

Definition:

These are values. which divide the data into ten equal parts. There are nine deciles i.e. 1st, 2nd, 3rd…9th denoted by D1. D2. and D3…D9 respectively.

D3 is defined as 3n/10th value after the data is arranged in ascending order.
Find Q1. Q3.

From the following data

x = 12. 8. 7. 11. 16. 21. 27. 18. 2. 4. 3. 20

arrange the above values in ascending order:

	1st.	2nd.	3rd.	4th.	5th.	6th.	7th.	8th.	9th.	10th.	11th.	12th
x =	2	3	4	7	8	11	12	16	18	20	21	27

Thus Q1 is the average of third and fourth value i.e.

$$Q1 = 1/2\,(3^{rd} + 4^{th})$$

$$Q1 = \frac{4+7}{2} = \frac{11}{2} = 5.5$$

In fact 3rd value = 12/4th = n/4th and

$$4^{th}\text{ value} = (12/4 + 1)\text{th} = (n/4 + 1)\text{th}$$

thus when n/4 is integer then

$$Q1 = \frac{1}{2}[\frac{n}{4}\text{th} + (\frac{n}{4} + 1)\text{th}]$$

similarly when 3n/4 is integer then

$$Q3 = \frac{1}{2}[\frac{3n}{4}\text{th} + (\frac{3n}{4} + 1)\text{th}]$$

Example of Decile:

A value that divides the data into ten equal parts i.e. there are nine deciles D1, D2, D3…D9, where

D1 = n/10th

D2 = 2n/10th

D5 = 5n/10th = n/2 Th i.e. D5 is equal to median.

Find D4 for above data. In analogous to quartile

Check 4n/10 is integer or non-integer.

$\frac{4(9)}{10} = 3.6$ non integer

thus D4 = ([$\frac{4n}{10}$] + 1)th = (3+1)th = 4th value

which is 7 D4 = 7

similarly analogy of quartile is used for integral value of any decile Jn/10th =integer

then D4 = $\frac{1}{2}$ [$\frac{Jnth}{10}$ + ($\frac{Jn}{10}$ + 1)th]

Percentile:

A value that divides the data into 100 equal parts. There are 99 percentiles. Denoted by Pj.

i.e. P1, P2, P3…P99.

Using the analogy of quartile for any Pj if Jn/100 is non-integer.

Then

Pj = ([$\frac{Jn}{100}$] + 1)th value.

if Jn/100 = integer

then Pj = $\frac{1}{2}$ [$\frac{jnth}{100}$ + ($\frac{Jn}{100}$ + 1)th]

find P6 for the above data.

Quartiles for discrete frequency distribution:

Find Q3, D7, and P60 for the data given below:

X Marks	10	15	27	54	61	72	87	95
F	16	27	24	50	18	15	10	3

Solution:

(1) x	(2) f	(3) cf
10	16	16
15	27	43
27	24	67
54	50	117
61	18	135
72	15	150
87	10	160
95	3	163

For Q3

$\frac{3n}{4} = 3(\frac{163}{4}) = 122.25$ which is non-integer.

Thus $Q3 = ([\frac{3n}{4}] +)th = (122 + 1)th = 123^{rd}$ which is 61

For D7 =

$\frac{7n}{10} = 7(\frac{163}{10}) = 114.1$ Non-integer.

thus

$D7 = ([\frac{7n}{10}] + 1)th = 115^{th}$ which is 54

for P60

$(\frac{60n}{100}) = [\frac{60(163)}{100}] = (97.8)$ non-integer.

Quantile for the continuous data:

Find Q1. D8. and P40 for the following marks.

Class Marks	30 – 39	40 – 49	50 – 59	60 – 69	70 – 79	80 – 89
f Students	10	20	22	15	12	8

Solution:

Class (1)	c.b (2)	f (3)	cf (4)
30 – 39	29.5 – 39.5	10	10
40 – 49	39.5 – 49.5	20	30
50 – 59	49.5 – 59.5	22	52
60 – 69	59.5 – 69.5	15	67
70 – 79	69.5 – 79.5	12	79
80 – 89	79.5 – 89.5	8	89

Find the class of Q1: -

Which is $n/4^{th} = \frac{87}{4} = 21.75^{th}$ which corresponds to the class 39.5 – 49.5. for single value of Q1

use linear Interpolation formulae on the same pattern as for median. Similarly find D8 and P40 quantiles are used as reference value for the values in the data for example the marks of the student can be viewed with reference to the Q1, median and Q3. Those below Q1 may be grouped as the weakest and those between Q1 and median are grouped as weak and between median and Q3 as good and above Q3 as excellent. Similarly percentiles are used for expressing the marks in percentile ranks.

Self-Assessment:

Q.1) Give short answers:

a) What are quantiles?
b) Define quartile decile and percentile

Q.2) Find the percentile rank of marks 60, 84, from the data on marks of 120 students.

Class (marks)	10 – 19,	20 – 29,	30 – 39,	40 – 49,	50 – 59,	60 – 69,	70 – 79,	80 – 89
f	20	18	21	15	12	8	10	9

Q.3) find Q3. D8 from the data in (Q.2)

Q.4) fill in the blanks from the choices given for each blank.

i. P50 is equal to ________ (Q1, median, Q3, D7)
ii. A decile divides the data into ___ _____ (4, 10, 20,100 parts)
iii. For Jn/10 is integer then Dj = _____(jn/10 th, (jn/10 + 1) th, 1/2[(jn/10)th + (jn/10 + 1)th]
iv. P40 is less than _______(Q1, D4, D7)

Answers of Q.4) (i) median (ii) 10 (iii) ½[(jn/10) th + (jn/10 + 1)th] (iv) D7