Cihan University/Sulaimanyia

Department of - General Education

University of Cihan- Sulaimaniya

Subject: Educational Statistics

Course Book – Year -2023-2024

Stage-Two

Lecturer's name - Prof.Dr. Obaid Mahmmood Mohsin

Academic Year: 2023/2024

Department of - General Education

University of Cihan- Sulaimaniya

Subject: Educational Statistics

Course Book – Year -2023-2024

Stage-Two

Lecturer's name - Prof.Dr. Obaid Mahmmood Mohsin

Academic Year: 2023/2024

Course Book

1. Course name	Educational Statistics
2. Lecturer in charge	Prof.Dr.Obaid Mahmmood Mohsin
3. Department/ College	General Education
4. Time (in hours) per week	3
5. Office hours	4 for morning
6. Course code	CUES 84027
7. Teacher's academic profile	*obaidmohsin808@gmail.com*
8. Keywords	

9- **Course overview :** At the end of this course the student will:-

1-Understand what is the meaning of Statistics, Population, Sample, Elements, Variables, and other definitions.
2- Learn how to collect, describe, summarize, classified and interpret data.
3 - Understand and apply the Measures of Central tendency, like the Mean , The Median, and the Mode.
4- Understand and apply the Measures of Dispersion, like Range (R),Variance (S^2), Standard deviation(S), and the coefficient of variation (CV).

10. Course objective:

This course is designed to help the students enrolled in the first semester how to use the Statistical Techniques. The course combines the best of the traditional methods with more recent modern approaches. The main objective is to help students how to use statistical techniques or methods to collect data and information from different sources and summarized in different ways, classified data in to different classes and groups , analyzed data by different statistical methods and interpret the results to take decisions.

11. Student's obligation

Note book
#Solve exercises -&Four Sets of problems &Its solving
#Midterm Examination
#Final examination

12. Forms of teaching

Give the title of lecture
Summary about it
Explain it
Solve examples
homework

13. Assessment scheme

ECTS 4.00

Units 3

Type S

14. Student learning outcome:

At the end of semester the students understand : what is Statistics ,why study Statistics ,Some definitions (population, element ,observation ,data set ,variable , sample , sampling) , types of : statistics, variables, data , samples ,Statistical method and its steps ,How tabulate and represent Qualitative and Quantitative data-frequency distribution ,relative frequency distribution, percentages distribution, bar graph, pie chart, histogram, polygon , cumulative distribution ,ogive .

Mean , median , mode, how calculate the value of every one of them for the grouped and ungrouped data and how explain them .

Variance , standard deviation, range ,how calculate the value of every one for grouped and ungrouped data and how to explain them.

Also solve more than **60** exercises

15. Course Reading List and References:

Key references: Statistics for Business and Economics- 2000/Anderson , Sweeny

- Useful references: Statistics for Management and Economics 2003 –Gerald Keller & Brian Warrak.
  -Statistics for Economics, Accounting and Business Studies /Michael Barrow.
  -Comprehensive Statistical Methods-2007/P.N.ARORA

Introductory Statistics
PREM S.MANN Sixth Edition---john wily &sons ,inc 2006

16. The Topics:

17. Peer review

Topic one : Introduction and definitions. & Exercises (Set of 14 questions and solving)

Topic two: Tabulating and Graphing Qualitative and Quantitative Data.& Exercises (Set of 15 questions and solving)

Topic Three: Measures of Central Tendency.& Exercises (Set of 15 questions and solving)

Topic four : Measures of Dispersion & Exercises (Set of 10 questions and solving)

Main Lecturer incharged **Head of The Department**

Cihan University-Sulaimanya Camp

Principles of statistics

Topic –ONE-

Introduction and definition

2023-2024

Prof.Dr. Obaid Mahmmood Mohsin

1. What is statistics?
2. Why study statistics?
3. Definition:

   Population, observation, element, variable, sample, sampling
4. types of statistics
5. types of variables
6. statistical method
7. types of samples
8. Exercises.

Topic 1

Introduction and definition

1. What is Statistics?

The word statistics has two meaning.

in the more common usage statistics refers to

Statistics as a Numerical facts:

- The numbers that represent the income of family

,- the age of student,

- the percentage of passes completed by the quarter of a football team,

- the starting salary of a typical college graduate,

are examples of statistics in this sense of the word.

Statistics:

Is a group of methods used to

-collect

- present

- analysis

- and **interpret** data

- and **make decisions**

2. Why study statistics?

There are at least three reasons for studying statistics.

a) Data are every where.

b) Statistical techniques are used to make many decisions that effect our live.

c) No matter what your career, you will make professional decisions that involve data.

An understanding of statistical methods will help you make these decisions more effectively.

3. Definitions:

Population:

Population: the collection of all elements of interested.

Or

Consists of all element- individuals, items, or objects or (statistical unites) whose characteristic are being studied. The population that is being studied is also called the target population.

Observation or (measurement):

the value of a variable for an element.

is called an observation or measurement.

Data set: Is a collection of all observations on one or more variable.

Element or member: An element or member of sample or population is a specific subject or object (for example) a person, firm, item, state, or country, about which the data is collected.

Variable: Is a characteristic under study that assumes different values for different elements. In contrast to a variable the value of a constant is fixed.

Example:

- The number of houses built in a city per month during the past year.
- The gross profits of companies.

- The number of insurance policies sold by a sales person per day during the past month.

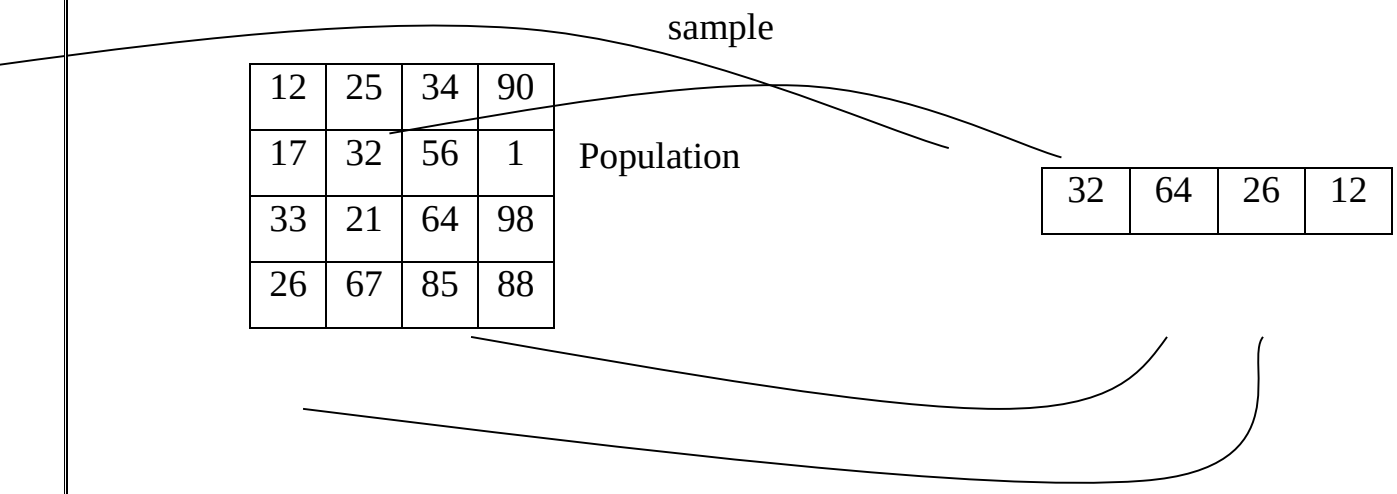
In general, variable assumes different values for different elements, and denoted by X, Y, Z, ….

Sample:

Apportion of the population selected for study or / the selection of a few elements for this population or / Is part of the population which was chosen to represent the population a real representation.

Sampling:

The operation of choosing a part from the units of the population in a way implied is representation of the whole population .Or. The operation of choosing a sample .



Example:

The following table lists the departments in Cihan University-Sulaimanya Camp and the number of students in 2016-2017

Department	Number of students
Law	250
Computer Science	75
Architecture	45
English	110
Banking	150
Accounting	350
Business	320
Accounting by IT	280

Computer Science is an element or member An observation or measurement

We have 8 observations

Data Set

4***. Types of statistics:***

1- Descriptive statistics:

Consists of methods for organizing displaying and describing data by using table, graphs and summary measurement

2- Statistical inference (inferential statistics):

Consists of methods that use samples results to help make decisions or predictions about population.

Or

It includes the statistical methods which aim to draw conclusions and evidence on the source from which the data were collected.

This division has two main subdivisions:

1- Estimation

2- test of hypotheses

5. Types of variables:

1- Quantitative variables:

A variable that can be measured numerically is called quantitative variables.

The data collected on quantitative variables are called *quantitative data*.

Example:

Price of homes,

number of accidents…

are examples of *quantitative variables* because each of them can be expressed numerically

.

Such quantitative variables may be classified either *discrete variable* or *continuous variable* .

Discrete variable:

A variable whose values are a countable. In other words, discrete variable can assume only certain value with no intermediate value.

Example:

- The number of cars sold on any day
- The number of people visiting a bank on any day.
- The number of student in a class.

Continuous variable:

A variable that can assume any numerical value over a certain interval numerical value over a certain interval or intervals.

Example:

- The time taken to complete an examination.
- The height of a person
- The weight of people.

Any variable that involving money is considered a continuous variable.

2- Qualitative or categorical variables:

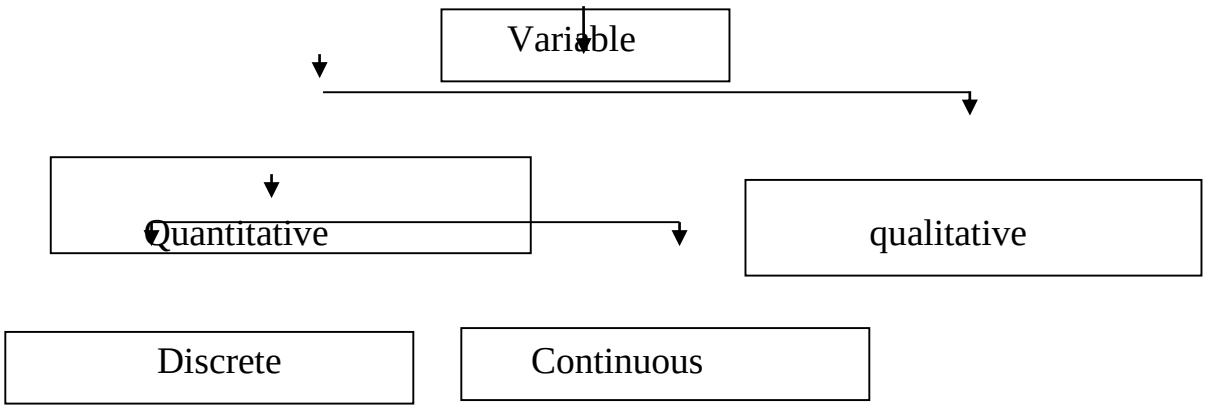
A variable that cannot assume a numerical value but can be classified into two or more non numerical categories is called qualitative or categorical variable .

The data collected on such a variable are called *qualitative data.*

Example:

The status of an undergraduate college student. The gender , hair color

The following figure illustrates the type of variable



6. ***Statistical method:***

Is the scientific method treats with the phenomena subject to quantitative analysis (number).

The steps of the statistical method according to its logical sequence are.

1- Data collecting
2- Data classification and tabulation
3- Data presentation
4- Analysis step
5- Explanation and forecasting step or decision making.

7. Types of samples:

Samples are dividing by size into:

- Large sample size
- Small size samples

Samples are divided according to technique into:

- probability samples
- non-probability samples

Probability samples:

- simple random sample
- stratified random sample
- systematic sample
- multistage sample

Non-probability samples:

- intentional sample

Simple random sample:

It is the sample taken form a population in which all its unites have been given the same chance of occurrence.

The most important condition is that the *population must be similar*

i.e : that similar units.

Example:

- A factory produces one type of electric light bulbs. The drawn sample is a simple random sample because the product is similar.
- A store contains one type of stored which received at the same time, hence the drawn sample is a simple random sample because the unites are similar.
- The study of lengths and weight of students in one of primary school. The drawn sample is a simple random sample because the population of this phenomenon (length or weight) is similar.

Stratified random sample:

When the statistical population under searching consisting of strata because of some considerations and criteria. Each strata has its own qualities that make it different from the others.

Hence the use of the simple random sample is impossible because of the unlike units.

The operation of choosing the stratified random sample is done by choosing simple random sample from each stratum. The total number of these samples represent the stratified random sample. Here we have to take into consideration the proportional structure of population of each stratum of the population strata in relation to the whole units of population .If we symbolize (n_h) to the size of the sample to each stratum (h),

N the size of population

n the size of the sample

N_h the size of the stratum from the population

Then

$$n_h = n * N_h / N$$

Example:

1- A factory produces two type of production A , B the production of A is two time of B the random sample is stratified random sample in condition that the number of unites drawn from A stratum is also two times of those drawn from B (the proportional representations) and the total number of the drawn samples from A and B is the stratified random sample.

2- A store contains three types of foods. The drawn sample is a stratified random sample. We take a simple random sample from each of the three types of foods, the total number represent the stratified random sample.

3- If you want to pull a sample with size (10) from population consists of:

20 doctors

30 engineers

50 students

1- What type of sample and why?

2- How are selected sample drawn selections?

Cihan University –Sulaimanya Camp

Principles of Statistics
Topic –ONE-

Introduction &Definitions

Exercises & solving

2023-2024

Prof.Dr. Obaid Mahmmood Mohsin

Principles of statistics

Topic 1

Exercises

1-What is statistics ?

2-Why study statistics ?

3- Define : population ,observation ,data set ,element ,variable ,sample , sampling ,statistics .

4-Explain the types of :

(i) Statistics
(ii) Variables and draw the flowchart

5-Answer with true or false .Correct the false statements .

(i) In the more common ,usage statistics refers to numerical facts .
(ii) Statistics is a group of methods used to collect and analyze data only .
(iii) One of the reasons for studying statistics : ((Statistics techniques are used to make many decisions that effect our life))
(iv) Population means the collection of all elements of interest .
(v) Data set is collection of observations on one variable only.
(vi) Sample means the operation of choosing a part from the units of the population in away implied areal representation of the whole population .
(vii) They are two types of statistics : descriptive statistics and statistical inference
(viii) Sampling is a part of the population which was chosen to represent the population real representation .
(ix) They are two types of variables quantitative and qualitative variables ,and quantitative variables also two types discrete and continuous variables .

6- Explain the statistical method and write its steps .

7- The following table lists the number of students in departments in Cihan university /Sulaimanya at 2012-2013:

Department	Number of students
Accounting by IT	400
Accounting	60
Law	300
Business Administration	65
English	80
Computer Science	75

(i)Show: variable ,element or member ,observation or measurement ,data set ,and population

(ii)How many :variables ,observations ,elements ,data sets ,population ?

8-If you want to pull a sample with size (30) from a population consists of :

60 students of law department

150 students of accounting by IT

90 students of business administration

(i) What type of sample and why ?
(ii) How is selected sample drawn selection ?
(iii) If you pull a sample from the students of accounting by IT only ,what type of sample and why ?

9-Explain the :

(i) Sample random sample .
(ii) Stratified random sample

And give two examples for every one

10-The following table lists the number of persons visit a bank through three weeks in the last month from four cities A ,B ,C and ,D .

	Number of persons visit a bank		
City	First week	Second week	Third week
A	250	200	150
B	200	80	30
C	300	200	100
D	90	50	40

(i) Define : observation , variable , population ,data set .
(ii) Show : element ,observation ,data set ,population ,variable ,type of variable .
(iii) How many : variables ,observations ,elements , data sets ?

11-What are the types of :

(i) Variables ?
(ii) Data ?
(iii) Statistics?
(iv) Samples ?

12-Give an example for every one of the following :

(i) Simple random sample .
(ii) Stratified random sample .
(iii) Continuous variables .
(iv) Discrete variable.

13-The following table lists the number of students in departments in Cihan university /Sulaimaniya at 2012-2013.

Department	Number of students in	
	Morning	Evening
Accounting by IT	250	200
Accounting	40	30
Law	100	200
Business Administration	50	40
English	30	20
Computer Science	60	0

(i) Define : the variable and show the variable .
(ii) Show: element , observation ,population ,data set .
(iii) How many :variables ,observations ,elements, data sets . populations ?
(iv) What are the type of the variables ?

14-If you want to pull a sample with size (100) from a population consists of :

1000 students in first stage

600 students ib second stage

400 students in the third stage

(i) What is the type of the sample and why?
(ii) How are selected sample drawn selection ?
(iii) If you choose a sample from the first stage only ,what is the type of sample and why ?

15-What is the type of the sample and why for every one of the following :

(i)Choose a committee of (8) students from a class contains (10) girls and (30) boys ?

(ii) A factory produces one type of electric light bulbs ?

(iii)A factory produces three types of production A , B ,and ,C , and the production of B , is two times of A , and the production of C , is the same production of B .

Principles of statistics

Topic 1

Solving the Exercises

1- The word statistics has two meaning :

(i) Statistics: Is a group of methods used to collect ,present ,analyze and interpret data ,and make decisions.

(ii) In the more usage statistics refers to numerical facts .

For example :the numbers that represent the income of family . Is example of statistics in this sense of the word .

2-There are at least three reasons for studying statistics :

(i)Data are every where .

(ii) Statistical techniques are used to make many decisions that effect our live.

(iii) No matter what your career ,you will make professional decisions that involve data .

Or:

an understanding of statistical methods will help you make these decisions more effectively .

3-population:the collection of all elements of interest.

Observation : the value of a variable for an element .

Data set : is a collection of observations on one or more variables .

Element : an element or member of a sample or population is a specific subject or object about which the data is collected .

Variable : Is a characteristic under study that assumes different values for different elements . In contrast to a variable that value of a constant is fixed . In general , variable assumes different values for different elements and denoted by X ,Y ,Z .

Sample : Is a part of the population which was chosen to represent the population a real representation .

Sampling :The operation of choosing a part from the units of the population in a way implied a real representation of the whole population .

Or : the operation of choosing the sample .

Statistics :Is a group of methods used to collect , present ,analyze ,interpret data and make decisions .

4- (i) Types of statistics :

#Descriptive statistics : consists of methods for organizing displaying and describing by using tables ,graphs and summary measures .

#Statistical Inference :or :inferential statistics :consists of methods that use sample results to help make decisions or predictions about population .

Or:

It includes the statistical methods which aim to draw conclusions and evidence on the source from which the data were collected . This division has two main subdivisions :

-Estimations -Test of hypotheses.

(ii)The types of variables :

#Quantitative variables :a variable that can be measured numerically .The data collected on quantitative variable are called quantitative data . Such quantitative variables may be classified either discrete variables or continuous variables .

Discrete variable : a variable whose values are a countable.

Ex :The number of cars sold on any day , .The number of people vesting a bank on any day .

Continuous variable : a variable that can assume any numerical value over a certain interval.

Ex: The time taken to complete an examination ., The weight of people .

Qualitative variables :a variable that cannot assume a numerical value but can be classified into two or more no numerical categories .

Ex : The gender of a person , hair color .

5-

(i) True
(ii) False –Statistics is a group of methods used to collect ,represent ,analyze and interpret data and make decisions .
(iii) True
(iv) True
(v) False ---data set is a collection of observations on one or more variables .
(vi) False ---Sampling .
(vii) True.
(viii) False ---Sample .
(ix) True

6- Statistical method : Is the scientific method treats with the phenomena subject to quantitative analysis (numerical). Its steps:

(i) Data collection .
(ii) Data classification and tabulation .
(iii) Data presentation .
(iv) Analysis step
(v) Explanation and forecasting step or decision making .

7-

(i) Variable =number of students -- ,Element =business administration ---, observation =400 ,---Data set ={ 400 ,60 ,300 ,65 ,80 ,75 }.
Population ={ACC byIT ,ACC ,Law Business administration ,English ,Computer Science }

(ii) We have one variable only .
We have (6) observations .
We have (6) elements
We have one data set only
We have one population only .

8-

(i) The type of sample is stratified random sample because the population consists of three strata .

(ii) N1 =60 ,N2=150 ,N3=90
N= 60 +150+90 =300
n=30
nh=n * (Nh/N)
n1=30*(60/3000) =6
n2=30 * (150/300) =15
n3 = 30 * (90/300)= 9
n =6 +15 +9 =30

9- Simple random sample

It is the sample taken from a population in which all its units have been given the same chance of occurrence.

The most important condition is that the population must be similar i.e :those similar units.

For example:

1-A factory produces one type of electric light bulbs .The drawn sample is a simple random sample because the product is similar .

2-A store contains one type of stored which received at the same time , hence the drawn sample is simple random sample because the units are similar .

Stratified random sample :

When the statistical population under search consists of strata because of some considerations and criteria .Each strata has its own qualities that make it different from the others. Hence the use of the simple random sample is impossible because of the unlike units . The operation of choosing

the stratified random sample is done by choosing simple random sample from each stratum , the total number of these samples represents the stratified random sample . Here we have to take into consideration the proportional structure of population of each stratum of the population. If we symbolize (n_h)to the size of the sample to each stratum (h)

N the size of population

n the size of sample

N_h the size of stratum from the population

Then:

$n_h = n * (N_h/N)$

and $n = \Sigma n_h = n_1+n_2+\ldots\ldots$

for example:

1-A factory produces two types of production A ,B the production of A is two times of B .

The drawn sample is stratified random sample in condition that the number of units drawn from a stratum A is also two times of those drawn from B (the proportional representations) and the total number of the drawn samples from A and B is the stratified random sample.

2-A store contains three types of food .The drawn sample is stratified random sample.

We take a simple random sample from each of the three types of food , the total number represent the stratified random sample.

10-

(i) observation :The value of variable for an element.

Variable :Is a characteristic under study that assumes different value for different elements. In contrast to a variable that value of a constant is fixed . In general variable assumes different values for different elements , and denoted by X ,Y ,Z ,….

Population : The collection of all elements of interest .

Data set : Is a collection of observations on one or more variables .

(iii) Element =A , Observation =80 , Data set = {200 ,80 ,200 ,50 ,150. 30 ,100 ,40 ,250 ,200 ,300 ,90 } , population= {A ,B ,C ,D } , Variable =number of persons visit a bank first week. , The type of the variable is quantitative discrete variable .

(iv) There are (3)variables . There are (12) observations . There are (4) elements . ,There are (one) data set .

11-

(i)There are two types of variables : 1- Quantitative variables which are classified into:--Discrete variables ---- continuous variables . –2-Qualitative variables.

(ii)There are two types of data : 1- Quantitative data . 2-Qualitative data .

(iii) There are two types of statistics : 1- Descriptive statistics. 2- Statistical Inference or Inferential statistics .

(iv) Samples are divided by size into : -Large sample size , Small –sized samples ,
Samples are divided according to technique into: Probability samples , Non-probability samples .

12-

(i) A factory produces one type of electric light bulbs. The drawn sample is a simple random sample .

(ii) A store contains three types of food .The drawn sample is a stratified random sample.

(iii) The time taken to complete an examination .

(iv) The number of people visiting a bank in any day .

13-

(i) Variable is a characteristic under study that assumes different values for different elements. The variable =number of students in morning .

(ii) Element =Accounting by IT , Observation=100 , population ={Accounting by IT ,Accounting ,Law , Business administration , English , Computer Science } ,
Data set = {250 ,40 ,100 ,50 ,30 ,60 ,200 ,30 ,200 ,40 ,20 , 0 } .

(iii) There are two variables . There are (6) elements . There are (12) observations . There is **one** data set only . There is **one** population only .

(iv) The types of variables are :-Quantitative discrete variables .

14-

(i)The type of sample is stratified random sample because the population consists of three strata .

(ii)Let -----, N1=students in first stage =1000. ---- ,N2=students in second stage =600

N3 =students in third stage =400 ,---- N=the size of population =1000+600+400 =2000

n= the size of stratified random sample =100

nh= n * (Nh /N)

n1= 100 *(1000/2000)= 50

n2= 100* (600 /2000)= 30

n3 = 100 * (400 /2000)= 20

The stratified random sample be : 50 students of the first stage.

30 students of the second stage.

20 students of the third stage.

(iv) The type of sample is simple random sample because the population is similar .

15-

(i) The type of sample is stratified random sample because the population consists of two strata.

(ii) The type of sample is simple random sample because the population is similar.

(iii) The type of sample is stratified random sample because the population consists of three strata.

Cihan University-Sulaimanya Camp

Principles of Statistics

Topic –TWO-

Tabular and Graphical Methods:

Prof.Dr. Obaid Mahmmood Mohsin

2023-2024

Frequency Distribution.

Organizing and Graphing Qualitative Data.

Organizing and Graphing Quantitative Data.

Topic 2

Tabular and Graphical Methods:

Frequency Distribution:

Is a Tabular summary of data showing the number (frequency) of limits in each several no overlapping classes.

Qualitative Data:

Organizing and Graphing Qualitative Data:

Frequency Distribution for qualitative data:

A frequency distribution for qualitative data lists all categories and the number of elements that belong to each of the categories.

Example:1

A sample of (100) students enrolled at a university were asked what they intended to do after graduation.

(44) Said they wanted to work for private companies/businesses.

(16) Said they wanted to work for the federal government.

(23) Wanted to work for state or local government.

(17) Intended to start their own businesses.

Find a frequency distribution table .

Type of employment students intended to engaging.

Type of employment	Number of students
private companies	44
federal government	16
state or local government	23
own business	17
Sum	100

The table lists the types of employment and the number of students who intend to engage in each type of employment.

In this table:

The variable is the type of employment, which is a qualitative variable.

The categories (representing the type of employment) listed in the first column are mutually exclusive.

(Each of (100) students belongs to one and only one of this categories.)

The number of students who belong to a certain category is called the frequency of that category.

Frequency Distribution exhibits how the frequencies are distributed over various categories.

The table is called a Frequency Distribution table or simply Frequency table.

Example:2

A sample of (30) employees from companies was selected and the employees were asked how stressful their jobs were. The responses of these employees are recorded below where very represents.

Very mean Very stressful

Some what means some what stressful

None means for not stressful at all.

Some what	None	Some what	very	very
Very	Some what	Some what	Very	Some what
Very	Some what	None	Very	None
Some what	Very	Some what	Some what	Very
Some what	Very	Very	Some what	None
None	Some what	Some what	None	Some what

Construct a Frequency Distribution table for these data.

Note that the variable in this example is how stressful is an employee's job. This variable is classified into three categories.

Very stressful

Some what stressful

Not stressful

Solution:

Frequency Distribution of stressful on job

stressful on job	tally	Frequency (f)
Very	//// ////	10
Some what	//// //// ////	14
none	//// /	6
Sum		30

Relative frequency and percentage Distributions:

The Relative frequency of a category is obtained by dividing the frequency of that category by the sum of all frequencies. Thus, the relative frequency shows what fractional part or proportion of the total frequency belongs to the corresponding category. A relative frequency distribution lists the relative frequencies for all categories.

Calculating Relative frequency of a category:

$$Relative\ frequency\ of\ a\ category = \frac{frequency\ of\ that\ category}{sum\ of\ frequencies}$$

$$R_i = f_i / \sum f_i$$

The percentage for a category is obtained by multiplied the relative frequency of that category by (100).

A percentage distribution lists the percentages for all categories.

Calculating percentage:

Percentage = (Relative frequency) * 100

Percentage (Pi) = (Ri)* 100

Example:3

Determine the relative frequency and percentage distribution for the data in the last example.

Solution:

stressful on job	Relative frequency	Percentage
Very Some what none	10/30 = 0.333 14/30=0.467 6/30= 0.200	0.333*100=33.3 0.467*100=46.7 0.200*100=20
Sum	1	100

We can state that 0.333 or 33.3% of the employees said that their job are very stressful .By adding the percentages for the first two categories, we can state that 80 % of the employees said their jobs are very or some what stressful. The other number in table can be interpreted the same way.

Notice ***that the sum of the relative frequency is always 1 or ~1.***

The sum of the percentages is always 100 or ~ 100*.*

Graphical presentation of qualitative Data:

The ***bar graph*** and the ***pie chart*** are two types of graphs used to display qualitative data.

Bar graph:

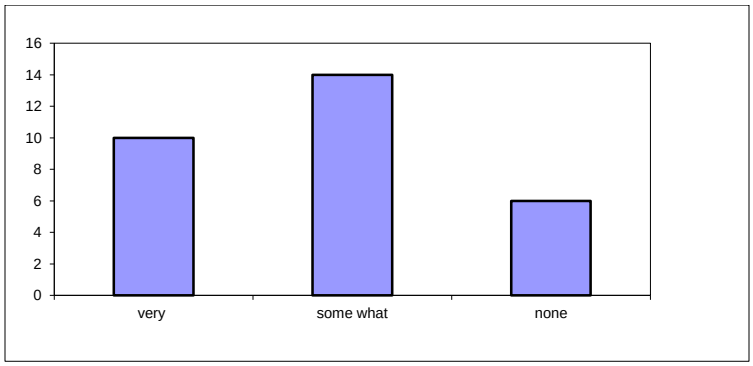
Def:

A graph made of bars whose heights represent the frequencies of respective categories is called a bar graph.

To construct a bar graph (also called bar chart), we make the various categories on the horizontal axis (as in the figure).Note that all categories are represented by intervals of the same width. We mark the frequencies on the vertical axis. Then we draw one bar for each category such that the height of the bar represents the frequency of the corresponding category. We leave a small gap between adjacent bars.

Example:4

Find the bar graph for the frequency distribution of the last example.



The bar graphs for relative frequency and percentage distribution can be drawn simply by marking the relative frequencies or percentages, instead of the class frequencies, on the relative axis.

Some times a bar graph is constructed by marking the categories on the vertical axis and the frequency on the horizontal axis.

Pie charts:

Def:

A circle divided into portions that represent the relative frequencies or percentage of a population or a sample belonging to different categories is called a pie chart.

A pie chart is more commonly used to display percentage although it can be used to display frequencies or relative frequencies. The whole pie(or circle) represent the total sample or population. Then we divide the pie into different portion that represent the different categories.

As we know, a circle contain (360) degree. To construct a pie chart, we multiply (360) by the relative frequency of each category to obtain the degree measure or size of the angle for the corresponding category.

Example:5

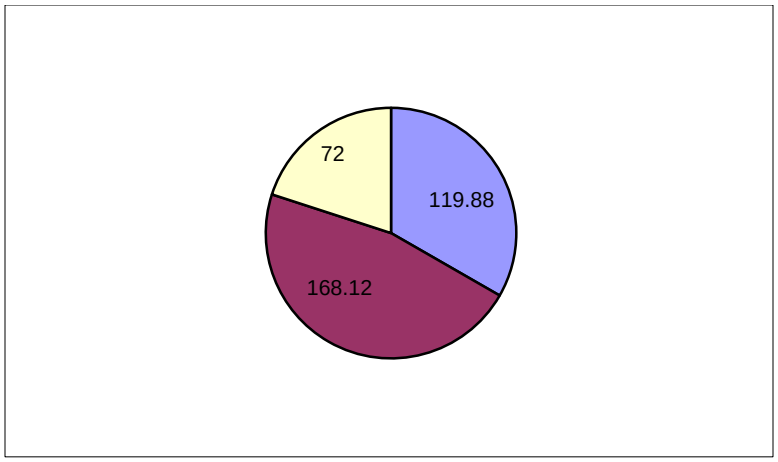
For the last example find pie chart.

Solution:

The following table shows the calculation of angle size for the categories.

Calculating angle size for the pie chart

Stress on job	Relative frequency	Angle size
Very	0.333	360*0.333=119.88
Some what	0.467	360*0.467=168.12
None	0.200	360*0.200=72.00
	Sum=1.00	Sum=360



Example 6:

The following data give the results of a sample survey the letters A, B and C represent the three categories.

A B B A C B C C C A

C B C A C C B C C A

A B C C B C B A C A

a) Prepare a frequency distribution table.

b) Calculate the relative frequencies and percentages for all categories.

c) What percentage of the elements in this sample belong to category B?

d) What percentage of the elements in this sample belongs to category A, C?

e) Draw a bar graph for the frequency distribution.

Example 7:

The following data give the results of a sample survey. The letters N, Y and D represent the three categories.

D N N Y Y Y N Y D Y

Y Y Y Y N Y Y N N Y

N Y Y N D N Y Y Y Y

Y Y N N Y Y N N D Y

a) Prepare a frequency distribution table.

b) Calculate the relative frequencies and percentages for all categories.

c) What percentage of the elements in this sample belong to category Y?

d) What percentage of the elements in this sample belongs to category N, D?

e) Draw a Pie chart for the relative distribution.

Organizing and Graphing Quantitative Data:

Frequency Distribution for quantitative Data:

A frequency distribution for quantitative data lists all the classes and the number of values that belong to each class.

Data represented in the form of a frequency distribution are called grouped data.

Class Boundary:

The class boundary is given by the midpoint of the upper limit of one class and the lower limit of the next class.

Class width:

Class width = upper boundary – lower boundary

Class midpoint or mark:

Class midpoint or mark = (lower limit + upper limit) / 2

For example:8

The following table gives the weekly earning of (100) employees of a large company, the first column lists the classes, which represent the (quantitative)

Weekly earning (variable)	Number of employees (frequency)
401 to 600	9
601 to 800	22
801 to 1000	39
1001 to 1200	15
1201 to 1400	9
1401 to 1600	6

variable weekly earnings. For quantitative data, an interval that includes all the values that falls within two numbers, the lower and upper limits is called a class.

Note that the classes always represent a variable. As we can observe the classes are non overlapping; that is, each value on earnings belong to one and only one class. The second column in the table lists the number of employees who have earnings within each class. For example, nine employees of this company earn
($ 401)to ($600) per week. The number listed in the second column is called the frequencies, which give the number of values that belong to different classes. The frequencies are denoted by (fi.
Weekly earning of employees of a company

For quantitative data, the frequency of a class represents the number of values in the data set that fall in the classes. The above table contains six classes. Each class has a lower limit and an upper limit. The values (401,601,801,1001,1201 and 1401)give the lower limits. And the values (600,800,1000,1200,1400 and 1600) are the upper limits. The data presented in the above table are an illustration of a frequency distribution table for quantitative data. Where as the data list individual values are called ungrouped data, the data represent in a frequency distribution table are called grouped data.

To find the midpoint of the upper limit of the first class and the lower limit of the second class in the table, we divide the sum of these two limits by 2 thus:

this midpoint is

$$= (600+601)/2 = 600.5$$

The value (600.5) is called the **upper boundary** of the first class and the **lower boundary** of the second class. By using this technique, we can convert the class limits of the table to class boundaries, which are also called real class limits .The second column of the table lists the boundaries for another table.

The difference between the two boundaries of a class gives the class width. The class width is also called the class size.

Thus from the table:

Width of the first class= 600.5-400.5 = 200

The class width for the frequency distribution of the above table are listed in the third column of the following table . Each class in the following table (and above table) has the same width of 200.

The class midpoint or mark

is obtained by dividing the sum of two limits (or the two boundaries) of a class by 2.

Thus, the midpoint of the first class in above table (or following table) is calculated as follows.

Midpoint of the first class = (401+600)/2 = 500.5

The class midpoints for the frequency distribution of above table are listed in the fourth column of the following table.

Class Boundaries, class widths, and Class midpoints for above table

Class limits	Class Boundary	Class width	Class midpoint
401 to 600	400.5 to less than 600.5	200	500.5
601 to 800	600.5 to less than 800.5	200	700.5
801 to 1000	800.5 to less than 1000.5	200	900.5
1001 to 1200	1000.5 to less than 1200.5	200	1100.5
1201 to 1400	1200.5 to less than 1400.5	200	1300.5
1401 to 1600	1400.5 to less than 1600.5	200	1500.5

Constructing frequency distribution table:

When constructing a frequency distribution table, we need to make the following *three major decisions*:

1- ***number of classes*****:**

Usually the number of classes for a frequency distribution table varies from 5 to 20, depending mainly on the number of observations in the data set.

It is preferable to have more classes as the size of data set increases. The decision about the number of classes is arbitrarily made by the data organizer.

One rule to help decide on the number of classes is Sturges formula.

$$C = 1 + 3.3 \log n$$

Where C is the number of classes and n is the number of observations in the data set.

2- ***Class width*****:**

First find the difference between largest value (X_L) and smallest value (X_S) in the data, then the approximate width of class is obtained by dividing this difference by the number of desired classes.

Approximate class width = (largest value – smallest value) / number of classe

3- ***Lower limit of the first class or the start point:***

Any convenient number that is equal to or less that the smallest value in the data set can be used as the lower limit of the first class.

Example:9

Construct a frequency distribution table for the following data.

135, 178 , 169 , 222 , 235 , 242 , 194 , 184 , 202 , 201 , 148 , 187 , 150 , 162 , 203 , 135 , 191 , 151 , 185 , 242 , 189 , 215 , 142 , 214 , 139 , 183 , 136 ,145 , 227 , 145

Relative Frequency and Percentage Distributions:

Ri = (frequency of that class) / (sum of all frequencies) = $\frac{f_i}{\sum f_i}$

Percentage (Pi) = (Ri)* 100

Graphing Grouped Data:

Grouped data can be displayed in a histogram or a polygon. We can also draw a pie chart to display the percentage distribution for a quantitative data set.

1- Histogram

Def: A histogram is a graph in which classes are marked on the horizontal axis and the frequencies, relative frequencies or percentages are marked on the vertical axis. The frequencies, relative frequencies, or percentages are represented by the heights if the bars. In a histogram the bars are drawn adjacent to each other.

Example:10

For the last example.

2-*Polygon*

Def: A graph formed by joining the midpoints of the tops of successive bars in a histogram with straight lines is called a polygon.

A polygon is another device that can be used to present quantitative data in graphic form. To draw a frequency polygon we first mark a dot above the midpoint of each class at a height equal to the frequency of that class. This is the same as making the midpoint at the top of each bar in a histogram. Next we mark two more classes, on at each end, and mark their midpoints.

Note that these two classes have zero frequencies.

In the last step, we join the adjacent dots with straight lines. The resulting line graph is called a frequency polygon, or simply a polygon.

Example:11

Find the frequency polygon for the frequency distribution of the last example.

Note:

For a very large data set, as the number of classes is increased (and the width of classes decreased) the frequency polygon eventually becomes a smooth curve. Such a curve is called a frequency distribution curve or simply a frequency curve. The following figure shows the frequency curve for a large data set with a large number of classes.

Less than method for writing classes:

The classes in the frequency distribution in the last example were written as
135 – 156
157 – 178 and so on
Alternatively, we can write the classes in a frequency distribution table using the less than method.
The technique for writing classes shown in the last example is more commonly used for data that do not contain fractional values.
The less than method is more appropriate when a data set contains fractional values.

Example: 12
According to the American petroleum institute, the state taxes (in cents) per gallon of gasoline as of April 2005 for all (50) sates are as follows.(C=6)

,8, 18 , 21.5 , 18, 22, 25, 23 , 14.5, 7.5,
,25, 19 , 18 , 20, 24 , 16 ,20 , 25.5 , 23.5,
23.5 , 19, 20, 18 , 17 , 27.5 ,25.5 , 25.4 , 23 , 18 ,14.5,
17, 31.5 , 26.6 ,21 , 26 , 16 ,24 , 31.1 , 30 , 16 ,
22 , 20 , 20 , 24.5, 20 , 17.5 , 28 , 20.5 , 32.5 , 32.9 ,14 ,

Construct a frequency distribution table. Calculate the relative frequencies and percentages for all classes.

Single – Valued classes:

If the observations in a data set assume only a few distinct (integer) values, it may be appropriate to prepare a frequency distribution table using Single- valued classes- that is, classes that are made of single value and not of intervals. This technique is especially useful in cases of discrete data with only a few possible values.

Example:13

The administration in a large city wanted to known the distribution of vehicles owned by households in that city. A sample of (40) randomly selected household from this city produced the following data on the number of vehicles owned.

5 ,1 , 1 ,2 ,0 , 1, 1 , 2 , 1 , 1,
1 ,3, 3, 0, 2 , 5 , 1 ,2, 3 , 4 ,
2 ,1 , 2 , 2 ,1 , 2 ,2 , 1 ,1 ,1 .
3 ,2 ,1 ,1 ,2 ,1 ,1 , 4 ,1 ,3,

Construct a frequency distribution table fore these data using single- valued classes.

Cumulative Frequency Distribution:

A Cumulative frequency distribution gives the total number of value that fail below the upper boundary of each class.

For the last example:.

Suppose we want to know how many that be a total of (200). Such question can be answered using cumulative frequency distribution. Each class in a cumulative frequency distribution table gives the total number of values that fall below a certain value .A cumulative frequency distribution is

Constructed for quantitative data only

In a cumulative frequency distribution table, each class has the same lower limit but a different upper limit.

Example:14

Using the following frequency distribution, prepare a cumulative frequency distribution.

Class	Frequency
135-156	10
157-178	3
179-200	7
201-222	6
223-244	4

Solution:

Class limit	Class boundaries	Cumulative frequency
135≤156	134.5 less that 156.5	10
135≤178	134.5 less that 175.5	10+3=13
135≤200	134.5 less that 200.5	10+3+7=20
135≤222	134.5 less that 222.5	10+3+7+6=26
135≤244	134.5 less that 244.5	10+3+7+6+4=30

From the above table, we can determine the number of observations that all below, the upper limit or boundary of each class.

Cumulative Relative and Percentage frequency:

The Cumulative relative frequencies are obtained by dividing the cumulative frequencies by the total number of observations in the data set.

$$\textit{Cumulative relative frequency} \frac{\textit{cumulative frequency of a class}}{\textit{total} \text{ observations in the data set}}$$

The cumulative percentages are obtain by multiplying the cumulative relative frequencies by (100)

Cumulative percentage = (cumulative relative frequency) * 100

Class limit	Cumulative relative frequency	Cumulative percentage
153≤156	10/30 = 0.333	33.3
135≤178	13/30 = 0.433	43.3
135≤200	20/30 = 0.667	66.7
135≤222	26/30 = 0.867	86.7
135≤244	30/30 = 1.000	100

Ogives:

Def:

An ogives is a curve drawn for the cumulative frequency distribution by joining with straight lines the dots marked above the upper of classes at heights equal to the cumulative frequencies of respective classes.

Example:15

To draw the ogive for the above table, the variable is marked on the horizontal axis and the cumulative frequencies on the vertical axis . Then the dots are marked above the upper boundaries of various classes at the heights equal to the corresponding cumulative frequencies.

The ogive is obtained by joining consecutive points with straight lines. Note that the ogive starts at the lower boundary of the first class and ends at the upper boundary of the last class.

Note:

We can draw an ogive for cumulative relative frequency and cumulative percentage distributions, the same way we did for the cumulative frequency distribution.

Cihan University –Sulaimanya Camp

Principles of Statistics
Topic –TWO-

Tabular and Graphical Methods

Exercises & solving

2023-2024

Prof.Dr. Obaid Mahmmood Mohsin

Principles of statistics

Topic 2

Exercises

1-Suppose we ask the same (50)students about their student status .The responses of the students are recorded in following .F(fresh man) ,So (sophomore) ,J(junior) and SE(senior)

J F SO SE J J SE J J J

F F J F F F SE SO SE J

J F SE SO SO F J F SE SE

SO SE J SO SO J J SO F SO

SE SE F SE J SO F J SO SO

(i) Prepare a frequency distribution table .
(ii) Calculate the relative frequencies and percentages for all categories .
(iii) What percentage of these students are juniors and seniors ?
(iv) Draw a bar graph for the frequency distribution.

2-The following data show the method of payment by (16)customers in a supermarket checkout line:

Hence C refers to cash

CC refers to credit card

CK refers to check

D refers to debit card

O refers to stands for other

C CK CK C CC D O C

CK CC D CC C CK CK CC

(i) Construct a frequency distribution table
(ii) Calculate the relative frequencies and percentages for all categories.
(iii) Draw a pie chart for the percentage distribution .

3-The response to a question has three alternatives A,B ,and C a sample of (120) responses provides 60 A, 24 B ,36 C ,show the frequency and relative frequency distribution.

4- A partial relative frequency distribution is given :

class	Relative frequency
A	0.22
B	0.18
C	0.40
D	

(i) What is the relative frequency of class D?
(ii) The total sample size is (200) ,what is the frequency of class D ?
(iii) Show the frequency distribution table .
(iv) Show the percentage frequency distribution .

5- A questionnaire provides (58) yes , (42) no , and (20) no opinion answers:

(i) In the construction of a (pie chart) how many degrees would be in the section of the pie showing (yes) answer ?
(ii) How many degrees would be in the section of the pie showing the (no) answer ?
(iii) Construct a pie chart .
(iv) Construct a bar chart .

6- Consider the following data :

14 19 24 19 16 20 24 20 21 22 24

26 22 23 25 19 18 16 15 24 21 16

20 22 22 16 16 16 21 25 19 24 20

18 19 17 21 23 23

(i) Develop a frequency distribution using class of , 12-14 ,15-17 ,18-20 ,21-23 ,and 24-26.

(ii) Develop a relative frequency distribution and percentage frequency distribution using the class in part (i).

7-Consider the following frequency distribution :

class	frequency
10-19	10
20-29	14
30-39	17
40-49	7
50-59	2

(i) Construct a cumulative frequency distribution and relative frequency distribution .

(ii) Construct a histogram and ogive for (i) .

8-A doctor' s office staff studied the waiting times for patients who arrive at the office with a request for emergency service the following data with waiting times in minute were collected over a one minute period:

2 ,5 ,10 , 12 , 4 , 4 , 5 , 17 , 11 , 8 , 9 , 8 , 12 , 21 , 6 , 8 , 7 , 13 , 18 , 3 ,

Use classes of : 0-4 ,5-9 and so on in the following :

(i) Show the frequency distribution .

(ii) Show the relative distribution .

(iii) Show the cumulative relative distribution .

(iv) What proportion of patients needing emergency service wait (9) minutes or less .

9-The following data provide the dollar amount of holiday spending for a sample of (25) consumers :

1200 850 740 590 340 450 890 260 610 350

1780 180 850 2050 770 800 1090 510 520 220

1450 280 1120 200 350

(i) What is the lowest holiday spending ? The highest ?

(ii) Use a class width of (250)$ to prepare a frequency distribution and a percentage frequency distribution for the data .
(iii) Prepare a histogram for (i).
(iv) What observation can you make about holiday spending ?

10-Storing through unsolicited e-mail and spam affects the productivity of office worker .An insight express survey monitored office workers to determine the unproductive time per day devoted to unsolicited e-mail and spam .The following data show a sample of time in minutes devoted to this task:

2 , 8 , 12 , 5 , 24 , 4 , 1 , 1 , 5 , 19 , 8 , 2 , 5 , 3 , 4 , 4 , 32 , 7 , 4 , 14 ,

Summarize the data by constructing the following :

(i) A frequency distribution , (class , 1-5 ,6-10 ,11-15 ,16-20 ,and so on).
(ii) A relative frequency distribution .
(iii) A cumulative frequency distribution .
(iv) A cumulative relative frequency distribution .
(v) An ogive.
(vi) What percentage of office workers spends (5) minuets or less on un solicited e-mail and spam ?What percentage of office workers spend more than (10) minuets a day on this task ?

11-The following data are the hours of personal computer usage during one week for a sample of (50) persons .

4.1 1.5 10.4 5.9 3.4 5.7 1.6 6.1 3.0 3.7 3.1 4.8 2.0 14 5.4 4.2 3.9 4.1 11.1 3.5 4.1 4.1 8.8 5.6 4.3 3.3 7.1 10.3 6.2 7.6 10.8 2.8 9.5 12.9 12.1 0.7 4.0 9.2 4.4 5.7 7.2 6.1 5.7 5.9 4.7 3.9 3.7 3.1 6.1 3.1

Summarize the data by constructing the following :

(i) A frequency distribution (use a class width of three hours) .
(ii) Relative frequency distribution.
(iii) A histogram .
(iv) An ogive .
(v) Comment on what the data indicate about personal computer usage at home .

12-The following data give the results of a sample survey .The letters A ,B and ,C represent the three categories :

A	B	B	A	C	B	C	C	C	A
C	B	C	A	C	C	B	C	C	A
A	B	C	C	B	C	B	A	C	A

(i) Prepare a frequency distribution table .

(ii) Calculate the relative frequency and percentage for all categories .

(iii) What percentage of the elements in this sample belongs to category B ?

(iv) What percentage of the elements in this sample belong to category A and C .

(v) Draw a bar graph for the frequency distribution .

13- The following data give the results of a sample survey. The letters Y ,N and ,D represent the three categories:

D	N	N	Y	Y	Y	N	Y	D	Y
Y	Y	Y	Y	N	Y	Y	N	N	Y
N	Y	Y	N	D	N	Y	Y	Y	Y
Y	Y	N	N	Y	Y	N	N	D	Y

(i)prepare a frequency distribution table.

(ii)Calculate the relative frequency and percentage .

(iii) What percentage of the elements in this sample belongs to category A and D .

(iv) Draw a pie chart for the relative distribution .

14-For the following table:

category	Fi	Ri
A	12	0.4
B		
C	5	
D	6	
E		0.1
sum		

(i)Complete the table.

(ii) Find the percentage distribution.

(iii)Draw the bar chare for the frequency distribution.

15- The following data give the results of a sample survey. The letters Y, and, Z ,and ,W represent the three categories :

Z W Y W Y Y W Y Z Y Y Y W W Y Y W Z Y Y Y W Y W Y W Y W W Y Z Y W Y W Y Y W W W

(i) Prepare a frequency distribution table.
(ii) Calculate the relative frequencies and percentages for all categories .
(iii) Show the : variable and its type , category ,frequency .
(iv) What percentage of the elements in this a sample belongs to category ,Z and ,W ?
(v) Draw a pie chart for the relative frequency distribution .

Principles of statistics

Topic 2

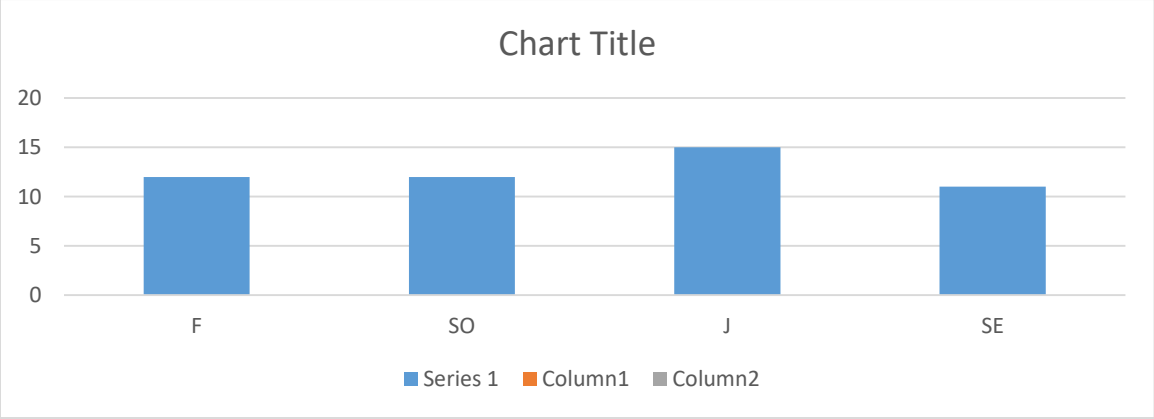
Solving the Exercises

1-

Category	Tally	fi	Ri	Pi
F	𝍸 𝍸 ||	12	12/50=0.24	24
SO	𝍸 𝍸 ||	12	0.24	24
J	𝍸 𝍸 𝍸	15	0.30	30
SE	𝍸 𝍸 |	11	0.22	22
		50	1	100

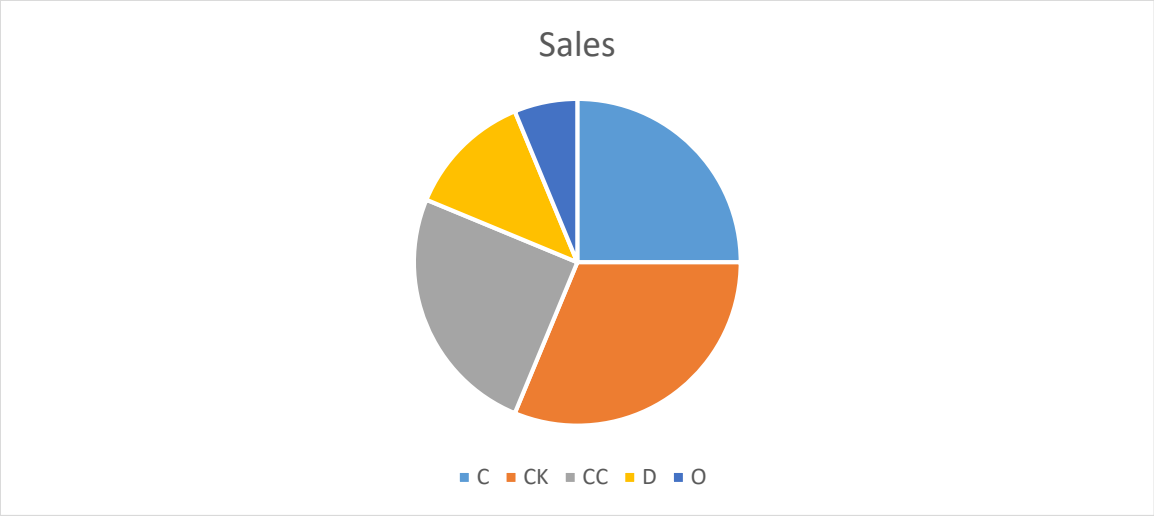
J+SE = 30%+22% =52%

Bar chart for frequency distribution



2-

Category	Tally	fi	Ri	Pi	Angel size
C	||||	4	0.25	25	90
CK	𝍸	5	0.31	31	111.6
CC	||||	4	0.25	25	90
D	||	2	0.125	12.5	45
O	|	1	0.063	6.3	22.7
sum		16	1	100	



3-

Category	fi	Ri
A	60	0.5
B	24	0.2
C	36	0.3
	120	

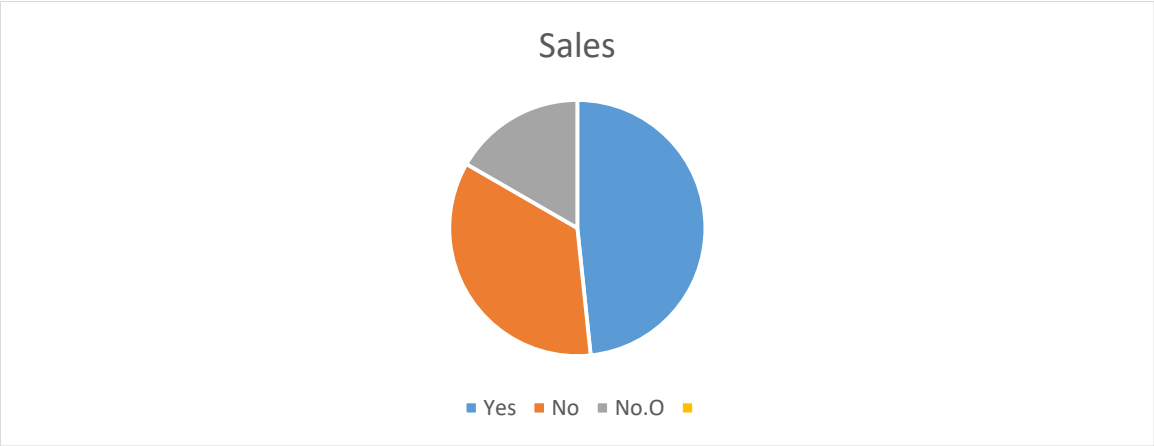
4- Ri of D = 1-{ 0.22 +0.18 +0.40 } = 0.20

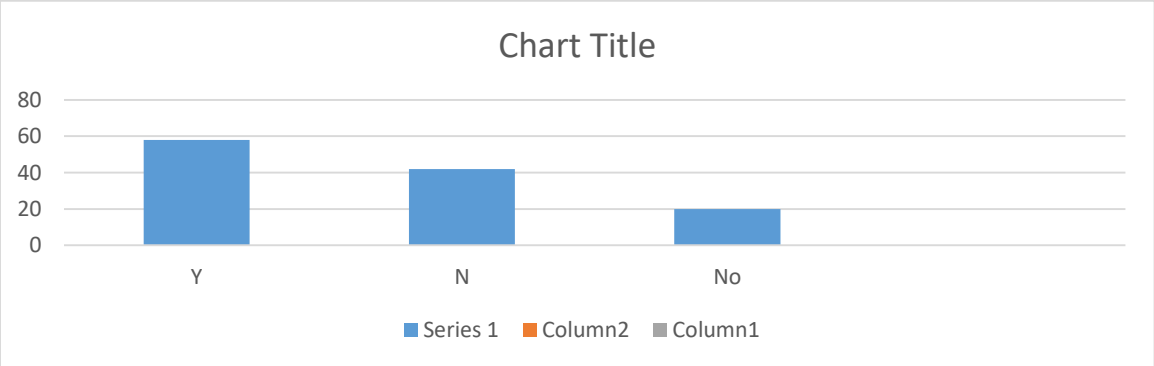
Fi of D = (0.20) * (200) =40

Class	fi	Ri	Pi
A	44	0.22	22
B	36	0.18	18
C	80	0.40	40
D	40	0.20	20
	200	1	100

5-

Category	fi	Ri	Pi	Angel size
Yes	58	0.483	48.3	174
No	42	0.35	35	126
No.O	20	0.167	16.7	60
	120	1		



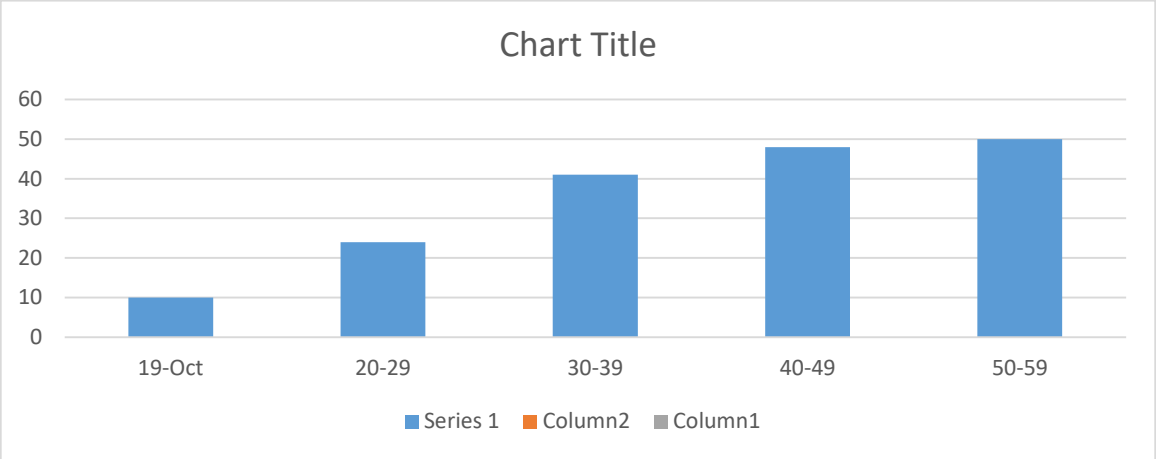


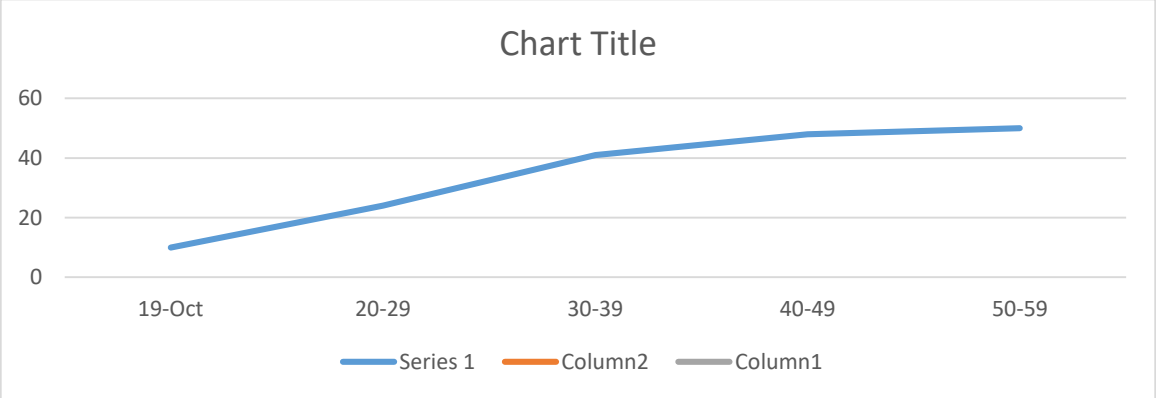
6-

Class	Tally	fi	Ri	Pi
12-14	|	1	0.026	2.6
15-17	𝍸 |||	8	0.205	20.5
18-20	𝍸 𝍸 |	11	0.282	28.2
21-23	𝍸 𝍸 |	11	0.282	28.2
24-26	𝍸 |||	8	0.205	20.5
		39	1	100

7-

Class	fi	Upper limit	Cum.	Ri C
10-19	10	≤19	10	0.2
20-29	14	≤29	24	0.48
30-39	17	≤39	41	0.82
40-49	7	≤49	48	0.96
50-59	2	≤59	50	1
	50			





8-

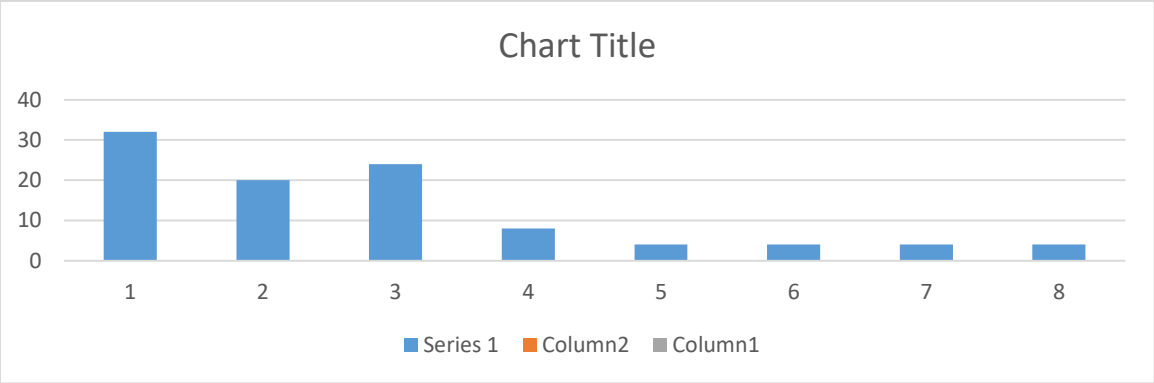
Class	Tally	fi	Ri	Upper limit	Cum	RiC
0-4	||||	4	0.2	≤4	4	0.2
5-9	𝍸|||	8	0.4	≤ 9	12	0.6
10-14	𝍸	5	0.25	≤ 14	17	0.85
15-19	||	2	0.1	≤ 19	19	0.95
20-24	|	1	0.05	≤ 24	20	1
		20				

20+40= 60%

9-The lowest holiday spending =180

The highest holiday spending = 2050

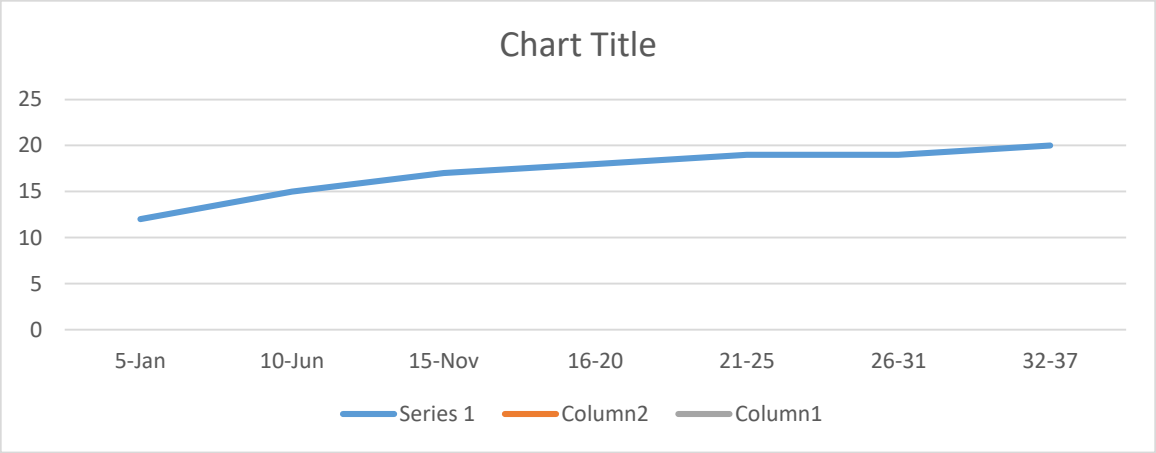
Class	Tally	fi	Ri	Pi
150-	𝍸|||	8	0.32	32
400-	𝍸	5	0.2	20
650-	𝍸|	6	0.24	24
900-	||	2	0.08	8
1150-	|	1	0.04	4
1400-	|	1	0.04	4
1650-	|	1	0.04	4
1900-2149	|	1	0.04	4
		25	1	100



We have 76% between 150-900 less than 900 .

10-

Class	Tally	fi	Ri	Cum	Cum Ri
1-5	𝍸 𝍸 ||	12	0.6	12	0.6
6-10	|||	3	0.15	15	0.75
11-15	||	2	0.1	17	0.85
16-20	|	1	0.05	18	0.9
21-25	|	1	0.05	19	0.95
26-31		0	0	19	0.95
32-37	|	1	0.05	20	1
		20	1		



≤ 5 = 60%

More than 10 = 25%

11-

Largest value =14

Smallest value =0.7

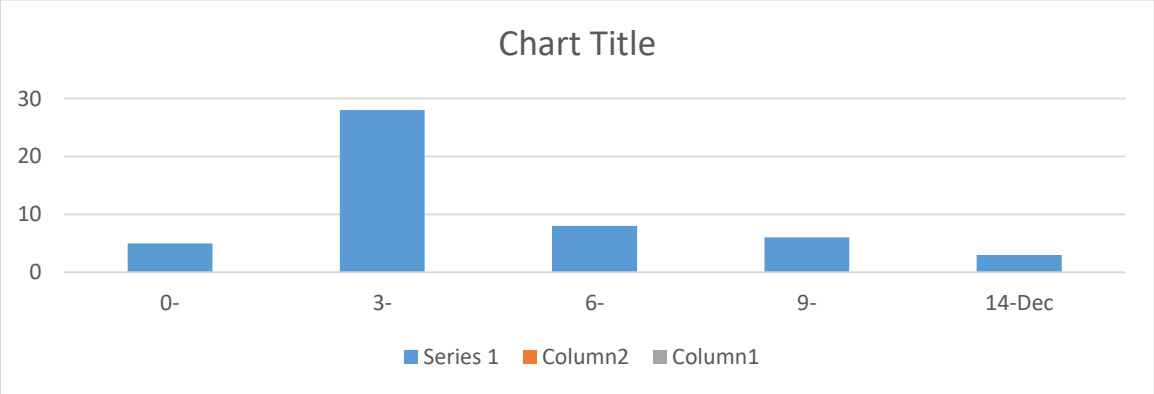
W =(14- 0.7) /C

3 = (14 -0.7)/C

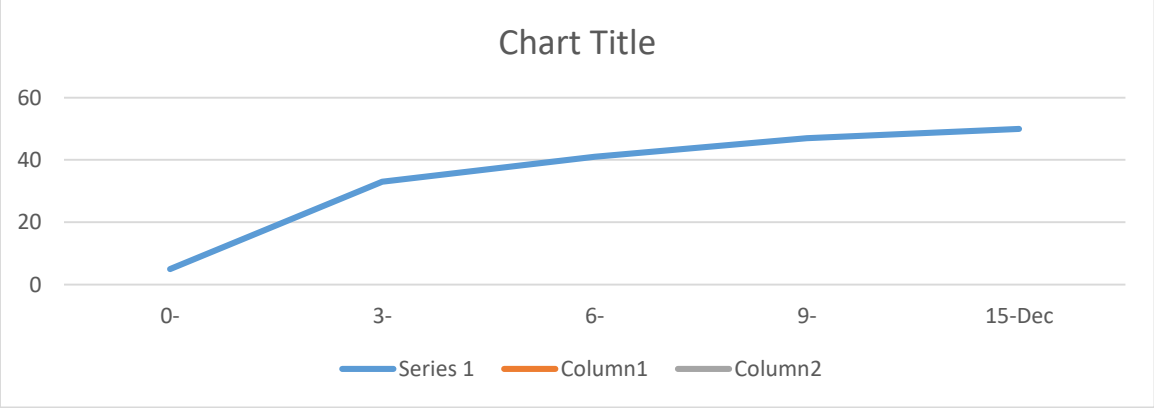
C=13.3/3= 5

Class	Tally	fi	Ri	Cum
0-	𝍸	5	0.1	5
3-	𝍸 𝍸 𝍸 𝍸 𝍸 |||	28	0.56	33
6-	𝍸 |||	8	0.16	41
9-	𝍸 |	6	0.12	47
12-14	|||	3	0.06	50
		50	1	

Histogram:



Ogive:

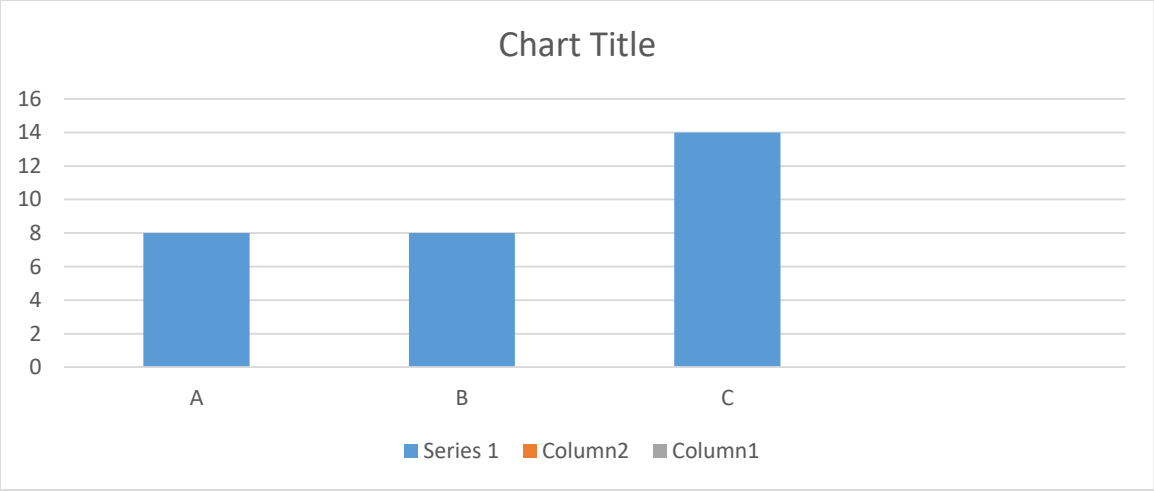


12-

Category	Tally	fi	Ri	Pi
A	𝍸 |||	8	0.267	26.7
B	𝍸 |||	8	0.267	26.7
C	𝍸 𝍸 ||||	14	0.467	46.7
sum		30		

(iii) Percentage of B = 26.7%

(iv) Percentage of A and C =26.7% +46.7% =73.4%



13-

Category	Tally	fi	Ri	Pi
Y	𝍸 𝍸 𝍸 𝍸 |||	23	Ri = fi/ Σ fi =0.575	(0.575)(100)=57.5
N	𝍸 𝍸 |||	13	13/40=0.325	32.5
D	||||	4	4/40=0.100	10.0
sum		40	1.00	100

(iii)percentage of Y =57.5%

Percentage of N and D = 32.5% + 10% =42.5 %

(iv)Angle size

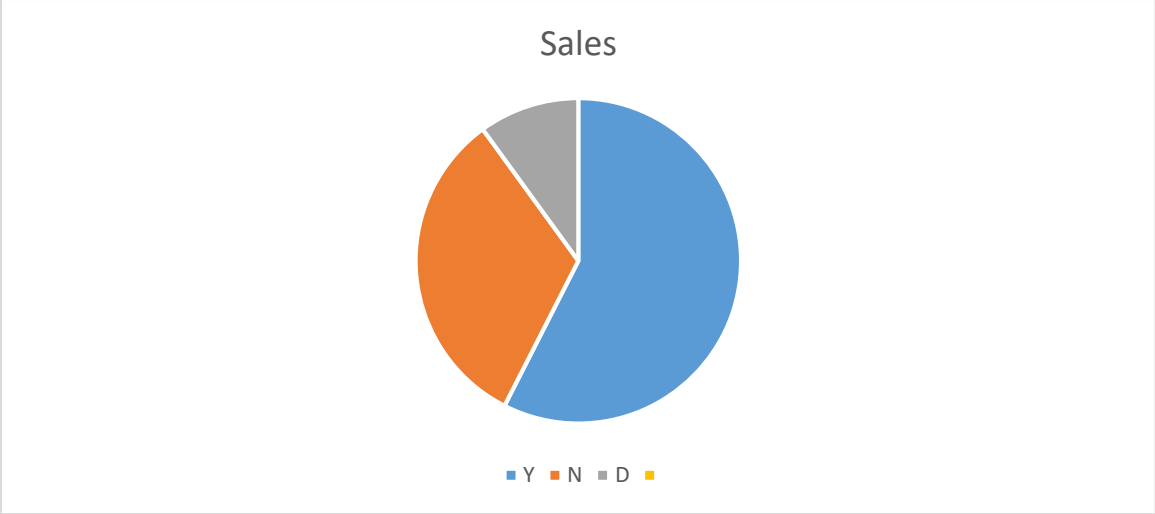
(360)(0.575) =207

(360) (0.325) =117

(360) (0.100) =36

(360) (0.325) =117

(360) (0.100) =36



14-

Ri = fi /Σfi-------R1 = f1/ Σfi----0.4=12/ Σfi -----Σfi =12/0.4 =30

R3 =f3/Σfi=5/30 =0.17

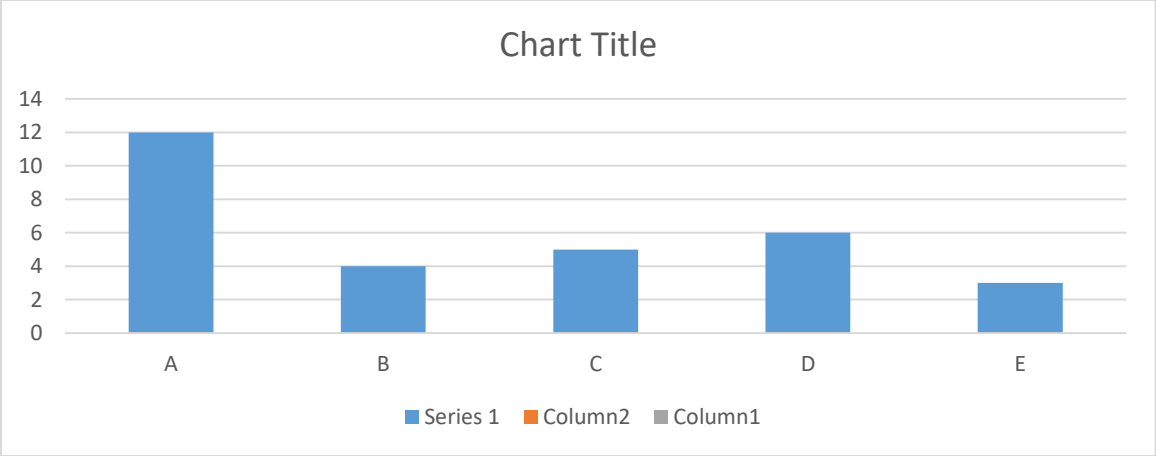
R4 =6/30 =0.2

R2 =1-{ R1 +R3 +R4 +R5 } = 1- {0.4 +0.17 +0.2 +0.1 } =0.13

R5 =f5/ Σfi ----------0.1 =f5/30 -------f5=3

F2 = 30 – { 12 +5 +6 +3 } =4

Category	fi	Ri	Pi
A	12	0.4	40
B	4	0.13	13
C	5	0.17	17
D	6	0.2	20
E	3	0.1	10
sum	30	1.0	100



15- (i)

Category	Tally	fi	Ri	Pi
Y	𝍸 𝍸 𝍸 𝍸	20	0.5	50
Z	||||	4	0.1	10
W	𝍸 𝍸 𝍸 |	16	0.4	40
sum		40	1	100

(ii)Relative frequency Ri = fi / Σfi

R1 =20/40 =0 .5

R2 =4/40 = 0.1

R3 =16/40 =0.4

Percentage Pi = (Ri) (100)

P1= (0.5) (100) =50

P1 =(0.1) (100)= 10

P3 = (0.4) (100)= 40

(iii)The variable : categories and its type is qualitative variable.

Category :Z

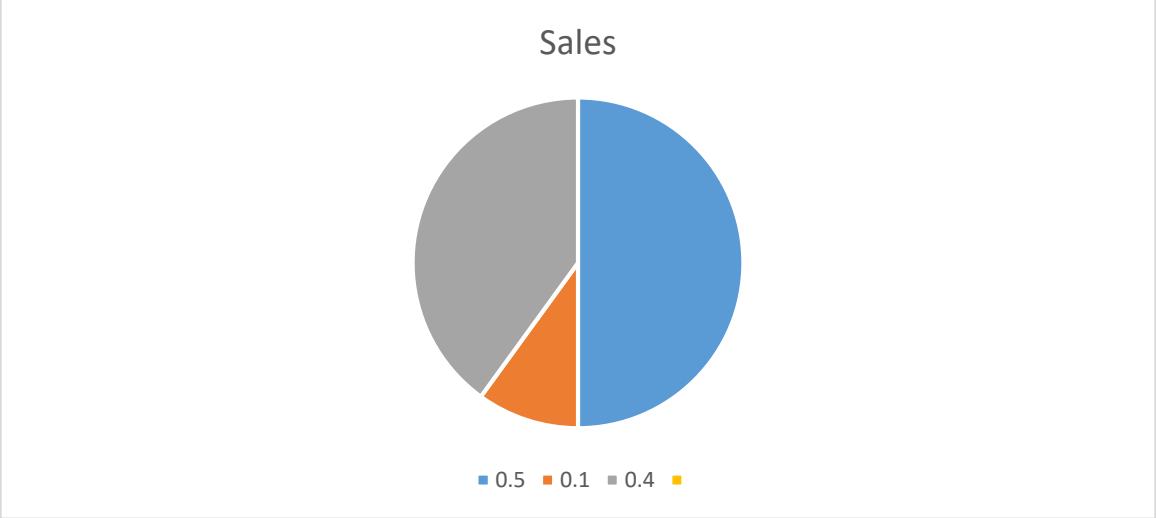
Frequency : 20

(iv)The percentage of the elements in this sample belong to category Y =50%

The percentage of the elements in this sample belong to category Z and W = 10%+40% =50%

(v)

Ri	Angle size
0.5	(360)(0.5)=180
0.1	(360)(0. 1)=36
0.4	(360)(0.4)=144
	360



Cihan University-Sulaimanya Camp

Principles of Statistics

Topic –THREE-

Measures of central tendency

Prof.Dr. Obaid Mahmmood Mohsin

2023-2024

Ungrouped data

i. Grouped data

ii. Exercises

Topic 3

Measures of Central Tendency

*I- **Ungrouped Data***

1- Mean, Arithmetic mean, Average

Mean = Sum of Values / Number of Values

Mean for Population $$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Mean for sample $$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Where

$\sum x_i$ The sum of all values

$(x_1 + x_2 + ... + x_N)$ or $(x_1 + x_2 + ... + x_n)$

N= is the Population size

n= is the sample size

Example:1

The following are the ages of all eight employees of a small company.

53, 32, 61, 27, 39, 44, 49, 57

i. Find the mean age of these employees.

ii. If we take sample of three employees from this company suppose the three values included in the sample are 32, 39, 57. Find the mean:

The weighted mean:

The weighted mean is a special case of the arithmetic mean.

$$\overline{x}_w = \frac{\sum w_i x_i}{\sum w_i}$$

Example: 2

A company pays its hourly employees $ 16.5, $ 17.5 or $18.5 per hour. There are 26 hourly employees, 14 are paid at the 16.5 rate, 10 at the 17.5 rate and 2 at the 18.5 rate. What is the mean hourly rate paid the 26 employees?

2- Median:

Another important measure of central tendency is the median. Its defined as follows:

Median is the value of the middle in a data set that has been ranked in increasing order.

As is obvious from the definition of the median, it divides a ranked data set into two equal parts. The calculation of the median consists of the following two steps.

1. Rank the data set in increasing order.
2. Find the middle term. The value of this term is the median.

Note:

If the number of observation in data set is odd. Then the median is given by the value of the middle term in the ranked data. However, if the number of observations is Even then the median is given by the average of the values of the two meddle terms.

Example:3

The following data give the weight lost (in Kg) by a sample of five member of a health club at the end of two months of membership.

10, 5, 19, 8, 3

Find the median.

Example:4

For the following data in increasing order as follows

11.669, 13.435, 14.413 , 21.088, 29.920, 42.082, 33.956 , 40.179, 18.215, 18.103, 26.343,40.769

Find the median.

3- Mode:

The value of the observation that appear most frequently.

Example:5

The following data give the speeds (in miles per hour) of eight cars.

77, 82, 74, 81, 79, 84, 74, 78

Sol:

Mode = 74 miles per hour

Example:6

Find the mode of each set of the following data set

i. 4, 3, 5, 7, 8, 7, 10, 15

ii. 2, 1, 7, 3, 2, 3, 10, 12

iii. 2, 1, 4, 8, 9, 3, 5, 14

II- Grouped Data

1- mean :

For population $\mu = \frac{\sum_{i=1}^{N} m_i f_i}{N}$

For sample $\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i}$ or $\frac{\sum_{i=1}^{n} m_i f_i}{n}$

m_i = is the midpoint of class i.

f_i = is the frequency of class i.

Example:7

The following table gives the frequency distribution of orders received each day during the past 50 days at the office of a mail – Order Company.

Number of orders(class)	Number of days (frequency)
10- 13- 16- 19-21	4 12 20 14
	50

Calculate the mean.

2- Median:

$$median\ (Me) = A + (\frac{T - f_1}{f_2 - f_1})w_i$$

Where

Li = lower class limit of median class

T = arrangement of the median and $T = \sum f_i / 2$

f_1 = cumulative frequency precedes median

f_2 = cumulative frequency succeeds median

w_i = width of median class

To calculate median for grouped data, we follow these steps:

1- We find the cumulative frequency distribution.

2- We find the median arrangement T that $T = \sum f_i / 2$

3- We determine median class interval. It is the one oppose the (cumulative frequency that succeeds median arrangement immediately)

Example:8

For the following frequency distribution table, compute median

classes	frequency
50-	8
60-	10
70-	16
80-	14
90-	10
100-	5
110-120	2

Solution:

Classes	frequency	Upper limits of classes	Cumulative frequency
50-	8	< 60	8
60-	10	<70	18
70-	16	<80	34
80-	14	<90	48
90-	10	<100	58
100-	5	<110	63
110-120	2	≤120	65

T = 65/2 =32.5

Li = 70

f_1 = 18

f_2 = 34

w_i = 10

$$Me = \text{A} + (\frac{T - f_1}{f_2 - f_1})w_i$$

$$Me = 70 + (\frac{32.5 - 18}{34 - 18}) * (10) = 79.1$$

3- Mode:

$$Mo = A + (\frac{d_1}{d_1 + d_2}) * w_i$$

Where

Li= lower class limit of the mode class interval

d_1 =difference between frequency of the mode class interval and the one precedes it .

d_2 difference between frequency of the mode class interval and the one succeeds it .

w_i = width of mode class

Example:9

For the last example find mode

Mode class = 70 (which has the more frequency)

w_i = 10

d_1 = 16-10 = 6

d_2 = 16-14=2

$$Me = 70 + (\frac{6}{6+2})*10 = 77.5$$

Example:10

The following table represents the monthly expense on medicine of a sample of families (80) families. That class intervals represent expense per(1000 ID), frequency represent families number.

Classes	Frequency
50-	10
70-	10
90-	30
110-	20
130-149	10
	80

Find the : Mean, Median ,and Mode

Cihan University –Sulaimanya Camp

Principles of Statistics
Topic –THREE-

Measures of central tendency

Exercises & solving

2023-2024

Prof.Dr. Obaid Mahmmood Mohsin

Principles of Statistics

Topic –3-

Exercises

1- The following data represent result obtained by repeated experiment ten-time :

37 , 35 , 36 , 38 , 36 , 36 , 35 , 37 , 37 , 36 .

Find : Mean ,Median and Mode .

2- The following data represent the weekly over wages paid for (36) employees in an industrial enterprise :

16 12 21 17 13 14 14 18 8 15 22 19 17 12 16 23 15 5

13 24 9 19 24 16 20 17 14 10 18 25 19 12 16 21 11 13

Required:

(i) Frequency distribution of three class .

(ii) Relative frequency and percentage and cumulative distribution table .

(iii) Mean , Median and , Mode for (i) .

3-For the following frequency distribution table :

Find , mean ,median ,mode

Class	fi
33-	10
38-	12
43-	51
48-	30
53-57	8
sum	111

4-The following data represent the period of storing a set of materials in a store per day :

45 , 49 , 10 , 11 , 22 , 39 , 25 , 33 , 20 , 24 , 45

Find : Mean , Median ,and ,Mode .

5-From the following frequency distribution table , find the mean , median ,and mode >

Class	fi
30-	1
40-	2
50-	5
60-	15
70-	25
80-	20
90-99	12
sum	80

6-The following table states frequency distribution of grades of (65) students .

Find : Mean , Median ,and ,Mode .

Class	fi
50-	8
55-	10
60-	16
65-	14
70-	10
75-	5
80-84	2
Σ	

7-The following table represents frequency distribution of overtime wages given to a sample of employees in a firm . Find : Mean , Median ,and ,Mode .

Class	fi
50-	8
60-	10
70-	16
80-	12
90-	12
100-	4
110-119	3
Σ	65

8- From the following frequency distribution table . Find : Mean ,Median .and ,Mode .

Class mark	fi
5	2
15	5
25	10
35	25
45	8
Σ	

9-The following frequency distribution table states the life time of electric bulbs production in a companies . Find : Mean ,Median ,and ,Mode .And what are the meaning

Class (age of bulb hour)	Fi (number of bulbs)
300-	14
400-	46
500-	58
600-	76
700-	68
800-	62
900-	48
1000-	22
1100-1199	6

10- The following frequency distribution table represents the weekly wages of employees in a factory . Find : Mean , Median ,and ,Mode .

Class (wages/$)	Fi (number employees)
60-	132
80-	159
100-	95
120-	42
140-159	22
Σ	450

11- The following data represent the period of storing specific product in a store per day during a year .

20 , 14 , 22 , 18 , 14 , 28 , 14 , 14 , 24 , 10 , 14 , 29

Find : Mean , Median ,and ,Mode .

12-In the following frequency distribution table classes represent the staying of a period of a sample of patients in a hospital per day , frequencies patients number .

Find : Mean , Median ,and ,Mode .

Class	fi
10-	6
15-	1
20-	2
25-	1
30-34	2
Σ	12

13-From the following frequency distribution table:

Class	fi
10-	8
20-	5
30-	7
Σ	20

Find : Mean , Median and ,Mode .

14- The following data represent the monthly income of commercial shop per thousand .

82 , 87 , 84 , 97 , 78 , 62 , 67 , 89 , 101 , 109 , 87 , 93 , 52 , 85 , 96 , 80 , 70 , 85 , 73 , 104 , 99 , 90 , 80

Find : (i) Find frequency distribution table (use 7 classes)

(ii) Relative and percentage distributions.

(iv) The mean , median ,mode for (i)

15-The following table represents the monthly expense on medicine of a sample of families (80) families . That class intervals represent expense per (1000 I.D) ,frequency represent families numbers .

Class	fi
50-	10
70-	10
90-	30
110-	20
130-149	10
Σ	80

Find : Mean ,Median ,and ,Mode

Principles of statistics

Topic 3

Solving the Exercises

1- $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ = (37+35+…..+36)/10 =36.3

Me :

35 , 35 , 36 , 36 , 36 , 36 , 37 , 37 , 37 , 38

Me = (36+36)/2

36

Mo =36

2-

(i)

X_L =25

X_S =5

C=3

W= (X_L - X_S)/C =(25-5)/3 =7

Class	Tally	fi	Ri	Pi	Upper lim	Cum
5-	𝍸	5	0.139	13.9	< 12	5
12-	𝍸 𝍸 𝍸 𝍸	20	0.556	55.6	< 19	25
19-25	𝍸 𝍸 |	11	0.305	30.5	≤ 25	36
sum		36				

(ii) Ri = (fi /Σ fi) , Pi = Ri * 100

(iii) $\bar{x} = \dfrac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i}$ or $\dfrac{\sum_{i=1}^{n} m_i f_i}{n}$

Class	fi	mi	mifi	Upper limit	Cum
5-	5	8.5	42.5	< 12	5
12-	20	15.5	310	< 19	25
19-25	11	22.5	247.5	≤ 25	36
sum	36		600		

Mean= (600/36)=16.7

Mo= L_i+ { (T-K)/K$'$ -K)} W_i

T = Σ fi /2 =36/2 =18

L_i =12- , W_i =7 , K =5 , K$'$ =25

Mo= 12 + { (18-5)/ (25-5) } (7) =16.55

Me = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =12- , d_1= 20-5=15 , d_2 =20-11=9 , W_i =7

Mo= 12 + { 15/(15+9) } (7) =16.38

3-

Class	fi	mi	mifi	Upper limit	Cum
33-	10	35.5	355	< 38	10
38-	12	40.5	486	< 43	22
43-	51	45.5	2320.5	< 48	73
48-	30	50.5	1515	< 53	103
53-57	8	55.5	444	≤ 57	111
	111		5120.5		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad = (5120.5/111)=46.1$$

$Me = L_i + \{ (T-K)/K' - K) \} W_i$

$T = \Sigma fi /2 = 111/2 = 55.5$

$L_i = 43-$, $W_i = 5$, $K = 22$, $K' = 73$

$Me = 43 + \{ (55.5-22)/ (73-22) \} (5) = 44.8$

$Mo = L_i + \{ d_1/(d_1+d_2) \} W_i$

$L_i = 43-$, $d_1 = 51-12=39$, $d_2 = 51-30=21$, $W_i = 5$

$Mo = 43 + \{ 39/(39+21) \} (5) = 46.3$

4- $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = (45+49+\ldots\ldots+24)/11=29.4$

Me : 10 , 11 , 20 , 22 , 24 , **25** , 33 , 39 , 45 , 45 , 49

Me =25

Mo=45

5-

Class	fi	mi	mifi	Upper limit	Cum
30-	1	35	35	< 40	1
40-	2	45	90	< 50	3
50-	5	55	275	< 60	8
60-	15	65	975	< 70	23
70-	25	75	1875	< 80	48
80-	20	85	1700	< 90	68
90-99	12	95	1140	≤ 99	80
	80		6090		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad = 6090/80 = 76.13$$

$Me = L_i + \{ (T-K)/K' - K) \} W_i$

$T = \Sigma fi / 2 = 80/2 = 40$

$L_i = 70-$, $W_i = 10$, $K = 23$, $K' = 48$

$Me = 70 + \{ (40-23)/ (48-23) \} (10) = 72.4$

$Mo = L_i + \{ d_1/(d_1+d_2) \} W_i$

$L_i = 70-$, $d_1 = 25-15=10$, $d_2 = 25-20=5$, $W_i = 10$

$Mo = 70 + \{ 10/(10+5) \} (10) = 76.7$

6-

Class	fi	mi	mifi	Upper limit	Cum
50-	8	52.5	420	< 55	8
55-	10	57.5	575	< 60	18
60-	16	62.5	1000	< 65	34
65-	14	67.5	945	< 70	48
70-	10	72.5	725	< 75	58
75-	5	77.5	387.5	< 80	63
80-84	2	82.5	165	≤ 84	65
	65		4217.5		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad = 4217.5/65 = 64.9$$

Me= L_i+ { (T-K)/K′ -K)} W_i

T = Σ fi /2 = 65/2 = 32.5

L_i =60- , W_i =5 , K =18 , K′ =34

Me= 60 + { (32.5-18)/ (34-18) } (5) = 64.5

Mo = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =60- , d_1= 16-10=6 , d_2 =16-14=2 , W_i =5

Mo= 60 + { 6/(6+2) } (5) = 63.75

7-

Class	fi	mi	mifi	Upper limit	Cum
50-	8	55	440	< 60	8
60-	10	65	650	< 70	18
70-	16	75	1200	< 80	34
80-	12	85	1020	< 90	46
90-	12	95	1140	< 100	58
100-	4	105	420	< 110	62
110-119	3	115	345	≤ 119	65
	65		5215		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad = 5215/65 = 80.2$$

Me= L_i+ { (T-K)/K' -K)} W_i

T = Σ fi /2 = 65/2 = 32.5

L_i =70- , W_i =10 , K =18 , K' =34

Me= 70 + { (32.5-18)/ (34-18) } (10) = 79.1

Mo = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =70- , d_1= 16-10=6 , d_2 =16-12=4 , W_i =10

Mo= 70 + { 6/(6+4) } (10) = 76

8-

Class mark	fi	mifi	Cum	
5	2	10	< 10	2
15	5	75	< 20	7
25	10	250	<30	17
35	25	875	< 40	42
45	8	360	≤49	50
	50	1570		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad =1570/50= 31.4$$

Me= L_i+ { (T-K)/K' -K)} W_i

T = Σ fi /2 = 50/2 = 25

L_i =30- , W_i =10 , K =187 , K' =42

Me= 30 + { (25-17)/ (42-17) } (10) = 33.2

Mo = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =30- , d_1= 25-10=15 , d_2 =25-8=17 , W_i =10

Mo= 30 + { 15/(15+17) } (10) = 34.7

9-

Class	fi	mi	mifi	Upper limit	Cum
300-	14	350	4900	< 400	14
400-	46	450	20700	< 500	60
500-	58	550	31900	< 600	118
600-	76	650	49400	< 700	194
700-	68	750	51000	< 800	262
800-	62	850	52700	< 900	324
900-	48	950	45600	< 1000	372
1000-	22	1050	23100	< 1100	394
1100-1199	6	1150	6900	≤ 1199	400
	400		286200		

$$\overline{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad =286200/400= 715.5$$

Me= L_i+ { (T-K)/K$'$ -K)} W_i

T = Σ fi /2 = 400/2 = 200

L_i =700- , W_i =100 , K =194 , K$'$ =262

Me= 700 + { (200-194)/ (262-194) } (100) = 708.8

Mo = L_i + { d_1/(d_1+d_2)} W_i

L_i =600- , d_1= 76-58=18 , d_2 =76-68=8 , W_i =100

Mo= 600 + { 18/(18+8) } (100) = 669.2

10-

Class	fi	mi	mifi	Upper limit	Cum
60-	132	70	9240	< 80	132
80-	159	90	14310	< 100	291
100-	95	110	10450	< 120	386
120-	42	130	5460	< 140	428
140-159	22	150	3300	≤ 159	450
sum	450		42760		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} = 42760/450 = 95.02$$

$Me = L_i + \{ (T-K)/K' - K) \} W_i$

$T = \Sigma fi / 2 = 450/2 = 225$

L_i =80- , W_i =20 , K =132 , K' =291

$Me = 80 + \{ (225-132)/ (291-132) \} (20) = 91.7$

$Mo = L_i + \{ d_1/(d_1+d_2) \} W_i$

L_i =80- , d_1= 159-132=27 , d_2 =159-59=64 , W_i =20

Mo= 80 + { 27/(27+64) } (20) = 85.9

11-

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = (20+14+\ldots\ldots+229)/12 = 18.4$$

Me : 10 , 14 , 14 , 14 , 14 , **14 , 18** , 20 , 22 , 24 , 28 , 29 ,

Me = (14+18)/2 =16

Mo =14

12-

Class	fi	mi	mifi	Upper limit	Cum
10-	6	12.5	75	< 15	6
15-	1	17.5	17.5	< 20	7
20-	2	22.5	45	< 25	9
25-	1	27.5	27.5	< 30	10
30-34	2	32.5	65	≤ 34	12
sum	12		230		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} = 230/12 = 19.2$$

Me= L_i+ { (T-K)/K′ -K)} W_i

T = Σ fi /2 = 12/2 = 6

L_i =15- , W_i =5 , K =0 , K′ =7

Me= 15 + { (6-0)/ (7-0) } (5) = 19.3

Mo = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =10- , d_1= 6-0=6 , d_2 =6-1=5 , W_i =5

Mo= 10 + { 6/(6+5) } (5) = 12.7

13-

Class	fi	mi	mifi	Upper limit	Cum
10-	8	15	120	< 20	8
20-	5	25	125	< 30	13
30-39	7	35	245	≤ 39	20
sum	20		490		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad = 490/20 = 24.5$$

$Me = L_i + \{ (T-K)/K' - K) \} W_i$

$T = \Sigma fi / 2 = 20/2 = 10$

L_i =20- , W_i =10 , K =8 , K' =13

$Me = 20 + \{ (10-8)/ (13-8) \} (10) = 24$

$Mo = L_i + \{ d_1/(d_1+d_2) \} W_i$

L_i =10- , d_1= 8-0=8 , d_2 =8-5=3 , W_i =10

Mo= 10 + { 8/(8+3) } (10) = 17.3

14-

(i) XL =109
XS =52
C=7
W= (XL - XS)/C =(109-52)/7 =9

(ii)

Class	Tally	fi	Ri	Pi
52-	|	1	0.044	4.4
61-	||	2	0.087	8.7
70-	|||	3	0.131	13.1
79-	𝍸|||	8	0.348	34.8
88-	|||	3	0.131	13.1
96-	𝍸	5	0.217	21.7
105-113	|	1	0.044	4.4
sum		23		

Ri = fi/Σfi , Pi = Ri * 100

(iii)

Class	fi	mi	mifi	Upper limit	Cum
52-	1	56.5	56.5	< 61	1
61-	2	65.5	131.0	< 70	3
70-	3	74.5	223.5	< 79	6
79-	8	83.5	668.0	< 88	14
88-	3	92.5	277.5	< 96	17
96-	5	101.5	507.5	< 105	22
105-113	1	110.5	110.5	≤ 113	23
sum	23		1974.5		

$$\overline{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} = 1974.5/23 = 85.9$$

Me= L_i+ { (T-K)/K′ -K)} W_i

T = Σ fi /2 = 23/2 = 11.5

L_i =79- , W_i =9 , K =6 , K′ =14

Me= 79 + { (11.5-6)/ (14-6) } (9) = 85.2

Mo = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =79- , d_1= 8-3=5 , d_2 =8-3=5 , W_i =9

Mo= 79 + { 5/(5+5) } (9) = 83.5

15-

To find Mean :

Class	fi	mi	mifi
50-	10	60	600
70-	10	80	800
90-	30	100	3000
110-	20	120	2400
130-149	10	140	1400
sum	80		8200

$$\overline{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad = 8200/80 = 102.5 \text{ (1000 I.D)}$$

To find Median :

Class	fi	Upper limit	Cum
50-	10	<70	10
70-	10	<90	20
90-	30	<110	50
110-	20	<130	70
130-149	10	≤149	80
sum	80		

$Me = L_i + \{ (T-K)/K' - K) \} W_i$

$T = \Sigma fi / 2 = 80/2 = 40$

L_i =90- , W_i =20 , K =20 , K' =50

Me= 90 + { (40-20)/ (50-20) } (20) = 103.3

To find Mode:

$Mo = L_i + \{ d_1/(d_1+d_2) \} W_i$

L_i =90- , d_1= 30-10=20 , d_2 =30-20=10 , W_i =20

Mo= 90 + { 20/(20+10) } (20) = 103.

Cihan University-Sulaimanya Camp

Principles of Statistics

Topic -FOUR

Measure of Dispersion:

Prof.Dr. Obaid Mahmmood Mohsin

2023-2024

1-Range

2-Variance and Standard deviation

3-Coefficient of variation

For ungrouped and grouped data

4-Exercise

Topic 4

Measure of Dispersion:

I – ungrouped Data

1-Range:

Range = largest value – Smallest Value

Example (1):

The following data represent the grades of (10) students in an examination.

57, 51, 61, 48, 84, 30, 44, 60, 69

Find the Range.

2- Variance and Standard deviation:

The variance is the total square of values deviations from their arithmetic mean divided by the sample size minus one.

If we have n of values $x_1, x_2, \ldots, x_n$ and we symbolized s^2 for variance then:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

And the arithmetical formula:

$$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)}$$

The standard deviation S be , $s = \sqrt{s^2}$

Example :2

The following data represent the over wages during a week for (5) of employees in a company.

10, 54, 21, 33, 15, 3

Find the variance and the Standard deviation.

Example :3

Following are the 2005 earnings (in thousands of dollars) before taxes for all six employees of a small company.

48.5, 38.5, 65.5, 22.5, 79.5, 54.5

Calculate the range, variance and standard deviation for these data.

Note:

If we consider the data as a population (not sample)

$$\delta^2 = \frac{N\sum x_i^{\,2} - (\sum x_i)^2}{N^2}$$

3- Coefficient of Variation (C.V):

The coefficient of variation (c.v) puts the standard deviation in the form of proportion from the arithmetic mean or:

$$C.V = \frac{s}{\overline{x}}$$

And the value of the coefficient variation has no unite.

Example :4

Assume we have two samples. We took the following information about them:

	Sample 1	sample 2
The age (year)	25	7
The lengths average (cm)	145	80
The standard deviation (cm)	10	10

Find the (C.V)

II- Grouped Data

1- Range :

The range is the difference between the upper class limit of the last class interval and the lower class limit of the first class interval.

The range = upper class limit of the last interval – lower class limit of the first interval

Example :5

For the following frequency distribution find range:

Classes	Frequency
Zero- 5- 10- 15- 20- 25-29	8 12 20 15 4 1
	60

Example :6

From the following frequency distribution table find the Range.

Class interval	5 15 25 35 45	
Frequency	2 5 10 25 8	50

2- *variance and Standard deviation:*

The variance is computed according to the following formula.

$$s^2 = \frac{\sum (m_i - \bar{x})^2 f_i}{\sum f_i - 1}$$

And the arithmetical formula:

$$s^2 = \frac{n\sum m_i^2 f_i - (\sum m_i f_i)^2}{n(n-1)}$$

Where:

mi = class mark

fi= class interval frequency

n= sum of frequencies or the sample size.

Example :7

From the following frequency table. Find the variance and the standard deviation.

classes	Frequency
10-	5
20-	19
30-	10
40-	13
50-	4
60-	4
70-79	2

Example :8

The following table give the frequency distribution of orders received each day during the post (50) days at the office of a mail-order company.

Find $s^2 and$ s

Class(No of order)	Frequency(No of days)
10- 13- 16- 19-21	4 12 20 14
	50

Cihan University –Sulaimanya Camp

Principles of Statistics
Topic –FOUR-

Measures of Dispersion

Exercises & solving

2023-2024

Prof.Dr. Obaid Mahmmood Mohsin

Principles of Statistics

Topic –4-

Exercises

1-For each of the following finds the range , variance ,and standard deviation , mean deviation and coefficient variation :

(i) 12 , 6 , 7 , 3 , 15 , 10 , 18 , 5
(ii) 10 , 8 , 2 , 4 , 14 , 4 , 10 , 12 , 8
(iii) -4 , 2 , -6 , 0 , -4 , 6 , 4 , 0
(iv) 2 , 3 , 3 , 3 , 5 , 6 , 2 , 2 , 8 , 10 , 12 , 20

2-The following data represent quantity incentive during a month of a set of employees in a company (1000 I.D):

82 , 87 , 97 , 65 , 70 , 67 , 88 , 100 , 107 , 85 , 93 , 50 , 90 , 80 , 95 , 80 , 70 , 85 , 72 , 109 , 95 , 90 , 70

Required:

(i) Frequency distribution table of (6) class interval.
(ii) Mean ,Median ,and ,Mode.
(iii) Range , variance ,standard deviation , mean deviation and coefficient variation .

3- A student got (70) in the computer skills , (80) in the mathematical . If you informed that the mean of whole grades is (60) and standard deviation is (2) in computer skills and the mean of mathematical grades is (75) and standard deviation is (2.5).In which of the two subjects the student's level is higher ?

4- Transform the following values into standard score .

6 , 4 , 3 , 8 , 7 , 5
7

5-The following frequency table represents the weekly over wages (1000 I.D) of employees of a factory :

Class (weekly wages (100 I.D))	Fi (employees number)
16-	15
22-	16
28-	25
34-	20
40-45	10

Find :

(i) Variance and standard deviation .
(ii) Mean deviation and semi-in quartile range.

6- Find the variance ,standard deviation ,mean deviation ,the range and semi-interquartile range of the following frequency table which represents the grades set of student.

Class (grades)	Fi (number of student)
30-	1
40-	3
50-	11
60-	21
70-	43
80-	32
90-99	9

7- For the following data find :
Mean , Median ,Mode ,Range ,Semi-interquartile range ,Variance ,standard deviation , Mean deviation , Coefficient variance .

(i) { 1 , 0 . 1 , 1 ,0 . 1 , 1 , 0 . 1 , 1 , 0 , 1 , 1 , }
(ii) { 2 , 1 , 0 , -4 , 5 , 6 , }

Principles of statistics

Topic-4-

Solving the Exercises

1-

(i) $R = X_L - X_S = 18-3 = 15$

(ii) $R = 14-2 = 12$

(iii) $R = 6-(-6) = 12$

(iv) $R = 20-2 = 18$

(I)

X_i	X_i^2
12	144
6	36
7	49
3	9
15	225
10	100
18	324
5	25
76	**912**

$$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)} = \{ (8)(912)-(76)^2 \} / \{ (8)(7)\} = (7296-5776)/56 = 1520/56$$

$S^2 = 27.14$, $S = \sqrt{S^2} = 5.21$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 76/8 = 9.5$$

C.V .= (S/ X_{bar}) $(100) = (5.21 /9.5) (100)$

C.V.= 54.84

(i)

X_i	X_i^2
2	144
-4	36
0	49
-6	9
6	225
-4	100
2	324
4	25
0	0
0	**128**

$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)}$ = { (9)(128)-$(0)^2$ } / { (9) (8)} = 16

S=√ S^2 = 4

$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ = zero

C.V .= (S/ X_{bar}) (100) = ignore

(iii)

X_i	X_i^2
2	4
-2	4
0	0
-6	36
6	36
-4	16
2	4
4	16
0	0
2	**116**

$$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)}$$ = { (9)(116)-(2)^2 } / { (9) (8)}

$S^2 = 14.4$

$S = \sqrt{S^2} = 3.8$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$ = 2/9 = 0.22

C.V .= (S/ X_{bar}) (100)

C.V.= 1727.3

(iv)

X_i	X_i^2
2	4
3	9
3	9
5	25
6	36
2	4
2	4
8	64
10	100
12	144
20	400
63	**799**

$$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)}$$ = { (11)(799)-(63)^2 } / { (11) (10)} = 4820/110 =43.8

$S^2 = 43.8$

$S=\sqrt{S^2} = 6.62$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 63/11 = 5.73$$

C.V .= (S/ X_{bar}) (100)

C.V.= 115.5

2-

(i) Largest value =109 , Smallest value =50

W=(L-S)/C =(109-50)/6 = 10

(ii)

Class	Tally	fi	mi	mifi	Upper limit	Cum
50-	|	1	55	55	< 60	1
60-	||	2	65	130	< 70	3
70-	||||	4	75	300	< 80	7
80-	𝍸||	7	85	595	< 90	14
90-	𝍸|	6	95	570	< 100	20
100-109	|||	3	105	315	≤ 109	23
		23		1965		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} = 1965/23 = 85.9$$

$Me= L_i + \{ (T-K)/K' -K)\} W_i$

$T = \Sigma fi /2 = 23/2 = 11.5$

L_i =80- , W_i =10 , K =7 , K' =14

Me= 80 + { (11.5-7)/ (14-7) } (10) = 86.4

Mo = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =80- , d_1= 7-4=3 , d_2 =7-6=1 , W_i =10

Mo= 80 + { 3/(3+1) } (10) = 87.5

(iii)

Range = 109-50 =59

Class	fi	mi	Mi^2	mifi	Mi^2fi
50-	1	55	3025	55	3025
60-	2	65	4225	130	8450
70-	4	75	5625	300	22500
80-	7	85	7225	595	50575
90-	6	95	9025	570	54150
100-109	3	105	11025	315	33105
	23			1965	171675

$$s^2 = \frac{n\sum m_i^2 f_i - (\sum m_i f_i)^2}{n(n-1)} = \{ (23) (171675) - (1965)^2 \} / \{ (23) (22) \}$$

=(3948525-3861225)/506 =172.53

S = √ S^2 =13.14

C.V. = S/Xbar (100) = 13.14 /85.4 (100)= 15.4

3- Let $X_1$the grades of computer skills

X_2 the grades of mathematical

Z_1= ($X_1 - X_{1bar}$)/S_1 = (70-60)/2 =5

Z_2= ($X_2 - X_{2bar}$)/S_2 = (80-75)/2.5 =2

The student's level in computer skills is higher .

4-

X_i	X_i^2
6	36
4	16
3	9
8	64
7	49
5	25
33	199

$$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)} = \{ (6)(199)-(33)^2 \} / \{ (6)(5) \}$$

$S^2 = 3.5$, $S=\sqrt{S^2} = 1.9$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 33/6 = 5.5$$

Xi	X_{bar}	S	$Z_i = (X_i - X_{bar})/S$
6	5.5	1.9	0.26
4			-0.79
3			-1.32
8			1.32
7			0.79
5			-0.26

5-

Class	fi	mi	mifi	mi^2	Mi^2fi
16-	15	19	285	361	5415
22-	16	25	400	625	10000
28-	25	31	775	961	24025
34-	20	37	740	1369	27380
40-45	10	43	430	1849	18490
	86		2630		85310

(i) R= 45-16 =29

$$s^2 = \frac{n\sum m_i^2 f_i - (\sum m_i f_i)^2}{n(n-1)}$$ = { (86) (85310) –$(2630)^2$ } / { (86) (85) }

=(7336660-6916900)/7310 =57.4

S = √ S^2 =7.6

$$\overline{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i}$$ or $$\frac{\sum_{i=1}^{n} m_i f_i}{n}$$ =2630/86= 30.6

C.V. = S/Xbar (100) = (7.6 /30.6) (100)= 24.84

6-

Class	fi	mi	mifi	mi2	mi2fi	Upper li.	Cum
30-	1	35	35	1225	1225	<40	1
40-	3	45	135	2025	6075	<50	4
50-	11	55	605	3025	33275	<60	15
60-	21	65	1365	4225	88725	<70	36
70-	43	75	3225	5625	241875	<80	79
80-	32	85	2720	7225	231200	<90	111
90-99	9	95	855	9025	81225	≤99	120
	120		8940		683600		

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i f_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \frac{\sum_{i=1}^{n} m_i f_i}{n} \quad = 8940/120 = 74.5$$

Me= L_i+ { (T-K)/K′ -K)} W_i

T = Σ fi /2 = 120/2 = 60

L_i =70- , W_i =10 , K =36 , K′ =79

Me= 70 + { (60-36)/ (79-36) } (10) = 75.6

Mo = L_i + { $d_1/(d_1+d_2)$} W_i

L_i =70- , d_1= 43-21=22 , d_2 =43-32=11 , W_i =10

Mo= 70 + { 22/(22+11) } (10) = 76.7

R=99-30=69

$$s^2 = \frac{n\sum m_i^2 f_i - (\sum m_i f_i)^2}{n(n-1)} = \{ (120) (683600) - (8940)^2 \} / \{ (120) (119) \}$$

=147.7

S = √ S^2 =12.2

C.V. = S/Xbar (100) = (12.2 /74.5) (100)= 16.4

7-

(i) ΣX=9 , Σ X^2 =9

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 9/13 = 0.7$$

Me : 0 , 0, 0 , 0 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 ,

Me = 1

Mo =1

R=1-0 =1

$$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)} = \{ (13)(9)-(9)^2 \} / \{ (13)\ (12)\}$$

$S^2 = 0.23$, $S=\sqrt{S^2} = 0.5$

C.V.= $(S/X_{bar})\ (100)$ =(0.5 /0.7 (100) = 71.4

(ii) $\Sigma X=10$, $\Sigma X^2 =82$,n=6

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad =10/6 =1.7$$

Me : -4 , 0, 1 , 2 , 5 , 6

Me = (1 +2)/2 =1.5

Mo =There is no mode

R=6-(-4) =10

$$s^2 = \frac{n\sum x_i^{\,2} - (\sum x_i)^2}{n(n-1)}$$

= { (6)(82)-$(10)^2$ } / { (6) (5)}

$S^2 = 13.1$, $S=\sqrt{S^2} = 3.62$

C.V.= (S/Xbar) (100) =(3.62 /1.7 (100) = 212.9