ACHIEVER'S EDUCATION CONSULTANT
08141710097
SUMMARY OF STT205 WITH PAST QUESTIONS

INTRODUCTION

It will be available to all first degree of the school of Business and Human resources Management at the National Open University, Nigeria. It will also be useful for those seeking introductory knowledge in STATISTICS FOR MANAGEMENT SCIENCES . The course consists of eighteen units that involved basic concepts and principles of statistics and decision making process, forms of data, methods of data collection, summarizing data, graphical presentation of data, measures of both central tendency and dispersion, set theory, permutations and combinations, some elements of probability concepts, probability distributions of both discrete and continuous random variables.

There are no compulsory prerequisites for this course, although a credit pass level in

This Course Guide tells you what the course is about, what course materials you will be using and how you can work your way through these materials. It suggests some general guidelines for the amount of time you are likely to spend on each unit of the course in order to complete it successfully. There are likely going to be regular tutorial classes that are linked to the course. It is advised that you should attend these sessions.

Details of the time and locations of tutorials will be communicated to you by National Open University of Nigeria. During the course, you will be exposed to nature of statistical information, collection and processing of statistical data. You will also be exposed to elementary concepts in probability and nature, characteristics and uses of some important probability distributions of both discrete random variable and continuous random variable. The course aims to give you an understanding of statistical information and presentation for decision-making.

[...]

You may want to refer to them during your study of the unit to check on your progress. In this way you can be sure you have done what was required of you by the unit. By meeting these objectives, you should have achieved the aims of the course.

[...]

Explain the principles underlying the application of various probability Compute measures for Probability Distributions. Each unit contains self-assessment exercises called Student Assessment Exercises, SAE. At some points in the course, you are required to submit assignments for assessment purposes. This course should take about 15 weeks to complete.

Major components of the course are

The last five units, module 4, teach the principles underlying the applications of some important probability distributions. Each unit consists of on week direction for study, reading material, other resources and summaries of key issues and ideas. 4t' ed., Prentice Hall of India Privatize Limited, New Delhi. In this tile, you will find the details of the work you must submit to your tutor for marking.

The marks you obtain for these assignments will count toward the final mark you obtain for this course. Further information on assignments will be found in the Assignment File itself and later in this Course Guide in the section on Assessment.

Assignment 1 - All question in Units 1 - 4

The presentation schedule included in your course materials gives you the important dates for this year for the completion of tutor-marking assignments and attending tutorials. Remember, you are required to submit all your assignments by due date. You should guide against falling behind in your work. In tackling the assignments, you are expected to apply information, knowledge and techniques gathered during the course.

The assignments must be submitted to your tutor for formal Assessment in accordance with the deadlines stated in the Presentation Schedule and the Assignments File. The work you submit to your tutor for assessment will count for 50 % of your total course mark. At the end of the course, you will need to sit for a final written examination of three hours' duration.

Tutor-Marked Assignments

You will submit all the assignments. You are encouraged to work all the questions thoroughly. When you have completed each assignment, send it, together with a TMA form, to your tutor. Make sure that each assignment reaches your tutor on or before the deadline given in the Presentation File.

If for any reason, you cannot complete your work on time, contact your tutor before the assignment is due to discuss the possibility of an extension. Extensions will not be granted after the due date unless there are exceptional circumstances. Final Examination and Grading The final examination will be of three hours' duration and have a value of 50% of the total course grade. The examination will consist of questions which reflect the types of selftesting, practice exercises and tutor-marked problems you have previously encountered.

You might find it useful to review your self-tests, tutor-marked assignments and comments on them before the examination.

COURSE OVERVIEW

This table brings the units, the number of weeks you should take to complete them and the assignment as follows.

15wks Assignments 4

Think of it as reading the lecture instead of listening to a lecturer. Just as a lecturer might give you an in-class exercise, your study units provides exercises for you to do at appropriate

points. The first item is an introduction to the subject matter of the unit and how a particular unit is integrated with the other units and the course as a whole. Next is a set of learning objectives.

These objectives let you know what you should be able to do by the time you have completed the unit. You should use these objectives to guide your study. When you have finished the unit you must go back and check whether you have achieved the objectives. If you make a habit of doing this you will significantly improve your chances of passing the course.

The main body of the unit guides you through the required reading from other sources. First, it will enhance your understanding of the material in the unit. Second, it will give you practical experience of using programs, which you could well encounter in your work outside your studies. In any event, most of the techniques you will study are applicable on computers in normal working practice, so it is important that you encounter them during your studies.

Self-tests are interspersed throughout the units, and answers are given at the ends of the units. Working through these tests will help you to achieve the objectives of the unit and prepare you for the assignments and the examination. You should do each self-test as you come to it in the study unit. The following is a practical strategy for working through the course.

If you run into any trouble, contact your tutor. Remember that your tutor's job is to help you. When you need help, don't hesitate to call and ask your tutor to provide it. Note the time you are expected to spend on each unit and how the assignments relate to the units.

Important information, e.g. details of your tutorials, and the date of the first day of the semester is available from study centre. You need to gather together all this information in one place, such as your dairy or a wall calendar. Whatever method you choose to use, you should decide on and write in your own dates for working breach unit. Once you have created your own study schedule, do everything you can to stick to it.

If you get into difficulties with your schedule, please let your tutor know before it is too late for help. The content of the unit itself has been arranged to provide a sequence for you to follow. As you work through the unit you will be instructed to read sections from your set books or other articles. Use the unit to guide your reading.

Up-to-date course information will be continuously delivered to you at the study centre. Well before the relevant due date , get the Assignment File for the next required assignment. Keep in mind that you will learn a lot by doing the assignments carefully. They have been designed to help you meet the objectives of the course and, therefore, will help you pass the exam.

When you are confident that you have achieved a unit's objectives, you can then start on the next unit. Proceed unit by unit through the course and try to pace your study so that you keep yourself on schedule. When you have submitted an assignment to your tutor for marking do not wait for it return `before starting on the next units. When the assignment is returned, pay particular attention to your tutor's comments, both on the tutor-marked assignment form and also written on the assignment.

Check that you have achieved the unit objectives and the course objectives . There are some hours of tutorials provided in support of this course. You will be notified of the dates, times and location of these tutorials. Together with the name and phone number of your tutor, as soon as you are allocated a tutorial group.

Your tutor will mark and comment on your assignments, keep a close watch on your progress and on any difficulties you might encounter, and provide assistance to you during the course. You must mail your tutor-marked assignments to your tutor well before the due date . They will be marked by your tutor and returned to you as soon as possible. The following might be circumstances in which you would find help necessary.

You should try your best to attend the tutorials. You can raise any problem encountered in the course of your study. You will learn a lot from participating in discussions actively.

[...]

We wish you success with the course and hope that you will find it both interesting and useful.

Basic Concepts in Statistics Assignment Question

You will realize that the activities of man and those of the various organizations, that will often be referred to as firms, continue to increase. This brings an increase in the need for man and the firms to make decisions on all these activities. The need for the quality and the quantity of the information required to make the decisions increases also. The management of any firm requires scientific methods to collect and analyze the mass of information it collects to make decisions on a number of issues.

Such issues include the sales over a period of time, the production cost and the expected net profit. In this regard, statistics plays an important role as a management tool for making decisions.

Describe the uses of statistics

Statistics can be defined as a management tool for making decision. It is also a scientific approach to presentation of numerical information in such a way that one will have a maximum understanding of the reality represented by such information. Statistics is also defined as the presentation of facts in numerical forms. Nonnumerical data cannot be subjected to statistical analysis except they are transformed to numerical data.

The Roles of Statistics You will realize that statistics is useful in all spheres of human life. A woman with a given amount of money, going to the market to purchase foodstuff for the family, takes decision on the types of food items to purchase, the quantity and the quality of the items to maximize the satisfaction she will derive from the purchase. For all these decisions, the woman makes use of statistics Government uses statistics as a tool for collecting data on economic aggregates such as national income, savings, consumption and gross national product. Government also uses statistics to measure the effects of external factors on its policies and to assess the trends in the economy so that it can plan future policies.

Business makes use of statistics in production, marketing, administration and in personnel management. Statistics is also used extensively to control and analyze stock level such as minimum, maximum and reorder levels. Management uses statistics to establish mathematical relationship between two or more variables for the purpose of predicting a variable in terms of others.

Basic Concepts In Statistics

Let us quickly define some of the basic concepts you will continue to come across in this course. In studying the nutritional well being of pupils in a primary school, the entity is a pupil in the school.

Examples of this is the hourly patronage of a restaurant

A discrete variable has "interruptions" between the values it can assume. These are called interruptions. There is always a continuum of values that the continuous variable can assume. The weight can be both whole values or decimal values such as 20 kilograms and 52 kilograms.

In studying the income distribution of students in the National Open University, the incomes of 1000 students selected for the study, from the population of all the students in the Open University will constitute the sample of the study. In this unit you have learned a number of important issues that relate to the meaning and roles of statistics. The various definitions and examples of concepts given in this unit will assist tremendously in the studying of the units to follow. What you have learned in this unit concerns the meaning and roles of statistics, and the various concepts that are important to the study of statistics.

Some surveys can be limited in the sense that they can be carried out with a few minutes of observation. Others can be detailed. When surveys are detailed, information from the surveys are more acceptable and valued than when they are limited.

Government Survey

Market research surveys - carried out for one particular client and not published in any form.

Research surveys - carried out by academicians and published in journals

Firms commission ad hoc surveys on a wide variety of subjects.

Survey methods consist of the following stages

Secondary statistics - We will need to calculate secondary statistics such as means and percentages. Observations This is one of the methods of collecting primary data. It can be used to know the use a particular facility is put. It can be used to study the behaviours of people in a work place.

Systematic observation

Mechanical observation In participant observation, the observer is involved in the activities he is trying to observe. This method can have serious influence in the entities the observer is observing. In systematic observation, the observer does not take part in the activity. Though the method is objective, it does not question the motives of people observed.

Sophisticated mechanical means such as television, film and tape recorders are used to provide more complex information. Selectivity- an observer can unintentionally become selective in perception, recording and reporting. Interpretation- the observer may impute meanings to the behaviour of people that the people do not intend. Chance- a chance event may be mistaken for a recurrent one.

Participation- the participation of the observer can influence the behaviour of people being observed.

Reporting especially in the radio and television

He must not express his own opinions and must not influence the answers of the respondents. The language of the question must be at the level of the respondents. If the questions are written in a language different from that of the respondents, the question must be translated to that of the respondent but the answers must be written by the interviewer in the language of the questions. This was the case during the last census in Nigeria.

The questionnaire used was written in English language. That is not the language many Nigerians understand.

It allows experienced interviewer to know when to make calls and recalls

A biased respondent, in a matter-affecting ego, may give false responses. It is very expensive and time consuming especially when respondents are widely scattered.

[...]

Most of the questions are answered The responses are always related to the objectives of the study. A major disadvantage of the structured questionnaire is the fact that it does not allow the views of the respondents, which may enhance the quality of the information collected. Unstructured questionnaire is a list of questions drawn on the study on which information is required. The questions are not accompanied by alternative answers as in the structured questionnaire.

The respondents are free to provide their own responses. Example of a question in an unstructured questionnaire is "What is your monthly salary?" Unstructured questionnaire are not difficult to construct since no question is accompanied by alternative answers. It has the following advantages. It allows for the views of the respondents The questionnaire are not difficult to construct since no question is accompanied by alternative answers.

However, the unstructured questionnaire has the following disadvantages. It is not easy to complete and analyze. Many of the responses supplied by the respondents may not be related to the objectives of the study. The respondents may not answer all questions, especially those considered difficult by them.

This will definitely reduce the quality of the information obtained. In drawing a standard questionnaire, the following must be considered. List all the questions you want to ask. Decide whether to use direct or indirect questions or both.

A question must be limited to an idea. Ask simple and interesting questions before the difficult and uninteresting ones. A question must mean the same thing to all the respondents. The language must be at the level of the respondents.

What you have learnt in this unit consists of the nature of statistical data, the sources of the data and the methods of collecting the data. The unit to follow will build upon this.

SUMMARISING DATA

You learned the methods of collecting data in the previous unit.

Construct relative frequency, cummulative relative frequency for both ungrouped and grouped data

When data are collected, they are collected in such a way that there is no particular arrangement of the values. This unordered array of values does not facilitate the process of analyzing the data. We can observe the distance between succeeding values in the data. The data can be further summarized by preparing the frequency distribution for the data.

A frequency distribution is a table showing the values of the data and the number of occurrence of each of the values. The values will be represented by Xi and the frequency is represented by Fi Fi This table shows the values of the variable and their respective frequencies. Frequency Distribution of Grouped Data It is possible to have the frequency in example 3.2 without grouping the data because the values are not many. In situation where the values are in thousands or millions, it may be difficult to analyze the values if they are not grouped.

For many of sophisticated analyses, we need to group the data before analysis commences. They are mutually exclusive in the sense that a value is Placed in one and only one class interval. It may even take either form. In this unit there will be more examples of class intervals.

The class should not be too few and should not be too many.

[...]

These concepts will be defined and computed now.

Class Boundaries

The lower class boundary of a class interval is the mean of the lower class limit of the interval and the upper limit of the preceeding interval. Let us compute the lower class boundary of the interval 11-20.

The lower class boundary of the class interval is therefore

The upper class boundary of a class interval is the mean of the upper class limit of the class interval and the lower class limit of the succeeding class interval. The upper class limit of the class interval is 20 while the lower class limit of the succeeding class interval is 21. From these values, you will see that the upper class boundary of a class interval is the lower class boundary of the succeeding class interval. The classes can also be given in terms of class boundaries rather than class limits.

When this is done, there is overlapping of the class intervals. For example 3.2, if the class boundaries are used, we will have the following as our frequency distribution. There is need to summarize what we have done so far into class limits, class boundaries, class marks and class width.

Class Limits

Class Width The relative frequency of a value is defined by the total frequencies of all the values

contained in the set of values.

Going through the relative frequencies table presented in the unit, you will realize that the sum of the relative frequencies for a set of values is 1

In this unit, you have learned construction of relative frequency for ungrouped data. A further summary of data can be in form of cumulative relative frequency. The cumulative relative frequency of a value is the cumulative frequency of the values divided by the total frequency of the values contained in the array of data.

GROUPED DATA

The cumulative frequency of grouped data is constructed the same way as that of ungrouped data.

Cumulative

In this unit you have learned how to arrange data in ordered form. You have also learned how to summarize data, both grouped and ungrouped data, into frequency distribution, relative and cumulative frequency distributions. You have also learned some basic concepts that will aid the analysis of grouped data latter in subsequent units. What you have learned in this unit will facilitate further presentation and analysis of the data in the subsequent units.

In this unit, we have summarized data using frequency distribution, relative and cumulative frequency distributions. Some concepts that will aid the graphical presentation of the unit that will follow this unit were also defined.

Graphing Frequency Distribution Histogram Frequency Polygon Bar Charts Simple Bar Charts

Multiple Bar Charts You learned in Unit 3 the summarizing of data using frequency distribution, relative and cumulative relative frequency distributions. In this unit, you will learn how to represent these data graphically. Graphical representation has a way of showing the salient features of data without having to interpret the column of numbers. Some other data that are not summarized into frequency distributions will be represented with charts in this unit.

Histogram

The vertical axis is used for the frequencies and the horizontal axis for the class boundaries. You still need to recall how the class boundaries are computed in unit 3 we will not discuss it here.

Fre que ncy

FIGURE 4.1 Histogram Showing The Data In Example 4.1.

30 Mode 40

The histogram can be used to estimate the mode of the distribution. You learned in unit 3 that class mark of a class interval is the mean of the lower and the upper class boundaries or limits of the class interval. If you look at table 4.1 that is of interest here, there is no class interval 5-9 at the beginning and there is no class interval 55-59 at the end of the distribution. We need to bring

these intervals in and assign a frequency of 0 to each of them.

Let us now compute the class mark for the class intervals. The next activity is to plot the frequency against the class marks. The frequency is on the vertical axis and the class mark on the horizontal axis.

Frequency polygon must be plotted on a graph sheet

This curve is very important in the determination of
median, quartriles, percentiles, semiinterquartile range, that will be discussed in some subsequent units.

Interval

Less than You will realize that the class interval 5-9 was introduced. A frequency of 0 was also assigned to the class interval since the original table did not show the class. This is done so that the ogive can take its origin from the horizontal line. From the table you will see that all the values that are less than 24.5 are contained jn class interval 5-9, 10-14, 15-19 and 20-24.

You should know that ogive can only be plotted on a graph sheet.

[...]

Another way of representing data graphically is by means of bar chart. A bar chart shows vertical bars with equal width to represent the values of a variable in some intervals of time. The bars must be drawn on graph sheet and they must have equal width. These will be discussed in this unit.

In a simple bar chart, we use the bars to represent the value of a variable in a period of time.

January

February You are required to represent the data on a simple bar chart. Another bar chart that shows the total value for a time period and the values of the components that makeup the total is the component bar chart. In this case, the bar for the total value for a period is divided into the values for the components that make up the total. Suppose a hotel has three departments A, B, C from where sales are made and the annual records of the net profit of the departments for three consecutive years are as presented below.

Net Profits in the Departments You are to represent the values of the net profits with the aid of a component bar chart. You will need to plot the values of the components A, B, and C for three years. The first year will show a bar of length 9.4cm divided into 3.2cm for A, 3.4cm for B, and 2.8cm for C. You will then repeat the exercise for the values of A, B and C in years 2000 and 2001. There must be a legend to show the shading of the component.

Using the values of net profit in example 4.5, construct a multiple bar chart for the hotel for the period three years. In this exercise, you will draw single bar for each component of every year. The values of the Items represented with pie chart are proportional to the area of the sectors that represent them. In the case of pie-chart, the sectorial angles are computed for the items based on their values and on the total values of the items.

Total Value of all items

You are required to represent the data on a pie char.

9625 x 3600 = 1260

This is to say that the monthly income of the worker is N27, 500. 360° is used in the calculation because the sum of angles at a point is 360°. If the sectoral angles are computed correctly, the sum of the angles must be equal to 360°. Suppose the basic elements of cost of a restaurant are expressed as percentages of sales as follows.

[...]

You learned in this unit the graphical representations that summarize the data further after the frequency distribution has been constructed.

[...]

Of all these graphs and charts only pie chart can be drawn on plan sheet of paper, all the others are to be constructed on graph sheets. This unit taught graphic representation of frequency distributions in terms of histogram, frequency polygon and ogive. The unit also taught graphical representation of other forms of data with bar and pie charts. Subsequent units will teach the computation of measures of central tendency from ungrouped and grouped data.

[...]

In unit 4, you learned how to construct graphs from frequency distribution and charts from raw data. The graphs and the charts gave us the trend and the patterns in the data. As beautiful and useful as these graphs and charts are, they do not give us any precise understanding of the reality of the information contained in the data. This will help us make quicker and better decisions without the need to consult our original observation.

List the properties, advantages and disadvantages of the arithmetic mean Compute weighted mean for some data

We use average many times to mean ~he arithmetic mean. We compute arithmetic mean for both ungrouped and grouped data. We also compute arithmetic mean, which we henceforth in this unit call mean, for both the sample and the population from where the sample is drawn. The mean of ungrouped data is the summation of the values in the set of data divided by the number of values in the set of data.

For a sample the number of values is denoted by n, that is the sample size, and for the population the population size is given as N...

Mean= x = 7+5+8+10+11+6+3+4+10+9+13+2

You will realize that there is no frequency distribution for this example.

Where Assumed mean A = 13

You will see that with the method of assumed mean, we should obtain the same value of mean

we had before. You have learned how to compute the arithmetic mean of ungrouped data. When the values are many in a set of data, there is the need to group them into class intervals. You learned about this in unit 3.

We need to take some time to compute the mean of grouped data. The earning per share of some firms is presented below with the frequency distribution. Earning per share You are required to calculate the mean of the distribution. The class mark becomes the Xi we will use in the computation.

Immediately this is done, the whole distribution is reduced to the form of an ungrouped data with frequency distribution. You should recall that class mark is the mean of the upper and lower class boundaries of a class interval.

From the computation, you will still realize that the mean of a grouped data is given as

The assumed mean method can also be used here.

[...]

We still obtain the same mean as before. For the distribution below, compute the mean, using the two methods we have used in this unit. Some times we are interested in showing the relative importance of some items. When we attach weighs to items we can only compute the weighted mean.

Xw = Weighted mean The percentage scores of a student in some courses and the credit units of the courses are given below:% Score

Calculate the weighted percentage mean for the scores.

Advantages And Disadvantages Of Arithmetic Mean

The arithmetic mean has the following advantages Mean is the best known of all the averages Mean can be used for further mathematical process. Mean is used to perform statistical procedures such as estimation and hypothesis testing.

[...]

The histogram drawn from a frequency distribution can reveal certain features about the distribution. It can reveal whether the distribution is symmetry or skewed.

A symmetry distribution has a peak at the middle of the distribution with the slopes on the either side of the peak virtually equal to each other

There can be positive and negative skews.

SK = 3 If the SK computed is positive, there is positive skew standard deviation

If the value is 0, there is a symmetrical It gives better understanding of a frequency distribution to know the relative position of mode, median and mean in skewed distribution.

On the basis of the peakedness of the distribution we can show three different types of

distribution

FIGURE 8.10 Platy Kurtic You learn how to compute the fractile that is quartiles, deciles and percentiles, in this unit you also learned skewness and Kurtosis in the unit. The subsequent unit will teach measures of dispersion to give more understanding about distribution.

Using the distribution shown below, compute The lower and the upper quartile The 48th and 72th percentile The 9th and the 5th decile

With the idea of histogram drawn from the task, show the type of skewness of the distribution.

MEASURES OF DISPERSION

The means of the distribution computed do not show whether the values of the distribution are clustered closely together or they spread. It is even not unusual to see two sets of values with same mean but different degrees of spread. The per capital income may be high and majority of the populace is still very poor.

What is the range of the following values?

It can be influenced by extreme values.

Mean = 420 = 21

Quartile Deviation In unit 8, we computed values of quartile from distribution. You may need to go back to this computation. While interquartile range measures how far we need to go from the median on either sides before we can include one half the values of the set of data, quartile deviation or semi-interquartile range measures the average range of one-fourth of the data.

Where Q3 = upper quartile

Q1= lower quartile The annual electricity bills in a household are given below and arranged in ascending order.

Quartile deviation = 730 - 570 = 80

A measure of dispersion that enables us determine, with high degree of accuracy, where the values of a frequency distribution are located in relation to the mean, is standard deviation. We can compute it for both the sample and the population from where the sample is taken. When it is from the samples, it is a statistic denoted by S. .

F¡

You should note that class mark is the mean of the upper and the lower class boundaries or limits.

S2 = 50 x 15735 – 8152

There may be the need to compare two distributions in terms of their variability. The standard deviation of the distribution alone can not be used to achieve comparison without knowing the means of the distribution. Again the unit of values used in the distribution may not be the

same. One unit may be in kilograms another unit may be in litre.

This makes comparison of the variability not easy since it is not easy to compare two things with different units.

Marks

No of Student You learn in this unit different measures of dispersion namely, range, mean deviation, quartile deviation, interquartile range and coefficient of variations.

No of workers

Levin R.1 Statistics For Management 4`h ed., Prentice-Hall of India Private Limited New Delhi.

Operations in Set Theory

In this unit you will learn the basic concepts in the set theory and solve some problems involving the use of Venn diagrams.

By the end of the unit, you should be able to

A Set is a collection of definite, distinct objects called elements or members of the set. For the set theory, the set is written in Capital letter, Set A. After each element there must be a comma. The common bracket is not allowed.

Where A

The two sets cannot overlap. Such sets that do not overlap are called disjoint sets. There are three factories J, K, L supplying goods to warehouses A, B, C, and D, the amount of supplies from the factories to the warehouses are shown in the table below.

The rest part of P C = 13 – x The rest part of P B = 14 – x The rest part of B C=5-x Physics only =34- =34-13+x-14+x-x 7+x Chemistry only = 36 - =8+x Biology only = 41- = 12+x

Now that we have obtained the values of all the components, we will add up the values. The sum of the values will be equal to 75.

[...]

The number that were neither Yoruba nor nutritionists The number that were neither Yoruba nor women nor nutritionists In this unit, you learned the definition of set and some basic concepts in set theory. You also learned the union and the intersection of sets. Venn diagram is treated to solve problems in set theory. Exercises and examples given are to assist you further in the understanding of the unit.

Set has been defined as collection of definite, distinct objects called elements or members of the set. The union and intersection of sets are treated to give a good understanding of how numerical problems in set theory could be solved.

[...]

In this unit you will learn other aspects of probability concepts, such as permutation and combination. Many times, the focus is on computing the probability of some event or that of events. Some times, the total number of possible events is large. Under this condition, you need to have some method for counting the number of such events.

[...]

With the use of factorials, you can solve questions on the number of ways, objects or persons can be arranged in a line. Suppose there are 3 waiters on a restaurant, standing in a line against a wall, waiting for quests. We can also use a graphical means to solve the problem with the aid of a tree diagram.

Computation of permutation for objects that are different

Permutation can be defined as an ordered arrangement of objects. 11.1 that shows possible arrangements of three objects taken three at a time gives the possible permutation of 3 objects taken 3 at a time. In some situations, as in many cases in Nigeria, there more people applying for the number of vacancies available. For instance, we may wish to fill three positions and 15 people are applying.

There are some ways we can solve the problem. There can be the use of logical reasoning and factorial method. We have 15 people available and 3 positions to fill. The first position can be filled in one of 15 ways, the second in one of the 14 ways and the third in one of 13 ways.

The resulting possible ways will be the product of 15,14 and 13 that is 15 x 14 x 13 = 2730 ways. The example we are considering can be solved mathematically if we make distinction between the three positions to be filled. What we are saying here is that we are having permutations of objects that are all different.

[...]

You have learned the permutations of objects when all the objects are different. This is not always the case in every situation. In some instances one or more subsets of the items cannot be distinguishable. In permutation we are required to determine how many distinguishable arrangements are possible.

In this case there are n objects from which the permutations is made, we have items, nr n2, n/k that are distinguishable. There are 6 positions available to be tilled by 6 persons from 3 tribes in Nigeria on equal basis. Combination is an arrangement of objects without regard to order. There are 6 ways of doing the arrangement.

[...]

Some time is devoted to this unit because of the importance of the concepts in this unit to the study of the subsequent units in this course The unit gives the permutations of r object, from n objects as n! when all the objects are different. It also gives the permutation of objects r1, r2, r3 rk from n objects to be when the objects are not r1!r2rl3 rk! all different.

SOME ELEMENTARY PROBABILITY CONCEPTS

While some events are certain to occur some others can never occur.

*** Understand elementary properties of probability**

Probability is defined as a value between 0 and 1 that shows the likelihood of occurrence of an event. If the probability is 1, it means the event is sure to occur. The probability that a man born of a woman will die one day is 1.

[...]

What is probability of having a success if the birth of a male child is regarded as a success.

N=5

The relative frequency concept is also used to define probability. Suppose a test is conducted for twenty students in the course Statistics for Management and the results of the test are as presented in the frequency distribution table below. Calculate the probabilities of the students scoring a particular score in the distribution. This is the frequency of 2 divided by the total frequency.

Relative frequency

This will be the probability that a student scores 6 or 7 or.

This will be equal to the sum of these: Probabilities =

Two events are mutually exclusive when they cannot occur together.

When two events are mutually exclusive, the intersection of the events will be zero A B

Equally likely events are those events in which there is no reason to expect one event rather than the others to occur. For instance, if a coin is not biased the head and tail are equally likely. If a die is not biased, 1, 2, 3, 4, 5, 6 are equally likely to occur. The first one is derived from the definition.

Probability of an event is defined as a value between 0 and 1 that shows the likelihood of occurrence of the event. Therefore probability of an event cannot be negative that is less than 0. The probability can also not be more than.

Therefore the probability of an event with characteristic E is 0

If events are both mutually exclusive and collectively exhaustive then the sum of the probabilities of the events is equal to 1.

Head and tail in a throw of a coin

With these examples, no other event can occur in each of the cases. Since the events cannot occur simultaneously, the events are said to be mutually exclusive.

P + P

You should note that if two events are mutually exclusive, their intersection will be zero.

CALCULATING THE PROBABILITY OF AN EVENT

In this section we will treat two types of probabilities - the unconditional and conditional probabilities. Let us explain these probabilities by solving a question that will bring the understanding of unconditional and conditional probabilities out properly.

P = n

Using the table in example 3.1, calculate the following probabilities. P is equal to the number that is common to both male and married divided by the total number of secretaries.

[...]

Suppose there are four types of appliances in the kitchen denoted by A1, A2, A3 and A4 and the colours of the appliances are of 3 types C l, C2, C3. From the table calculate probability that an appliance picked at random is of.

Colour C2 given type A4

The unit stated the properties of probability. The unit presented the calculation of the probability of an event and worked out problems involving conditional and unconditional probabilities. You learned the three different points of view, classical, relative frequency and subjective, from which we can discuss probability.

That is P + P + P ------ P = 1

The table below shows the programmes in the National Open University and the types of Employers that will employ the graduates of this programmes. It also shows the number of graduates the employer's employ.

When the two events are mutually exclusive P = 0

Let us go back to the table we had for 3.3 in unit 12 to illustrate numerically the addition rule.

In unit 12 under conditional probability, we discussed that the probability of event A given B is equal to P= P for P ≠0

This relationship holds if P is not equal to 0. If P is equal to 0, the definition of a probability 0 < P < 1 will be violated. Then P=P P. This is obtained by rearranging the preceding equation. What is the probability of joint occurrence of M1 and S2 in the table used for example 4.1.

What it remains is for us to compute E2 .

[...]

This unit has built more on the concepts and techniques of probability by showing the application of these concepts to the relative frequency of distribution of a discrete random variable. The unit has also giver the means of computing mean and variance of the probability distribution of a discrete random variable. Future units will show more of the types of probability distributions both of discrete and continuous random variables. The table below shows the

number of industrial accidents suffered by 50 employees of a large manufacturing firm in a year.

[...]

This unit treats Binomial distribution, which is one of the most widely used probability distributions in statistics. The unit will treat the properties of Binomial distribution, Binomial tables and measures of central tendency and dispersion in the distribution.

[...]

The probability of a success is arbitrarily denoted as P the probability of failure is 1 - P. which is denoted as q.

porP+q=1

The trials are independent, that is the outcome of a trial does not affect the outcome of any other trial. For instance if you want to have four children and you regard the birth of a male as a success then the value of success will be 0,1,2,3,4. You shall know that if we have n trials where n = 1,2,3,4 --------- n, we can have X successes where x=0,1,2,3 - - - - n. Suppose probability of a success is P, that of failure will be 1-P.

Therefore the probability of X success will be P x qn-x This is the general formula for binomial distribution.

3!x2x1

If you know the number of trials n and number of successor X, you see and multiplying it by probability of success raised is number of successor and multiply the product by probability of failure raised to the number of failure. If there are n trials and x successors number of failure will be failure. Suppose it is known that the probability that a person in a profession will be at home in particular time of the day is 0.6 or 60%.

= l0 x

We can have the possible arrangements for the 2 successes. This is equal to.

[...]

You will realize that we also obtain the same value as we obtained before. For n trials, X success and P as probability of success, we can have probability distribution for the binomial distribution. You should know that when we obtain probability distribution for all possible success, the sum of the probabilities is equal to 1.

[...]

Suppose the probability that a family will have a male child is given as 0.4. If a couple in the family decides to have five births, what is the probability of having?.

More than three male births

At most 4 male births Binomial Tables Binomial tables are used to solve problems involving

probability of success.

Cummulative binomial table

having known this we will locate the table of interest for the question being solved. In the table, for any x, the probability of x > xl is given. For example, if we have five trials that is n=5 and the p=0.4, we will locate the table of n=5, p=0.4. We are expected to have 0,1,2,3,4,5 successes when we have 5 trials.

In the table we will have probabilities against x=1,2,3,4 and 5, that of 0 is not given. You should note that the value against 0 is not given but it is 1.00.

Between 5 and 11 successes

The mean of a binomial distribution is given as n x p, that is the number of trials multiplied by the probability of success For n = 5 and p = 0.

MEASURES OF BINOMIAL DISTRIBUTION

The variance of a binomial distribution is given as n p q where n is number of trials, p is probability of a success and q=1-p is probability of a failure. This unit treats a binomial distribution as a distribution in which one of the two possible mutually exclusive outcomes must occur in a trial that is derived from Bernoulli, process. The units teaches the computation of probability of successes given x=number of, successes, given n= number of trials and p= probability of success.

POISSON DISTRIBUTION

This unit will teach the definition of Poisson distribution, the Poisson process, the use of Poisson table and use of Poisson distribution as an approximation to a binomial distribution.

Typographical errors in a book

Poisson Process In Poisson distribution, there is said to be random occurrences of some event, and that this is brought about by the Poisson process.

Daniel and Terrel gave the characteristics of this process as follows

The table you will be using is that of cumulative Poisson Distribution.

For example, for value of x = 5 we will have a probability value showing P = P+P+P

Let us now solve problems involving Poisson Distribution by using the table of cumulative Poisson Distribution.

Here = np = μ

Np was given in unit 15 as equal to the mean of binomial distribution. Generally it is accepted that for n to be large here n must be greater than or equal to 20, that is n > 20. " For p to "be small, it must be less than or equal to 0.05, that is p < 0.05. We will answer it using concepts of binomial distribution and Poisson distribution and compare the results.

1 - p

Since n = 25 is considered to be large and P = 0.02 is small then we can approximate the binomial distribution with Poisson distribution. If you compare the corresponding values, you will realize that the Poisson distribution can actually approximate binomial distribution when n is large n > 20 and p is small p < 0.05. In this unit, you learned that the Poisson distribution is a distribution of a discrete random variable, which is widely used in business. You also learned the characteristics of a Poisson process.

The unit taught the use of a Poisson distribution table p and also showed Poisson distribution as an approximation to the binomial distribution when n > 20 and p< 0.05.

[...]

The assumption made in binomial distribution is that the probability of a success, P must be constant for all trials. This assumption can hold where there is infinite population that is sampled or when a finite population is sampled with replacement. In practice, replacement of elements is not done. This will then create problem with analysis of binomial distribution of finite population when replacement is not done.

The hyper geometric distribution then solves the problem of unequal value of probability of success, p, that may occur when sampling finite population without replacement. This unit therefore treats the nature of hypergeometric function, solving of problems involving hypergeometric distribution and computation of measures for hypergeometric distribution.

Explain the concept of hypergeometric distribution

Solve problems involving hypergeometric distribution Compute measures for hypergeometric distribution. Suppose 12 of the balls are yellow and 8 of them are blue. If two balls are to be drawn successfully, the probability that the first ball will be yellow is 12 . If the first ball is taken away and it is yellow, there will remain 11 yellow balls and 8 blue balls.

With this in mind the assumption made for binomial distribution may not be valid for finite population except sampling is done with replacement. Sampling with replacement is not practical with many real life situations, especially when investigation requires the destruction of elements. hypergeometric model is used in situations where we wish to determine the probability of a given number of success or failures. Solution Of Problems Involving Hypergeometrie Distribution Use Of Tree Diagrams This is one of the ways to solve problems involving hyper geometric distribution.

We start by denoting defective by D and not defective by N. We will all put the successive probability of selecting the defective or not defective one in bracket. From the above tree diagram, we can identify the following arrangements with their probabilities.

NDN = 10 X 2 X

We have now obtained the distribution let us answer the questions now For O detective, we have NNN, that all of them are not defective. The required probability is 0.5455.

The formula for P under these conditions is given as

If four of the cocoa beans are selected, what is the probability that there are?.

2 defective beans?

The unit teaches how the weakness of constant value of probability of a success in binominal distribution could be solved when sampling a finite population without replacement. Future unit will discus probability distribution of a continuous random In a crate of eggs containing 30eggs, 4 eggs are found to be defective. If 3 of such eggs are picked at random what is the probability of getting.

NORMAL DISTRIBUTION

In this unit, you will learn this important probability distribution of continuous random variable and some of its applications.

Read the areas under the curve of a normally distributed population using the normal curve areas

The total area under the curve above x-axis is equal to l. This means that 50% of the area is to the left of the mean and the other 50% to the right of the mean. The area enclosed by μ _ is 68% of the total area.

68% of the area is between 1 standard deviation from the mean in

95% of the area under the curve is enclosed between 2 standard deviations from the mean in both directions. These are the characteristics of the normal distributions that will still be of serious use in the other courses that you may encounter in this programme. The Standard Normal Distribution The Normal Distribution, as already said is determined by the values of its and μ . Any change in the values of either or μ will give another normal distribution curve.

In view of this, there is a need to have a standard normal distribution that is determined by a mean of O and a standard deviation of 1. Now that the values of X in normal distribution has been transformed to values of Z in standard normal distribution, we still need to find the area under the curve for a given value of Z for this, we use the Normal Curve Areas. In finding the area under the curve for a value of Z, we get the area between O and that given value of Z = P . For instance if we are looking for area under the curve when Z = 2.5.

The area will be P . If the value of Z = 0.55, we will check Z = 0.5under 0.05. This is equal to For Z = 1.43, we look at the column 1.4 under 0.03. We will observe that 35 is equidistant from 1.

We divide by 2

Let us now apply the use of table to solving some problems. For a distribution, that is normal distribution, what is the probability the a Z picked as random from the population of Z's will have a value Z < 2. There are two areas involved here, the one to the right of the mean and the one to the left of the mean this is half of the area under the curve and is 0.5. The area to the left is P 0 < z < 2.3which is equal to 0.4893 we obtained before.

The required area is 0.5 + 0.4893 = 0.9893.

FBQ1: Any measure indicating the Centre of a set of data, arranged in an increasing or decreasing order of magnitude, is called a measure of: _____
Answer: Central tendency

FBQ2: Scores that differ greatly from the measures of central tendency are called:_____
Answer: Outliers

FBQ3: The total of all the observations divided by the number of observations is called: _____
Answer: Arithmetic mean

FBQ4: The sample mean X- is an example of a: _____
Answer: Statistic

FBQ5: The population mean μ is an example of a:
Answer: Parameter

FBQ6: The arithmetic mean is highly affected by: _____
Answer: Extreme values

FBQ7: If a constant value is added to every observation of data, how would the value of the arithmetic mean behave? _____
Answer: Increased by the constant

FBQ8: The median is considered a robust measure because it is resistant to: _____
Answer: Outliers

FBQ9: What effect will the elimination of extreme scores at the bottom of a data set have on the mean? _____
Answer: Increase the mean

FBQ10: The elimination of extreme scores at the top of the set has the effect of:
Answer: Reduce the mean

FBQ11: The sum of deviations taken from mean is: _____
Answer: 0

FBQ12: The sum of the squares of the deviations about mean divided by the number of observations is: _____
Answer: Variance

FBQ13: If $\sum i=110Xi-50=100$ then sample mean X- will be: _____
Answer: 60

FBQ14: For a certain distribution, if $\sum(X -20) = 25$, $\sum(X- 25) =0$, and $\sum(X-35) = -25$, then X- is equal to:
Answer: 25

FBQ15: The sum of the squares of the deviations of the values of a variable is least when the deviations are measured from: _____
Answer: Arithmetic mean

FBQ16: If X-=100 and Y=2X – 200, then mean of Y values will be: _____
Answer: 0

FBQ17: Step deviation method or coding method is used for computation of the_____
Answer: Arithmetic mean

FBQ18: If the arithmetic mean of 20 values is 10, then sum of these 20 values is:_____
Answer: 200

FBQ19: Ten families have an average of 2 boys. How many boys do they have together?
Answer: 20

FBQ20: If the arithmetic mean of the two numbers X1 and X2 is 5 if X1=3, then X2 is:_____
Answer: 7

FBQ21: Given X1=20 and X2=-20. The arithmetic mean will be: _____
Answer: 0

FBQ22: The mean of 10 observations is 10. All the observations are increased by 10%. The mean of increased observations will be: _____
Answer: 11

FBQ23: The frequency distribution of the hourly wage rate of 60 employees of a paper mill is as follows: The mean wage rate is:N_____
Answer: 59.00

FBQ24: The sample mean X- of first n natural numbers is: _____
Answer: (n+1)/2

FBQ25: The sum of deviations is zero when deviations are taken from:_____
Answer: Mean

FBQ26: When the values in a series are not of equal importance, we calculate the: _____
Answer: Weighted mean

FBQ27: When all the values in a series occur the equal number of times, then it is not possible to calculate the:_____
Answer: Weighted mean

FBQ28: The mean for a set of data obtained by assigning each data value a weight that reflects its relative importance within the set, is called:_____
Answer: Weighted mean

FBQ29: The arithmetic mean of 10 items is 4 and the arithmetic mean of 5 items is 10. The combined arithmetic mean is: _____
Answer: 6

FBQ30: The midpoint of the values after they have been ordered from the smallest to the largest or the largest to the smallest is called:____
Answer: Median

FBQ31: The first step in calculating the median of a discrete variable is to determine the: _____
Answer: Array

FBQ32: The suitable average for qualitative data is: _____
Answer: Median

FBQ33: If the smallest observation in a data is decreased, the average which is not affected is: ____
Answer: Median

FBQ34: Sum of absolute deviations of the values is least when deviations are taken from: _____
Answer: Median

FBQ35: The frequency distribution of the hourly wages rate of 100 employees of a paper mill is as follows: The median wage rate is:N ____
Answer: 59.00

MCQ1: The values of the variate that divide a set of data into four equal parts after arranging the observations in ascending order of magnitude are called:
Answer: quartiles

MCQ2: The lower and upper quartiles of a symmetrical distribution are 40 and 60 respectively. The value of median is:
Answer: 50

MCQ3: If in a discrete series 75% values are less than 30, then:
Answer: Third quartile =30

MCQ4: The probability of the amount X (in million Naira) of investment in the shares of ABC Company is given as follows:Find E (X).
Answer: 35/18

MCQ5: The mean of first 2n natural numbers is:
Answer: (2n+1)/2

MCQ6: If X-1, X-2, X-3,….X-k be the arithmetic means of k distributions with respective frequencies n1, n2, n3, … ,nk, then the mean of the whole distribution X-c is given by:
Answer: ∑nX-∑n

MCQ7: The combined arithmetic mean of two sets of means is calculated by which formula?
Answer: n1X-1+n2 X-2n1+n2

MCQ8: Extreme scores will have the following effect on the median of an examination
Answer: They may have no effect

MCQ9: The probability of the amount X (in million Naira) of investment in the shares of ABC Company of Adewale is given as follows:E(X2). is actually 5.89. What is the variance of X?
Answer: 2.11

MCQ10: The grouped frequency distribution shown below is to be used to answer the following questionWhich class is the modal class?
Answer: 20 - 24

MCQ11: The grouped frequency distribution shown below is to be used to answer the following questionWhich class is the median class?
Answer: 20 - 24

MCQ12: The grouped frequency distribution shown below is to be used to answer the following questionWhat is the cumulative frequency of the modal class?
Answer: 14

MCQ13: For a standard normal distribution, what is the values of the mean and variance?
Answer: Mean = 0, variance = 1

MCQ14: Given the set of numbers: 15, 16, 12, 11, 19, 18, 13 then åx is,
Answer: 1600

MCQ15: The mean of 63, 19, 52, 10, 95, 18 is
Answer: 42.83

MCQ16: The median of 63, 19, 52, 10, 95, 18 is
Answer: 35.5

MCQ17: The mode and the range of the above data are
Answer: 40, 40

MCQ18: The geometric mean of 6, 8, 10 and 16 is
Answer: 9.36

MCQ19: The harmonic mean of 6, 7, 8 and 9 is
Answer: 7.33

MCQ20: Given that the mean of a distribution is 160, the mode is 150 and the standard deviation is 25. Find the coefficient of skewness
Answer: 0.4

MCQ21: A set of sales from an outlet produced the following: 16, 14, 18, 10, 12 compute the variance
Answer: 8

MCQ22: The coefficient of variation for data set whose mean is 10 and variance 100 is
Answer: 100%

MCQ23: For a symmetric distribution
Answer: The mean, median and the mode are equal

MCQ24: Which statistics is found by summing all the values and dividing by the number of observations?
Answer: The arithmetic mean

MCQ25: How would you describe the skewness of a distribution whose mean is smaller than the median?
Answer: Negatively skewed

MCQ26: What level of measurement is required for the median?
Answer: ordinal

MCQ27: The Nigeria Stock Exchange (NSE) index increased from 961 in 1980 to over 9,500 in 2003. The annual rate of increase is best described by the
Answer: Geometric mean

MCQ28: What is the shape of a frequency distribution with an arithmetic mean of 12,000 pounds, a median of 12,000 pounds, and a mode of 12,000 pounds?
Answer: symmetric

MCQ29: Given that the mean of a distribution is 60, the mode is 50 and the standard deviation is 25. Find the coefficient of skewness:
Answer: 0.4

MCQ30: A set of experimental animals was fed in a special diet for one week and produced the following gains in weight: 6, 4, 8, 10, 12 compute the variance:
Answer: 8

MCQ31: The coefficient of variation for data set whose mean is 16 and variance 10 is
Answer: 19.8

MCQ32: Given the mean = 60 and variance is 625, find the coefficient of variation
Answer: 41.7%

MCQ33: Suppose A and B are independent events with PA=0.2, PB=0.6..What is PAB=?
Answer: 0.2

MCQ34: In a shipping organization, it is observed that the total number of items imported is 400 units. If you are to categorise these items into types of commodity with the aid of a pie chart, what angle would 160 units of chemical take?
Answer: 144 degree

MCQ35: The data collected by questionnaires are usually classified as what type of data?
Answer: Primary data