Victoria University of Bangladesh

Final Assessment

Summer Semester 2023

Name : MD Sujan ali

ID : 1119460091

Program : BBA

Batch : 46th

COURSE CODE : STA 220

COURSE TITLE : Business Statistics

Submitted to : TANIYA NASHIN

Course title: Business Statistics
Course code: STA 220

Ans: to the Q.N. ①

① Define variable:-

A variable is a symbolic name or identifier that represents a value or data stored in a computer program. It is a fundamental concept in programming and can hold different types of data, such as numbers, text, or complex objects. Variables allow programmers to manipulate and work with data in their code. Variable is a characteristic that can be measured and that can assume different values

P.t.o

Height, age, income Province or country of birth, grade obtained at school and type of housing are all examples of variables. variables may be classified into two main

Different types of variable:

There are several types of variables including:

* Numeric variables:
* Integer: Holds whole numbers without decimal points.
* Floating-point: Stores numbers with decimal points.
* Double: Similar to float but with higher precision,

p.t.o

but

* Long : Stores large whole numbers

. character variables =>

* char : stores single character like letters or symbols.

* Boolean variables →

* Boolean : Represents true or false values.

String variables →

* String : Holds sequences of characters such as words or sentences.

Array variables →

* Array : Stores multiple values of the same type in a single variable

p.t.o

Pointer variables →

* Pointer: Stores memory addresses, often used in low-level programming.

Constant variables →

* constant: Represents values that cannot be changed once defined

Enum variables →

* Enum: Represents a set of named integer constants.

object variables →

* object: Used in object-oriented Programming to store instances of user-defined classes.

p.t.o

Variant variables →

* Variant :- A dynamic data type that can hold values of various types (common in some scripting languages :).

In Programming, variables can be categorized into different types based on the kind of data they can hold. Here are some common types of variables and example :-

(Int) :- This type stores whole numbers without decimal point. For example, int age = 25; assigns the value 25 to the variable age

p.t.o

(float) :- The variables can store numbers with decimal points. For instance, float temperature = 98.6; stores the temperature value with a decimal point.

(double) :- Similar double variables can store floating-point numbers with greater accuracy. For example, double pi = 3.14159265359; stores the value of pi.

(char) :- char variables hold single characters, such as letter or symbol. For instance, char grade = 'A'; assigns the character 'A' to

p.t.o

The variable grade.

(String):- String variables are used store sequences of characters, like words or sentences. For example, string name = "Jonh"; store the name "Jonh" in the variables name.

(boolean):- two values: true or false. They are often used for logical conditions. For instance, boolean is Raining = true; indicates wheth it's raining.

P.t.o

(Arrays): An arrays is a collection of values of the same data type. You can think of it as a list of variables. For example int numbers = {1,2,3,4,5}; creates an integer arrays containing five values.

Objects: variables of object types can store complex data structures with attributes and methods. For instance, Person person = new Person(); creates an object variable person of the class Person.

Programming, objects are instances of classes.

P.t.o

(Enum) :- Enum are user-defined data types that consist of a fixed that consist of a fixed set constant values. They are often used to represent a set of related values. For example, you could define an enum for days of the week.

Pointer (in languages like c/c++):
Pointers store memory addresses, allowing you to work with data indirectly. They are a more advanced concept in some languages and require careful handling.

p.t.o

these are some of the fundamental variable types used in Programming. and their usage may vary depending on the Programming language you are working with. Choosing the right variable type is important for efficient memory usage and accurate representation of data in your Programs. ordinal variable can be logically ordered or ranked higher or lower than another but do not necessarily establish a numeric difference between each such

P.t.o

Ans° to the Q.N (2)

(a) That is why we will study statistics

Statistics is the study collecting, analyzing, interpreting Presenting and organizating data it serves several important Purposes:

Data interpretation: Statistics helps us make sense of large amounts of data by summarizing and presenting it in a meaningful way. This is crucial for decision-making in various fields. There some kind fields them out raise the statistics crucial for the

P.t.o

Inference: It allows us to draw conclusions and make prediction based on data. For example, in medical research, statistics help determine if a new drug is effective.

Risk Assessment: Statistics is used to assess and manage risks. In insurance, for instance, it helps determine premium rates based on risk factors.

Scientific Research: In scientific research, statistics is essential for designing experiments, analyzing results, and drawing valid conclusions.

P.t.o

Business and Economics: Statistics is crucial in market research, financial analysis, and forecasting, aiding businesses in making informed decisions.

Public Policy: Governments use statistics to inform policies, allocate resources, and address societal issues like healthcare, education, and infrastructure.

Quality control: In manufacturing and quality control, statistics is used to monitor and improve processes to ensure products meets standards.

p.t.o

Social Sciences: In fields like Sociology and Psychology, statistics helps researchers analyze human behavior and social phenomena.

The study of statistics fundamental about making informed decisions drawing conclusions from data, growing and advancing knowledge in various disciplines. It is a valuable tool in today's data-driven world, in work whose country of different kind

P.T.O

also there are five major reasons to study statistics:

The first reason is to be able to effectively conduct research. Without the use of statistics it would be very difficult to make decisions based on the data collected from a research project. For example, in the study cited in chapter one, in the difference in recorded absenteeism between psychiatric and obstetrics nurse large enough to conclude that there is meaningful difference in absenteeism between the two units. There are two possibilities, the first possibility is that the diff-

P.t.o

exercise the between the two groups is a result of chance factors. In reality, the two jobs have approximately the same amount of absenteeism.

A second point about research should be made. It is extremely important for a research to know what statistics they want to use before they collect their data. otherwise data might be collected that is uninterpretable. Unfortunately, when this happens it results in a loss of data, time and money. How many a student may be ashamed to themselves but I never plan on doing my research.

P.T.O

The second reason to study statistics is to be able to read journals. Most technical journals you will read contain some form of statistics. Usually, you will find them in something called the results section. Without an understanding of statistics, the information contained in this section will be meaningless. An understanding of basic statistics will provide you will the fundamental skill necessary to read and evaluate most result section. The ability to extract meaning from journal articles and the ability to critically evaluate research from a statistics perspective are fundamental skill that enhance

p.t.o

The third reason is to further develop critical and analytic thinking skills. most student completing high school and introductory undergraduate coursework have at their disposal a variety of critical thinking and analytic skills. the study of statistics will serve to enhance and further develop these skills. To do well in statistics one must develop and use formal logical thinking abilities that are both high level and creative. The more these statistics can reason the different them often there.

p.t.o

The fourth reason to study statistics to be an informed consumer like any other tool, statistics can be used or misused. Yes, it is true that some individuals do actively lie and mislead with statistics. More often, however, well meaning individuals unintentionally report erroneous statistics conclusions. If you know some of the basic statistical concepts, you will be in a better position to evaluate the information you have been given. some can used them statistics

P.t.o

The 5th reason to have a working knowledge of statistics is to know when you need to have a statistics. Most of us know enough about our cars to know when to take it into the shop. Usually we don't attempt the repair ourselves because we don't want to cause any irreparable damage. Also we try to know enough to be able to carry on an intelligible conversation with the mechanic to insure that we don't get a whole new engine job should be the same way about hiring a statistician.

Ans:- to the Q.N. (3)

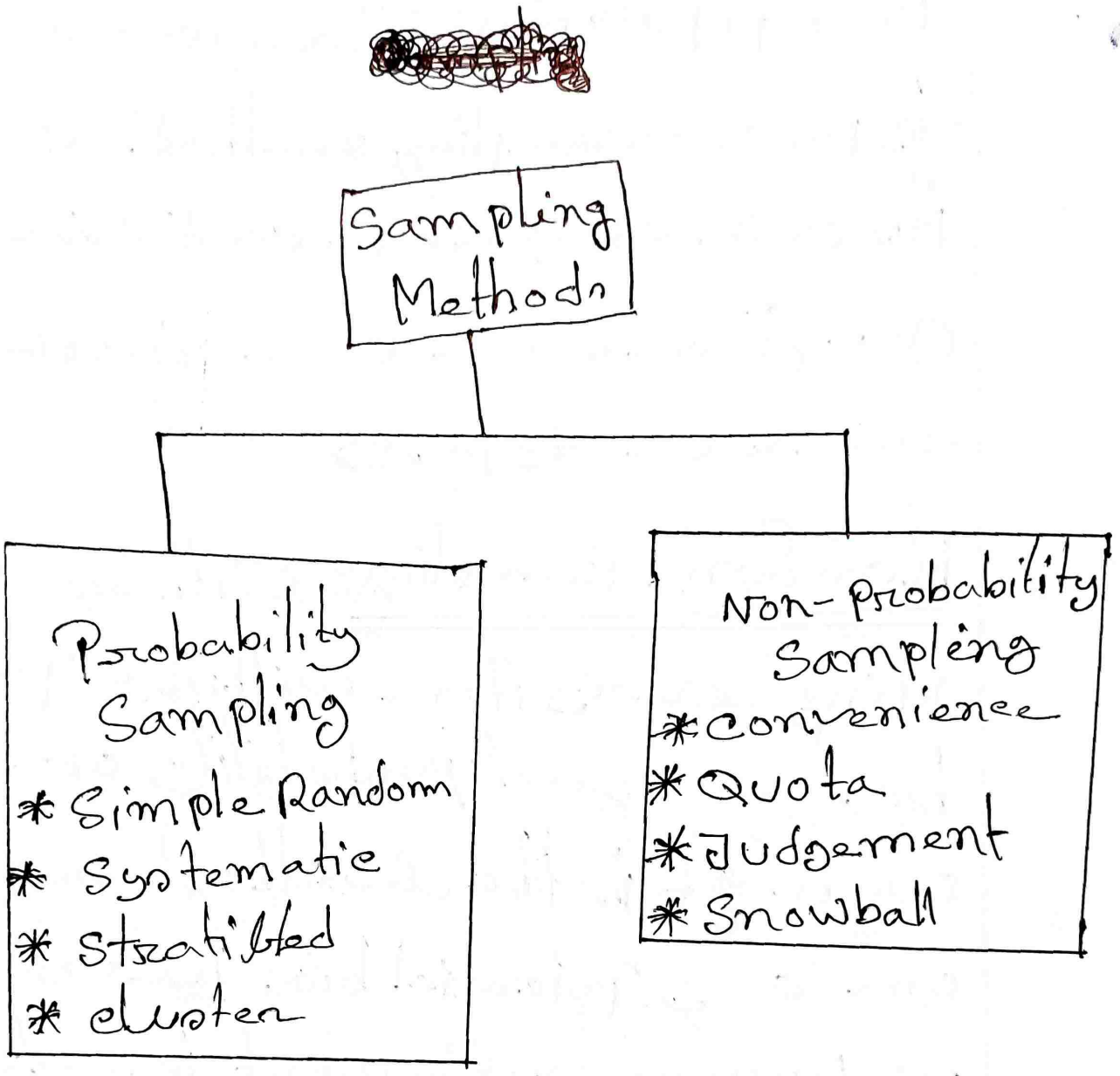
(3) **Sample:-** A sample is a subset of a larger population or group that is selected for research, analysis or testing. It is used to make inferences or draw conclusions about the entire population based on the characteristics or data collected from the sample. Samples are commonly used in fields like statistics, research and market analysis to gather insights and make predictions without having to study the entire population which may be impractical or impossible to do think there are statistics.

Sampling Works

It can be difficult for research to conduct accurate studies on large population. In some case, it can be impossible to study every individual in the group. That is why they often chose a small portion to represent the entire group. This is called a sample. Samples allow researchers to use characteristics of the small group. This is called a sample. Samples allow researchers to use characteristics of the small group to make estimates of the large population. The chosen sample should be a fair representation of the en-

p.t.o

when taking a sample from a large population, it is important to consider how the sample is chosen to get a representative sample, it must be drawn randomly and encompass the whole population. For example, a lottery system could be used to determine the average age of student in university by sample 10% of the student body. Sampling is commonly used when studying large portions of the population for economic purpose. The monthly employment report in olves the use of sample. p.t.o



Discuss different types of sample

Sampling methods are essential in research and statistics for gathering data from a population when it's not feasible to collect data from the ent-

p.t.o

ize population. There are several types of sampling methods, each with its own advantage and disadvantages. Here's an overview of some common types:→

Random Sampling:- With random sampling, every item within a population has an equal probability of being chosen. It is the furthest removed from any potential bias because there is no human judgement involved in selecting the sample.

For example, a random sample may include choosing the names of 25 employees out of a batina company of 250 employees.

p.t.o

Stratified sampling:- The population is divided into subgroups or strata, and random samples are taken from each each stratum. This method is useful when you want to ensure representation from various subgroups within the population.

Systematic sampling:- Researchers select every nth member from a list of the population. For example, if you have a list of 1,000 individuals and want a sample of 1000, you might choose every 10th person on the list. The sampling interval is calculated as the population size divided by the sample size. p.t.:

Despite the sample population being selected in advance, systematic sampling is still considered random if the periodic interval is determined beforehand and the starting point is random. Assume that an auditor reviews the internal controls related to a company's cash account and wants to test the company policy that stipulates that checks exceeding $10,000 must be signed by two people. The population consists of every company check exceeding $10,000 during the fiscal year, which, in this example was 300.

P.T.O.

Cluster Sampling: The population is divided into clusters is selected. Then, all members within the chosen clusters are included in the sample. It's useful when it's difficult to create a comprehensive list of the population. cluster sampling is a form of probability sampling. When researchers conduct cluster sampling they divide the population into smaller groups. They then select individuals randomly from these groups to form their samples and conducts their studies. This kind of sampling is used when both the overall population and sampling size

p.t.o

Convenience Sampling:

This method involves selecting participants who are easy to reach or readily available. It is quick and convenient but may introduce bias, as it may not represent the entire population.

Purposive Sampling:

Researcher select specific individuals or groups intentionally based on certain characteristics or criteria. It is often used in qualitative research when the focus is on specific traits or experiences.

P.T.O

Quota Sampling: Researchers divide the population into subgroups and then select participants non randomly to meet predefined quotas for each subgroups. It is similar to stratified sampling but not truly random.

Snowball Sampling: This method is commonly used in situations where the population is hard to reach like hidden or marginalized groups. One participant refers the researcher to another, creating a snowball effect.

p.t.o

Judgmental Sampling:

Researchers rely on their judgment or expertise to select participants who they believe are the most relevant or representative of the popul.

Opportunistic Sampling: Also know as accidental sampling this involves sampling individuals who happen to be available at the time of data collection. It's not a systematic approach and can introduce bias.

The choice of sampling method depends on various factors, including the research objectives available resources, and the nature of the population. Researchers must carefully consider these factors to ensure the sample accurately

Ans° to the Q.N ④

④ Discuss Mean and standard Deviation

Mean and standard deviation are two important statistical measures used to describe a set of data :→

Mean (Average)

The mean, often referred to as the average, is calculated by adding up all the values in a data set and then dividing by the number of values.

Formula: Mean (M) = $(\Sigma X) / N$, Where ΣX represents the sum of all data values, and N is the total number

P.t.o

values.

The mean provides a measure of central tendency, giving you a single value that represents the "typical" value in the dataset

standard Deviation.

The standard deviation measures the amount of variation or dispersion in a data set.

Formula: standard Deviation (σ)

$= \sqrt{[\Sigma (X-\mu)^2 / N]}$, where X is each individual data point, μ is the mean Σ denotes summation, and N is the number of data points.

P.t.o

A higher standard deviation indicates that data point are more spread out from the mean, implying greater variability. Conversely, a low standard deviation suggests that data points are closer to the mean indicating less variability. In summary, the mean provides a measure of central tendency, while the standard deviation quantifies the spread or dispersion of data points around the mean. Together these two statistics offer valuable insights into the characteristics

P.t.o

of a data set and are frequently used in various fields, including statistics, finance, and science to summarize and analyze data.

Standard Deviation from ungrouped data :⇒ The standard deviation is a summary measure of the differences of each observation from the mean. If the differences of each themselves were added up, the positive would exactly balance the negative and so their sum would be zero. Consequently the squares of the differences are added. p.t.o

The sum of the squares of them divided by the number of observations minus one to give the mean of the squares, and the squares root is taken to bring the measurements back to the units we started with. The division by the number of observations minus one instead of the number of observations if we're to obtain the mean square is because "degrees of freedom" must be used. In these circumstances there are one less than the total. The theoretical justification for this

P.t.o

need not trouble the use in practice to gain an intuitive feel for degrees of freedom, consider choosing a chocolate from a box of n chocolates. Every time we come to choose a chocolate we have a choice until we come to the last one and then we have no choice. Thus we have n-1 choices, or "degrees of freedom."

The calculation of the variance is illustrate in table 2.1 with 15 readings in the preliminary study of variance instead of concentrations.

p.t.o

table 2.1 calculation of standard deviation

(1) lead concentration µmol/24 hr	(2) Differences from mean $X-\bar{X}$	(3) Differences squared $(X-\bar{X})^2$	(4) Observations in col (1) squared x^2
0.1	−1.4	1.96	0.01
0.4	−1.1	1.21	0.16
0.6	−0.9	0.81	0.36
0.8	−0.7	0.49	0.64
1.1	−0.4	0.16	1.21
1.2	−0.3	0.09	1.44
1.3	−0.2	0.04	1.69
1.5	0	0	2.25
1.7	0.2	0.04	2.89
1.9	0.4	0.16	3.61
1.9	0.4	0.16	3.61
2.0	0.5	0.25	4.00
2.2	0.7	0.49	4.84
2.6	1.1	1.21	6.76
3.2	1.7	2.89	10.24
total 22.5	0	9.96	43.71

The sum of the squares of the differences (or deviations) from the mean 9.96, is now divided by the total number of observations minus one to give the variance thus.

$$\text{variance} = \frac{\Sigma (x - \bar{x})^2}{n - 1}$$

in this case we find

$$\text{varian} = \frac{9.96}{14} = 0.7114 \ (\text{mmol/24h})^2$$

Finally, the square root of the variance provide the standard deviation;

P.t.o

$$SD = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

from which we get

$$\sqrt{0.7114} = 0.843 \text{ mmol/(24)}$$

This procedure illustrates the structure of the standard deviation in particular that the two extreme values 0.1 and 3.2 contribute most to the sum of the differences squared.

Calculator Procedure:

Most inexpensive calculators have procedures that enable one to calculate the mean and standard deviations directly. Using the "SD" p.t.o

mode. For example on modern casio calculators one presses SHIFT and a little "SD" symbol should appear on the display. On earlier casios one presses INV and MODE whereas on a sharp 2nd F and stat should be used, the data are stored via the M+ button. Thus, burning set the calculators into the "SD" or "set" mode, from table 2.1 we enter 0.1 M+ 1.0 a M+ etc. When all the data are entered, we can check that the correct number of observations have been fed by.