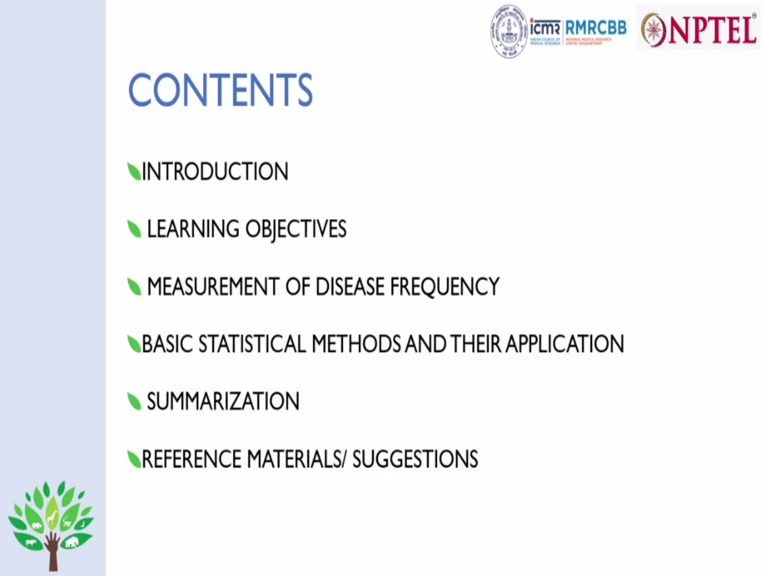
One Health
Prof: Dr. Srikanta Kanungo
Department of
Indian Institute of Technology –

Lecture - 18
Basic Statistical Methods and their Application and the Measurement of Disease Frequency

Hello everybody. Today in the series we are going to discuss basic statistical methods and their application and the measurement of disease frequency in the unit, Applied Epidemiology and Public Health in One Health Research. So, I am Dr. Srikanta Kanungo, Scientist working at RMRC Bhubaneswar. So, I am going to discuss these things.

(Refer Slide Time: 00:41)



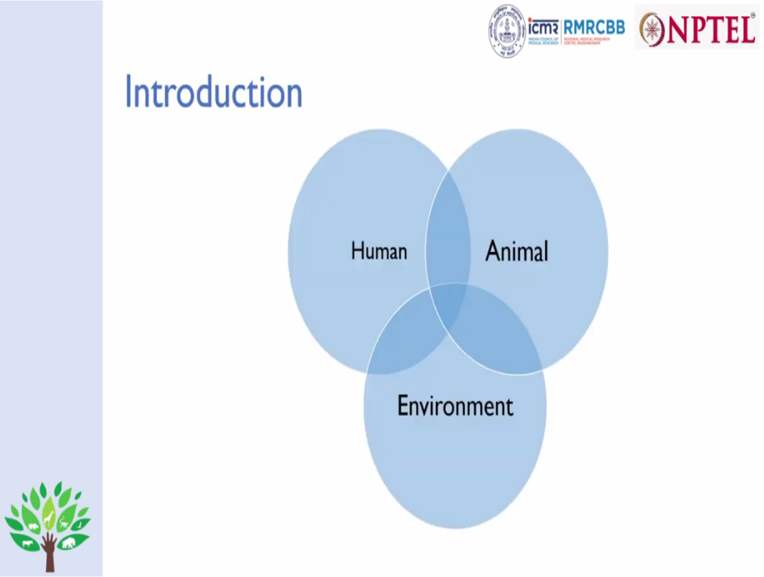
CONTENTS

- INTRODUCTION
- LEARNING OBJECTIVES
- MEASUREMENT OF DISEASE FREQUENCY
- BASIC STATISTICAL METHODS AND THEIR APPLICATION
- SUMMARIZATION
- REFERENCE MATERIALS/ SUGGESTIONS

icmr RMRCBB
NPTEL

So, the contents of this lecture will be I will introduce the topic to you then the learning objectives that what are the objectives you are going to achieve in this lecture series and the measurement of disease frequency followed by basic statistical methods and their application and finally I will summarize what I have taught.

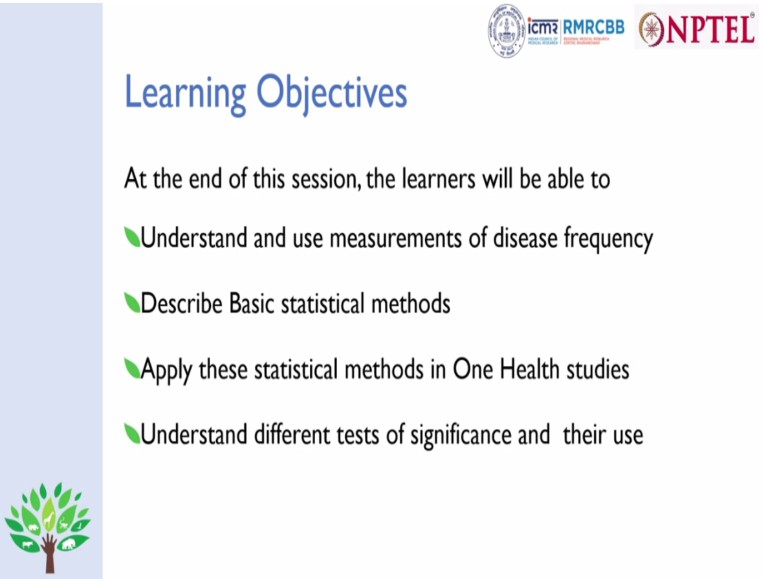
(Refer Slide Time: 01:02)



Introduction

So, as you know that one health is a collaborative and multidisciplinary approach involving participants from human, animals and environment.

(Refer Slide Time: 01:13)



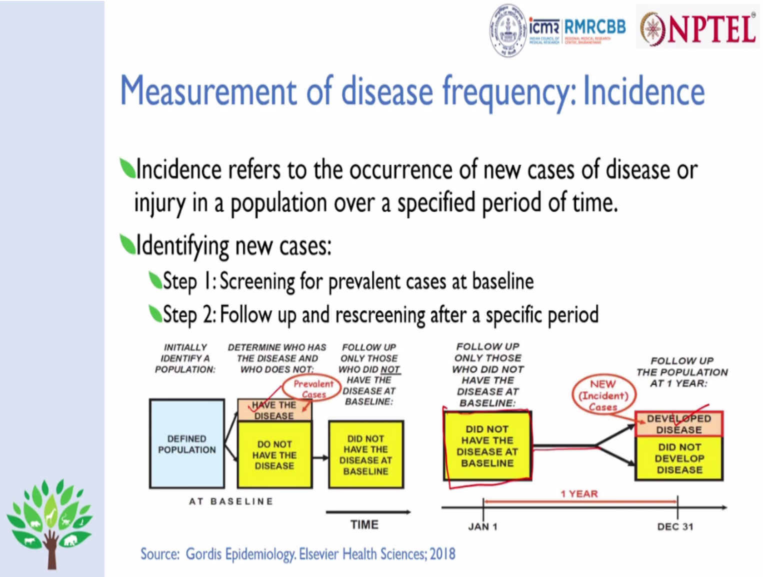
Learning Objectives

At the end of this session, the learners will be able to

- Understand and use measurements of disease frequency
- Describe Basic statistical methods
- Apply these statistical methods in One Health studies
- Understand different tests of significance and their use

So, at the end of this session the learners will be able to understand and use measurements of disease frequency and describe the basic statistical methods, apply the statistical method in one health research and understand the different tastes of significance and their use.

(Refer Slide Time: 01:29)



Measurement of disease frequency: Incidence

- Incidence refers to the occurrence of new cases of disease or injury in a population over a specified period of time.
- Identifying new cases:
  - Step 1: Screening for prevalent cases at baseline
  - Step 2: Follow up and rescreening after a specific period

INITIALLY IDENTIFY A POPULATION:
DETERMINE WHO HAS THE DISEASE AND WHO DOES NOT:
FOLLOW UP ONLY THOSE WHO DID NOT HAVE THE DISEASE AT BASELINE:
DEFINED POPULATION
HAVE THE DISEASE
Prevalent Cases
DO NOT HAVE THE DISEASE
DID NOT HAVE THE DISEASE AT BASELINE
AT BASELINE
TIME
FOLLOW UP ONLY THOSE WHO DID NOT HAVE THE DISEASE AT BASELINE:
DID NOT HAVE THE DISEASE AT BASELINE
NEW (Incident) Cases
FOLLOW UP THE POPULATION AT 1 YEAR:
DEVELOPED DISEASE
DID NOT DEVELOP DISEASE
1 YEAR
JAN 1
DEC 31

Source: Gordis Epidemiology. Elsevier Health Sciences; 2018

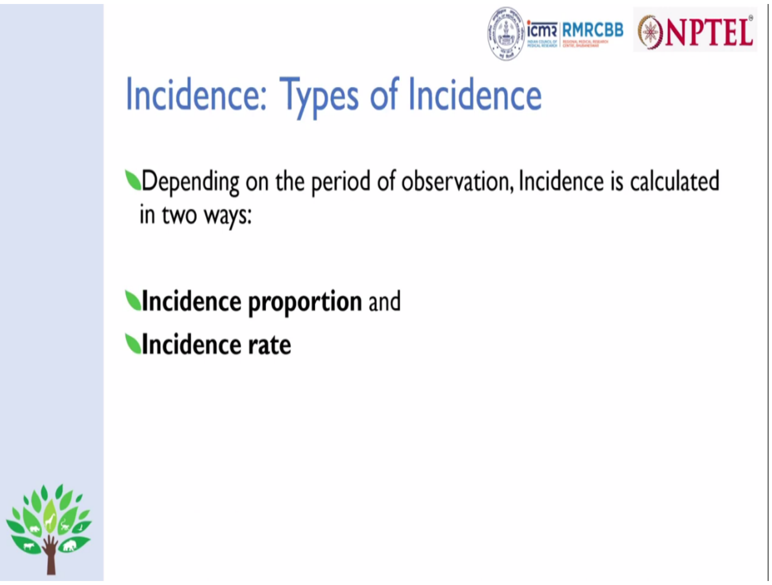
So, starting with the disease frequency I will be talking about the incidence. So, the incidence refers to the occurrence of the new cases of disease or injury in a population over a specific period of time. So, these are the new cases that means while starting up this any research the population is not having the disease or these are all the conditions and with follow up they will develop the cases that means those are the new cases.

So, we call them the new cases or the incidence cases. So, as I have said the step one will screen about the prevalent cases, the prevalence is something I will talk about in the next slide and the step two will follow up and rescreening after a specific period whether the people without having the disease or disorders are developing the diseases. So, this picture of the schematic diagram I have presented from Gordis Epidemiology.

And you can see that the initially the identify populations we have identified and in that first at the baseline we have done a survey and we have seen this much of people are having the disease. So, these are the people prevalent cases, these are the old cases and then we will follow up of the population without having the disease or the desired then if these populations after the follow up they are developing an illness or the disease or our desired disease or desired conditions.

We call them the new cases or the incidence cases. So, you can see that these prevalent cases these are the baseline. These are the prevalent cases and the old cases and after follow up of those who did not have the disease they are developing after the one year this much of the cases they are developing the new cases, this population are the new or incidence cases so this is incidence.

(Refer Slide Time: 03:31)

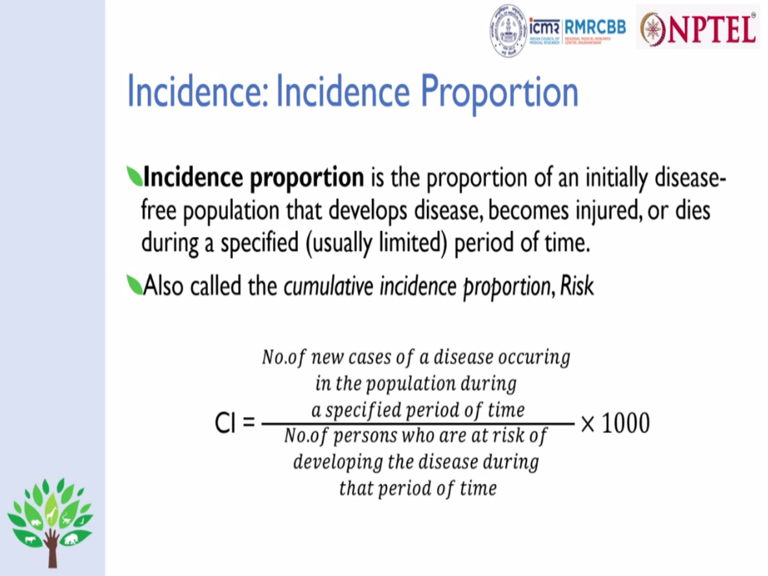


Incidence: Types of Incidence

- Depending on the period of observation, Incidence is calculated in two ways:
- **Incidence proportion** and
- **Incidence rate**

So, incidences are two types. One is incidence proportion another is incidence rate.

(Refer Slide Time: 03:37)



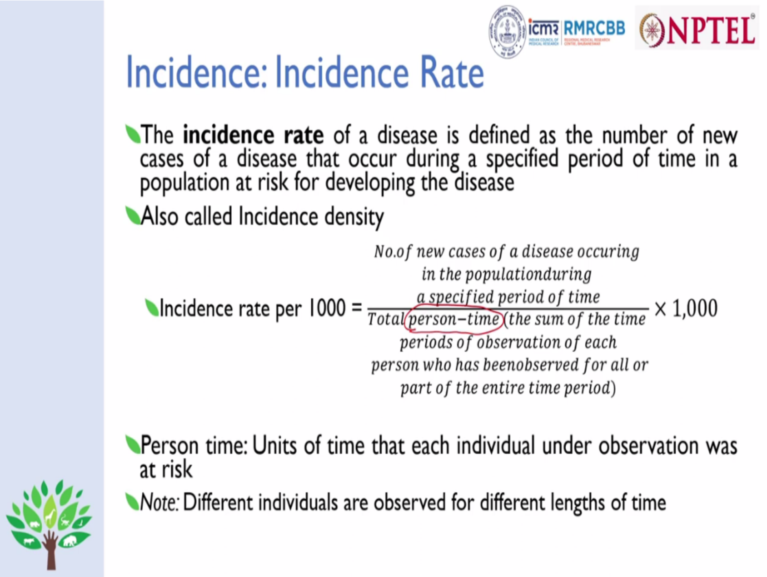
Incidence: Incidence Proportion

- **Incidence proportion** is the proportion of an initially disease-free population that develops disease, becomes injured, or dies during a specified (usually limited) period of time.
- Also called the *cumulative incidence proportion, Risk*

$$\text{CI} = \frac{\textit{No.of new cases of a disease occuring in the population during a specified period of time}}{\textit{No.of persons who are at risk of developing the disease during that period of time}} \times 1000$$

So, incidence proportion is the proportion of an initially disease free population that develops disease becomes injured or dies during a specified period of time. So, also called the cumulative incidence proportion risk. So, cumulative incidence which is represented as CI here. So, number of new cases of a disease occurring in the population during specified period of time divided by number of person who are at risk of developing the disease during that period of time into 1000. So, this is the cumulative incidence or incidence proportion.

(Refer Slide Time: 04:13)



Incidence: Incidence Rate

- The **incidence rate** of a disease is defined as the number of new cases of a disease that occur during a specified period of time in a population at risk for developing the disease
- Also called Incidence density
- Incidence rate per 1000 =

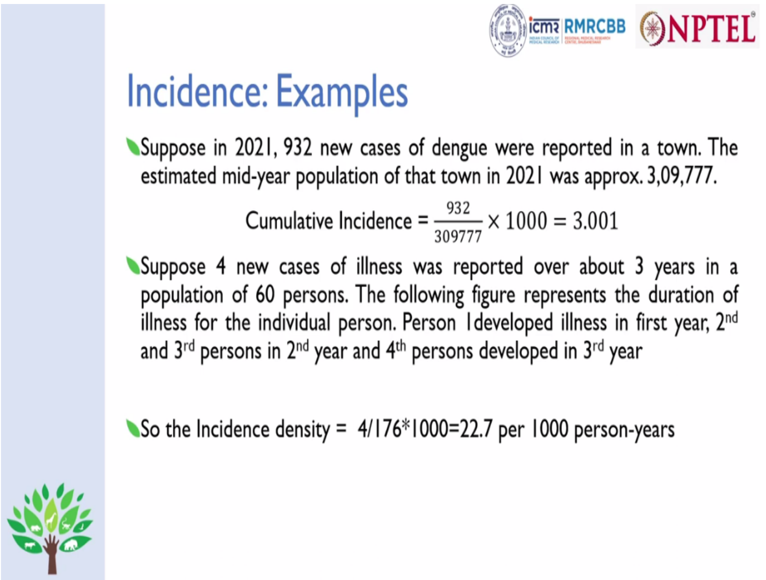
$$\frac{\begin{array}{c}\textit{No.of new cases of a disease occuring}\\ \textit{in the populationduring}\\ \textit{a specified period of time}\end{array}}{\begin{array}{c}\textit{Total person−time (the sum of the time}\\ \textit{periods of observation of each}\\ \textit{person who has beenobserved for all or}\\ \textit{part of the entire time period)}\end{array}} \times 1{,}000$$

- Person time: Units of time that each individual under observation was at risk
- *Note:* Different individuals are observed for different lengths of time

Whereas incidence rate which is also known as the incidence density. It is defined as the number of new cases of a disease that occurs during a specified period of time in a population at risk of developing disease. So, the incidence density or rate per thousand is equal to number of new cases of a disease occurring in the population during a specified period of time divided by total person time.

So, this is very important to know the denominator here we are using person time instead of the number of persons. The denominator of the total person time that is the sum of the time period of observation of each person who has been observed for all or part the entire time period into 1000. So, the person time is the unit of time that each individual under observation was at risk. So, I will come to one example so it will be more clear to you.

(Refer Slide Time: 05:10)



Incidence: Examples

- Suppose in 2021, 932 new cases of dengue were reported in a town. The estimated mid-year population of that town in 2021 was approx. 3,09,777.

$$\text{Cumulative Incidence} = \frac{932}{309777} \times 1000 = 3.001$$

- Suppose 4 new cases of illness was reported over about 3 years in a population of 60 persons. The following figure represents the duration of illness for the individual person. Person 1developed illness in first year, 2nd and 3rd persons in 2nd year and 4th persons developed in 3rd year

- So the Incidence density = 4/176*1000=22.7 per 1000 person-years

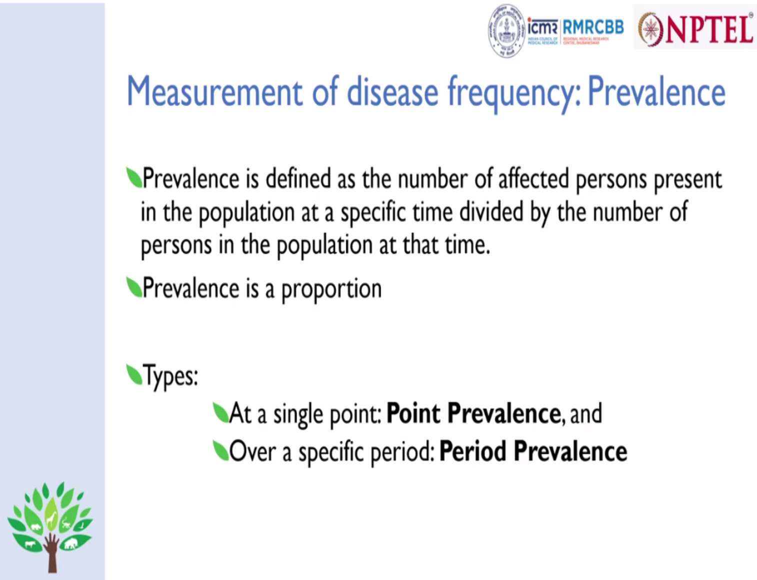
So, let us see suppose in 2021 there are 932 new cases of dengue were reported in a town and the estimated media population of the town is in 2021 is approximately 3,09,777. So. the cumulative incidence or the incidence proportion will be 932 divided by the total population which are at risk that is the total population of the town that is 3.09,777 into 1000 is equal to 3.001.

At the same time suppose in another scenario suppose 4 new cases of illness was reported over about 3 years in a population of 60 persons. So that means we are following 60 persons and this 60 persons are at risk. The following figure represents the duration of the illness of individual of the person. Suppose, person 1 develop illness in the first year and second and third person developed in the second year and the fourth person developed in the third year.

So, we are following up these 60 people up to 3 years. So, that means total 60 into 3, 180 person year we are following up, but you can see the first person developed illness in the first year that means we are not following up that cases when it has been developed a disease that means it has lost two following year. So, we are not following these first cases for the next 2 years.

The second and third cases developed in the second year that means we have not developed or not followed them in the next year that is the third year and the fourth person developed in the third year. So, here we have not followed of 180 person year, we have followed up 170 person year. So, the incidence density will be 4 divided by 176 into 1000 is equal to 22.7 per 1000 person year. So, this is the difference between incidence rate and incidence per person.

(Refer Slide Time: 07:15)

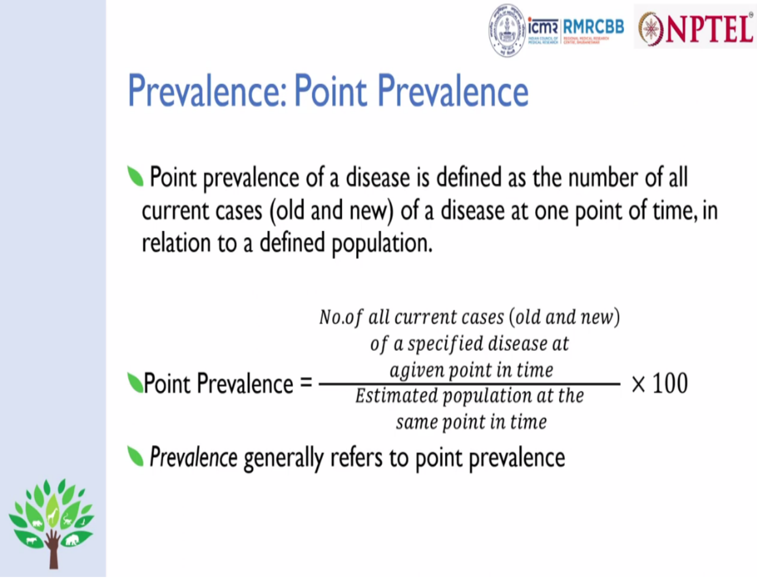


Measurement of disease frequency: Prevalence

- Prevalence is defined as the number of affected persons present in the population at a specific time divided by the number of persons in the population at that time.
- Prevalence is a proportion
- Types:
  - At a single point: **Point Prevalence**, and
  - Over a specific period: **Period Prevalence**

So, next is measurement of disease frequency that is prevalence. So, prevalence is defined as the number of affected person present in the population at a specific time divided by the number of person in the population at that time. So, prevalence is a proportion. So, there are two types of prevalence generally it is point prevalence and period prevalence.

(Refer Slide Time: 07:36)



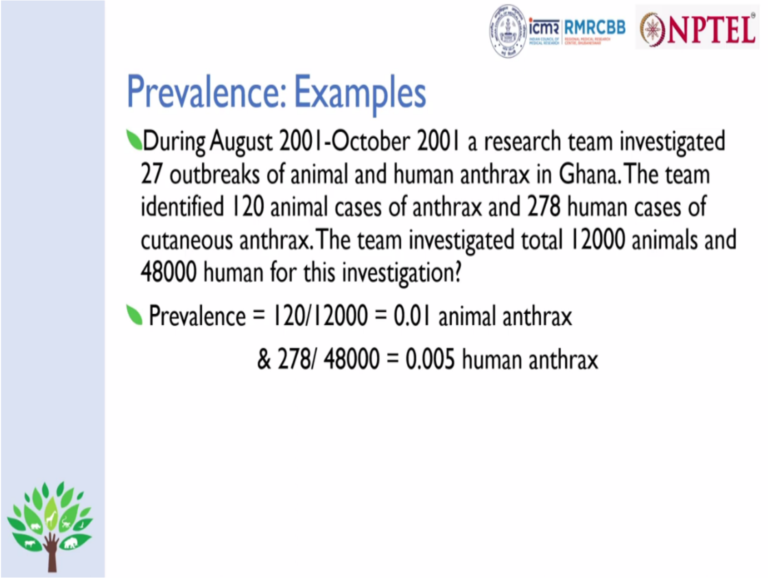
Prevalence: Point Prevalence

- Point prevalence of a disease is defined as the number of all current cases (old and new) of a disease at one point of time, in relation to a defined population.
- Point Prevalence = $\dfrac{\begin{gathered}\textit{No.of all current cases (old and new)}\\ \textit{of a specified disease at}\\ \textit{agiven point in time}\end{gathered}}{\begin{gathered}\textit{Estimated population at the}\\ \textit{same point in time}\end{gathered}} \times 100$
- *Prevalence* generally refers to point prevalence

So, point prevalence of a disease the number of all current cases; current cases means old and new cases that means those cases are already been diagnosed are in the community or in the population those who have developed recently those are the new cases of a disease at one point of time as the name suggested point prevalence. So, at a point of time what is the prevalence is known as point prevalence.

So, this is the formula of point prevalence so in that number of all current cases of a specified disease at a given point of time divided by estimate population at the same point time into 100 as prevalence is a proportion so it is we usually multiplied which is 100. So, prevalence generally refers to a point prevalence.

(Refer Slide Time: 08:22)



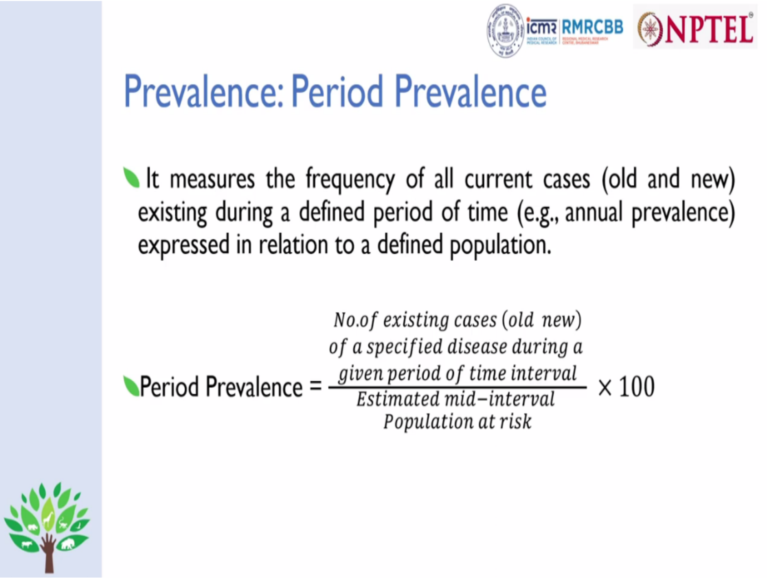
Prevalence: Examples

- During August 2001-October 2001 a research team investigated 27 outbreaks of animal and human anthrax in Ghana. The team identified 120 animal cases of anthrax and 278 human cases of cutaneous anthrax. The team investigated total 12000 animals and 48000 human for this investigation?
- Prevalence = 120/12000 = 0.01 animal anthrax

  & 278/ 48000 = 0.005 human anthrax

So, next is example of this point prevalence. So, suppose in August 2001 to October 2001, a research team investigated 27 outbreaks of animals and human anthrax in Ghana and the team identified 120 animal cases of anthrax and 278 human cases of cutaneous anthrax. So, the team investigated total 12,000 animals and 48,000 humans for their investigation and the prevalence will be 120 that is the number of cases that is old plus new cases of animal divided by total number of animals who has been investigated as 12,000 is equal to 0.01 animal anthrax. So, 278 divided by 48,000 is equal to human anthrax prevalence.

(Refer Slide Time: 09:12)



Prevalence: Period Prevalence

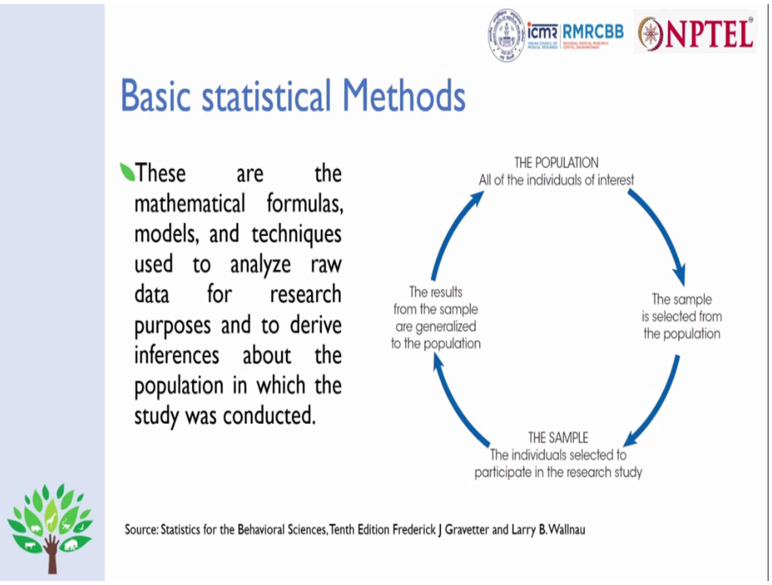
- It measures the frequency of all current cases (old and new) existing during a defined period of time (e.g., annual prevalence) expressed in relation to a defined population.

- Period Prevalence = $\dfrac{\begin{array}{c}No.of\ existing\ cases\ (old\ \ new)\\ of\ a\ specified\ disease\ during\ a\\ given\ period\ of\ time\ interval\end{array}}{\begin{array}{c}Estimated\ mid{-}interval\\ Population\ at\ risk\end{array}} \times 100$

So at the same time the period prevalence is it measures the frequency of all current cases existing during a defined period of time example annual prevalence. Suppose, there will be 28 number of asthma cases reported in a one year suppose from 2021 January to 2021 December. So, there will be annual prevalence that is a period prevalence. So, expressed in relation to a defined population.

So, the period prevalence formula will be number of existing case of a specified disease during a given period of time interval divided by estimated mid interval population at risk into 100.

(Refer Slide Time: 09:49)



Basic statistical Methods

- These are the mathematical formulas, models, and techniques used to analyze raw data for research purposes and to derive inferences about the population in which the study was conducted.

THE POPULATION
All of the individuals of interest

The sample
is selected from
the population

THE SAMPLE
The individuals selected to
participate in the research study

The results
from the sample
are generalized
to the population

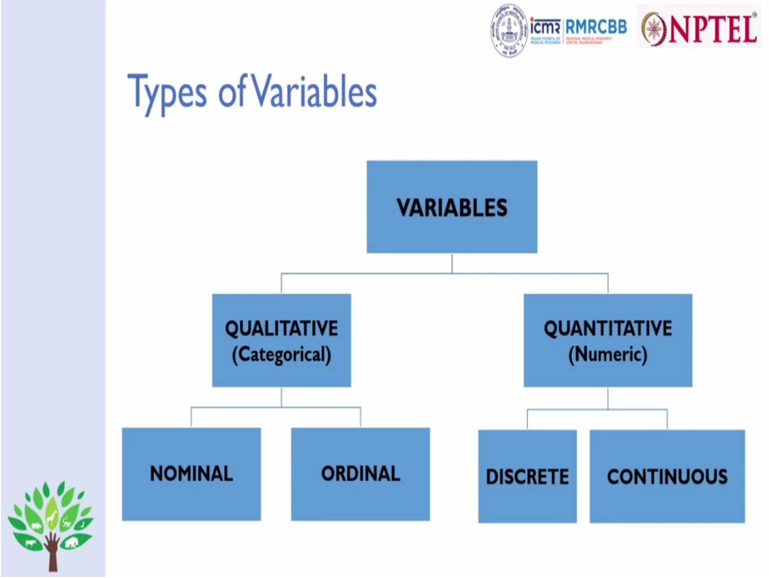
Source: Statistics for the Behavioral Sciences, Tenth Edition Frederick J Gravetter and Larry B. Wallnau

As today we are discussing about the quantitative researches. So, now I am going to discuss about basic statistical method. So, in any research we start with a population. So, the population we call the study population that means all the individual of our interest is known as the study population. So, it is not possible to do the study or measure the whole population. So, that is why we have to take a sample from the population.

And the sample which we will take from the whole population is known as the sample and we usually measure the individuals inside the sample those were selected to do our study, to do our research. So, when we will get some result of the sample we have to in turn we have to generalize the result to the whole population. So, this is a two way procedure that means we will take a sample from the population.

And we will measure that sample and whatever result we will get from the sample as per our interest of the variable. So, we will generalize that result to the whole populations. So, this is the basics of the statistics. So, the statistical methods are nothing, but the mathematical formula model and techniques used to analyze raw data of research purposes and to derive inferences about the population in which the study was conducted.

(Refer Slide Time: 11:18)



Types of Variables

So, for that actually we should know the variable. So what is a variable? Variable are characteristics of the conditions of an individual that can change from individual-to-individual because we have to measure our interest, the outcome, the exposure in our study. So, we should know what are the different type of variables? Variables can be a qualitative variable or a quantitative variables.

So, qualitative variable nothing, but the categories. Suppose you are married or not? So, the answer will be yes or no there are two categories, but quantitative, if I will ask those who are married the number of years you have got married so that is the quantitative variable so that means a numeric. So, in qualitative variable it can be nominal or ordinal. Nominal means simple names.

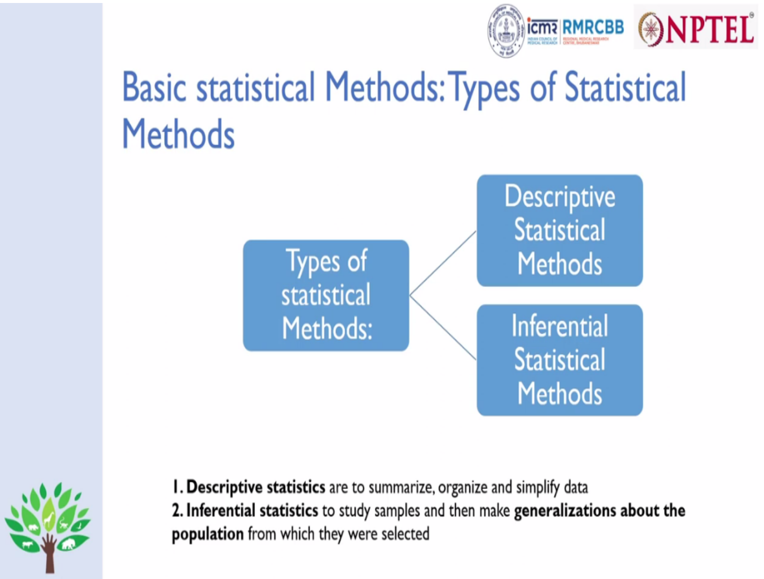
So, suppose yes, no or the name of the city Bhubaneswar, Mumbai, Chennai. So, these are the names so this is nominal variable, but an ordinal variable there will be certain orders mild, moderate, severe, taller, smaller. So, these are the different orders. So, the qualitative variable can be nominal or ordinal, but at the same time quantitative variable are discrete or continuous based on the counting.

If you can count the quantitative variable it is a discrete if we are have to measure that variable it is a continuous variable as for example suppose we want to see the antimicrobial resistance the number of drugs a sample is resistant. So, suppose there will be the sample is resistant to two antibiotics. So, we can set two antibiotics. So, suppose one sample is resistant to three antibiotics.

So, these are the counts of antibiotics the sample is resistant. So, this is a discrete variable, but continuous variable is suppose age, my age is 23.5 years. Suppose, my height is 178.5 centimeter so we have to measure it. So, these are continuous variable. So, if I will summarize this variables are two types, one is qualitative variable another is quantitative variable.

Qualitative variable can be nominal variable and ordinal variable and quantitative variable can be discrete variable or continuous variable.

(Refer Slide Time: 13:47)



Basic statistical Methods: Types of Statistical Methods

Types of statistical Methods:

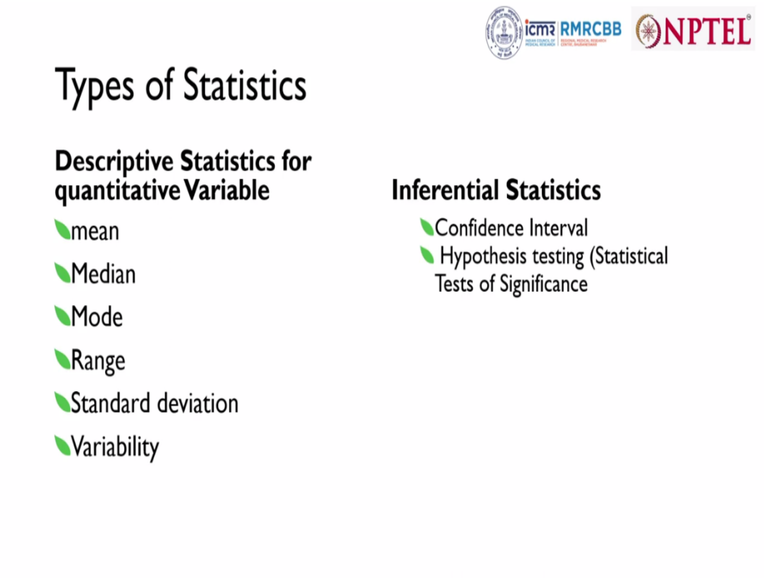
Descriptive Statistical Methods

Inferential Statistical Methods

1. **Descriptive statistics** are to summarize, organize and simplify data
2. **Inferential statistics** to study samples and then make **generalizations about the population** from which they were selected

So, basic statistical method which are used in the different research, these are two types. One is descriptive statistics another is inferential statistics. So, descriptive statistics are to summarize, organize or simplify the data and inferential statistics to study the sample and then generalize about the population from which they were selected.

(Refer Slide Time: 14:11)



Types of Statistics

Descriptive Statistics for quantitative Variable

- mean
- Median
- Mode
- Range
- Standard deviation
- Variability

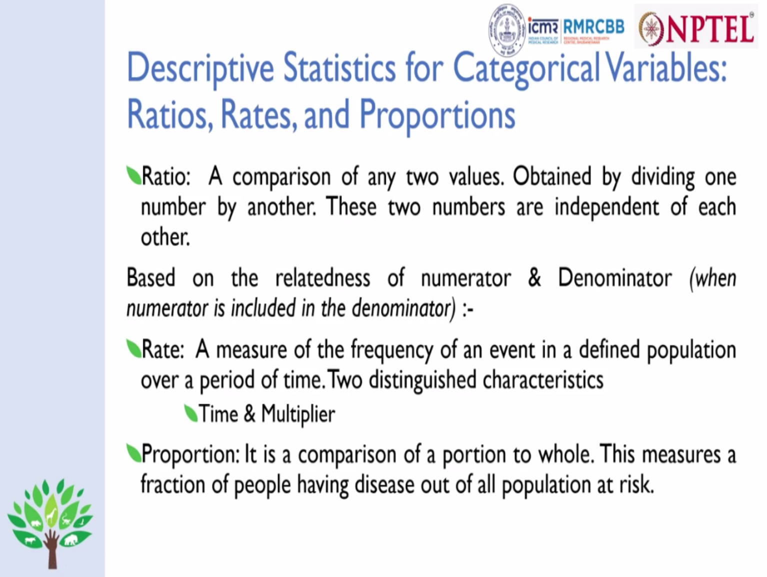
Inferential Statistics

- Confidence Interval
- Hypothesis testing (Statistical Tests of Significance

So, if we will see the types of statistics generally the descriptive statistics for the quantitative variable are mean, median, or mode these are the central tendency, these are the central value. So, mean is the simple average of the distribution the values of the distribution and median is the central value and mode is the most frequently occurred values in the distribution and there are dispersions are maybe range that means from minimum to maximum value or the standard deviation or the variability.

So, standard deviation is the individual values how it is deviated from the central value that is called standard deviations. So, in the inferential statistics there are two things, one is confidence interval another is hypothesis testing that is statistical test of significance.

(Refer Slide Time: 14:59)



Descriptive Statistics for Categorical Variables: Ratios, Rates, and Proportions

- Ratio: A comparison of any two values. Obtained by dividing one number by another. These two numbers are independent of each other.

Based on the relatedness of numerator & Denominator *(when numerator is included in the denominator)* :-

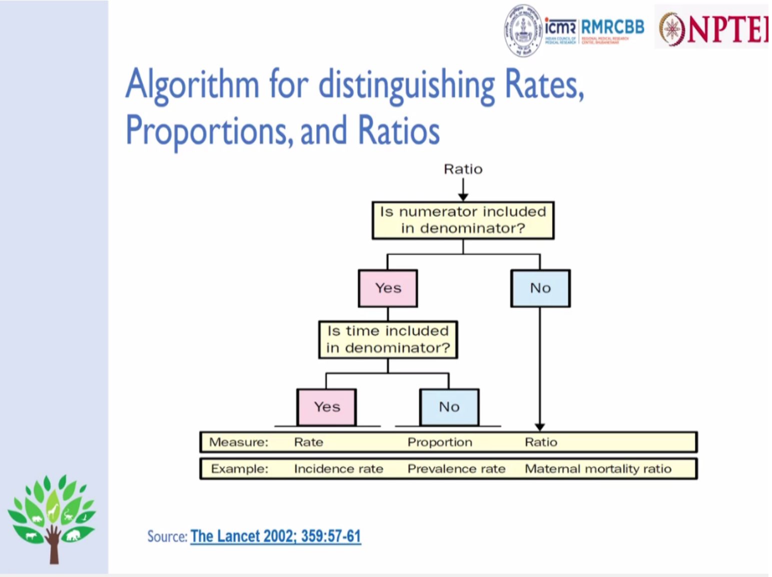
- Rate: A measure of the frequency of an event in a defined population over a period of time. Two distinguished characteristics
    - Time & Multiplier
- Proportion: It is a comparison of a portion to whole. This measures a fraction of people having disease out of all population at risk.

So, first we will discuss about descriptive statistics for the categorical variable. So, descriptive statistics for categorical variable can be expressed in term of ratio, rates and the proportion. So, the ratio that is nothing, but a comparison of two values that is obtained by dividing one number by simple another number and these two numbers are independent of each other then that time we call this is a ratio.

So, based on the relatedness of the numerator and the denominator when the numerator is included in the denominator, we call these things are rates or proportion. So, rate a measure of frequency of an event in a defined population over a period of time or is known as rate and it has two distinguished characteristics. One is the time and the multiplier, that means that rate must have one time factor and a multiplier.

In the proportion it is a comparison of a population to whole that means in proportion the numerator is a part of the denominator and this measures a fraction of people having disease out of all the population at the risk. So, you can say that prevalence is a proportion, but when we will calculate the incidence in a certain period of time so it will be a rate.

(Refer Slide Time: 16:14)



Algorithm for distinguishing Rates, Proportions, and Ratios

Ratio

Is numerator included in denominator?

Yes

No

Is time included in denominator?

Yes

No

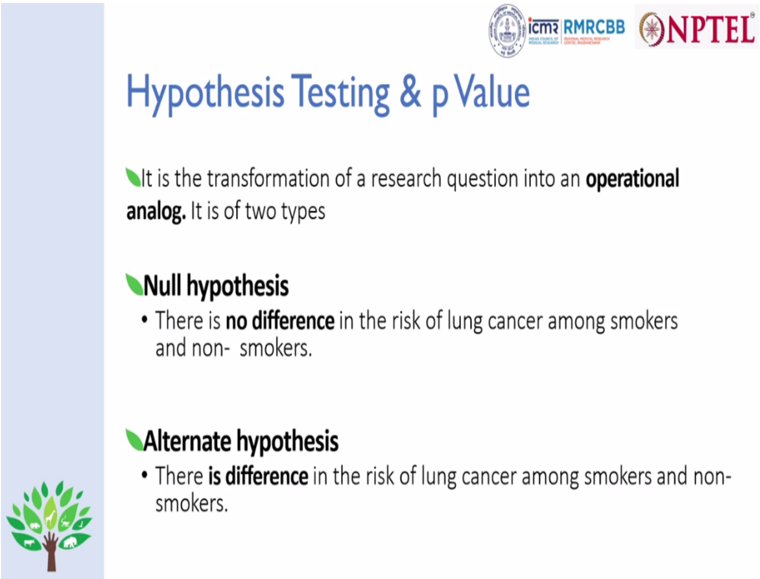
Measure:	Rate	Proportion	Ratio
Example:	Incidence rate	Prevalence rate	Maternal mortality ratio

Source: The Lancet 2002; 359:57-61

So, this is an algorithm for distinguished rates, proportion and ratio. So, ratio is simple when the numerator is not the part of the denominator that is a ratio, but when numerator is a part of the denominator that time it can be a rate or it can be a proportion. So, if the time is included in the denominator we call it is the rate, but when the time is not included in the denominator it is a proportion. So, these are the examples of rate, ratio and proportion.

This is very important in the quantitative research and the expression of the results of the quantitative research.

(Refer Slide Time: 16:50)



Hypothesis Testing & p Value

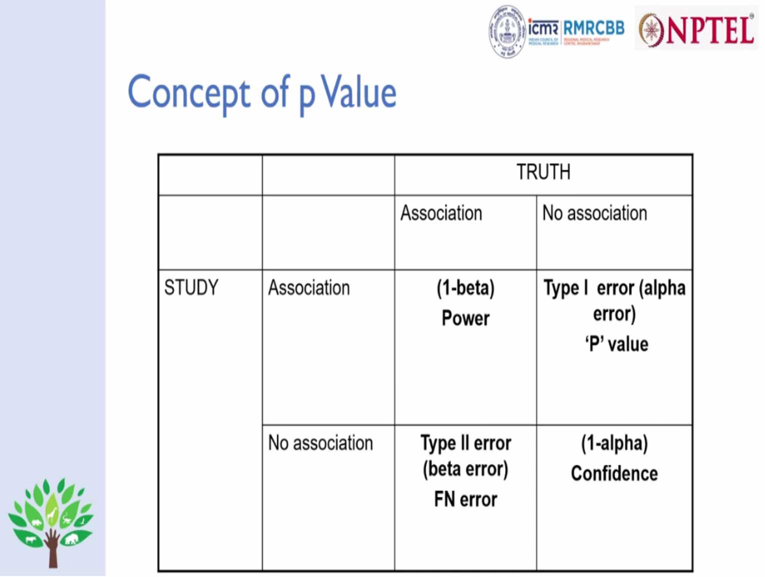
- It is the transformation of a research question into an **operational analog.** It is of two types
- **Null hypothesis**
  - There is **no difference** in the risk of lung cancer among smokers and non- smokers.
- **Alternate hypothesis**
  - There **is difference** in the risk of lung cancer among smokers and non-smokers.

So, next we are going to discuss about hypothesis testing and p value. So, before that I will talk about what is hypothesis? So, hypothesis it is the transformation of a research question into an operational analog that means what we are going to prove. Suppose as for example, misuse of antibiotic can lead to antimicrobial resistance or the Methicillin-resistant Staphylococcus aureus can cause more death than Methicillin-sensitive Staphylococcus aureus infection.

So, these are the different research questions that can be put in a form of statement that need to be proved or disproved so that is the hypothesis. So, in negative sense we can call it a null hypothesis there is no difference in the risk of lung cancer and smokers and non-smokers. So, this is an example of null hypothesis, but alternate hypothesis this is the difference there is a difference in the risk of lung cancer among smokers and non-smokers.

Usually in the statistics which starts from the null hypothesis we usually put the null hypothesis statement and we collect information and evidence to disprove or prove it. So, based on our evidences we test whether the null hypothesis is rejected or accepted.

(Refer Slide Time: 18:08)



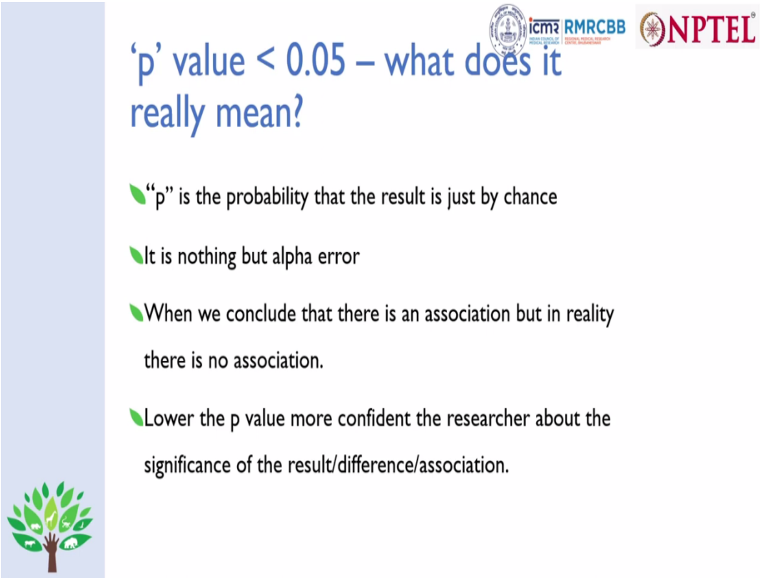
Concept of p Value

		TRUTH	
		Association	No association
STUDY	Association	**(1-beta)** **Power**	**Type I error (alpha error)** **'P' value**
	No association	**Type II error (beta error)** **FN error**	**(1-alpha)** **Confidence**

So, in this process usually there are four outcomes can happen that means suppose the null hypothesis is really true, but we will find it is not true or null hypothesis is false, but we will find it true. So, in these cases there can be some errors. So, these are alpha errors or beta errors. You can see in the example suppose in truth there is no association between one exposure variable another outcome variable.

But in a study we will find there is association that means we commit one mistake or one error so that error is known as type 1 error or alpha error and if there is no association, but we will find there is a association, but we will find there is no association. So, in that condition that is called type 2 error.

(Refer Slide Time: 18:59)



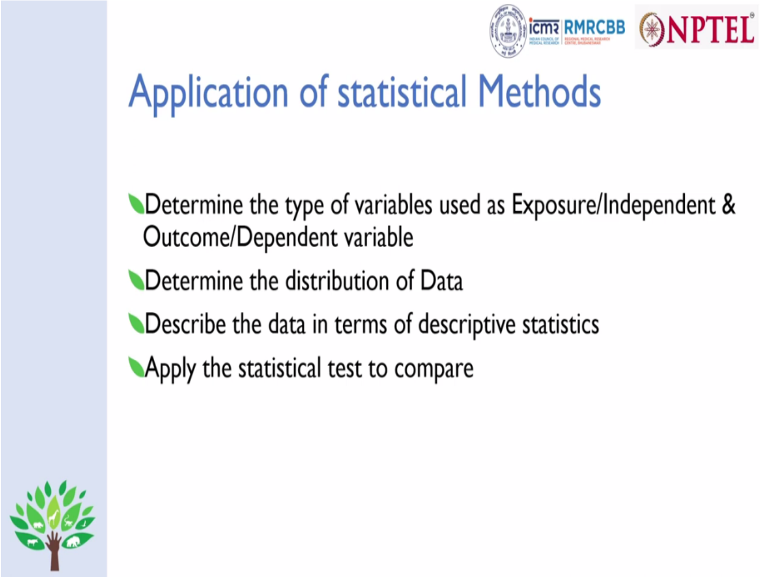
icmr RMRCBB
NPTEL

‘p’ value < 0.05 – what does it really mean?

- “p” is the probability that the result is just by chance
- It is nothing but alpha error
- When we conclude that there is an association but in reality there is no association.
- Lower the p value more confident the researcher about the significance of the result/difference/association.

So, p value, what is p value? So, p value is the probability that the result is just by chance and it is nothing, but the alpha error and when we conclude that there is an association, but in reality there is no association that is an alpha error and that is the probability of the alpha error is known as p value. So, lower the p value, more confident that the researcher about the significance of the result difference and the association.

(Refer Slide Time: 19:26)



icmr RMRCBB
NPTEL

Application of statistical Methods

- Determine the type of variables used as Exposure/Independent & Outcome/Dependent variable
- Determine the distribution of Data
- Describe the data in terms of descriptive statistics
- Apply the statistical test to compare

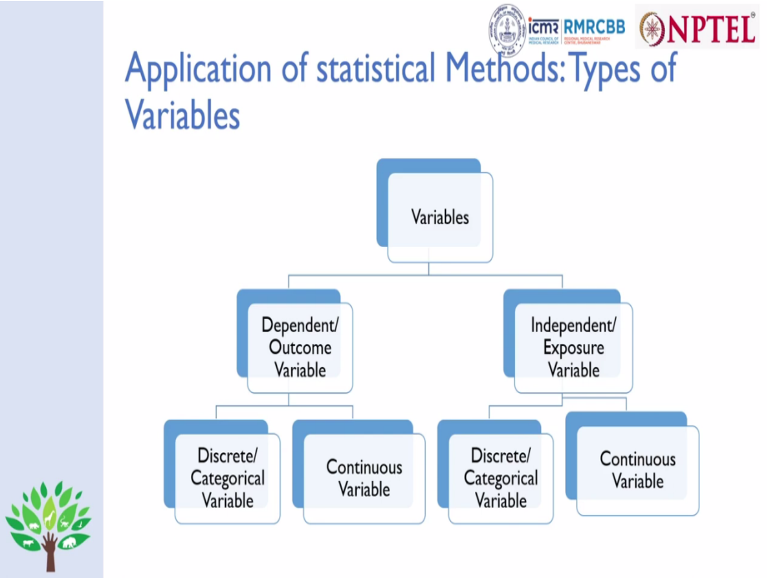
Application of the statistical methods. First, there are different steps how to use the statistical methods or the test of significance. So, first determine the type of variable used as exposure or independent and outcome on the dependent variable. So, let me explain what are the exposure variable and what is the outcome variable? In hypothesis, suppose we are putting on hypothesis that misuse of antibiotic can lead to antimicrobial resistance in the population.

So, in that case, the exposure is misuse of antibiotics and the outcome is that antimicrobial resistance. If I am putting another hypothesis that is Methicillin-resistant Staphylococcus aureus the infection has bad outcome than the Methicillin sensitive Staphylococcus aureus that means there is a association between the methicillin resistance or sensitive with the mortality.

So, in that cases the exposure will be Methicillin-resistant or Methicillin sensitive and the outcome will be the death. So, this is the exposure and the outcome variable. We have to find out what type of exposure variable is there and what type of outcome variable is there and then we have to find out the distribution of the data. So, in this session I am not going to tell about the distribution of the data, but we must understand the distribution of any variable or any continuous variable it can be a normal distribution or it is non-normal distribution.

There are certain methods we usually do that to find out the normality of a distribution then we have to describe the data in term of the descriptive statistics that means we have to use based on the variable what kind of descriptive statistics method we can use or the measurement we can use and then we have to finally apply the statistical test to compare.

(Refer Slide Time: 21:24)



Application of statistical Methods: Types of Variables

Variables

Dependent/ Outcome Variable

Independent/ Exposure Variable

Discrete/ Categorical Variable

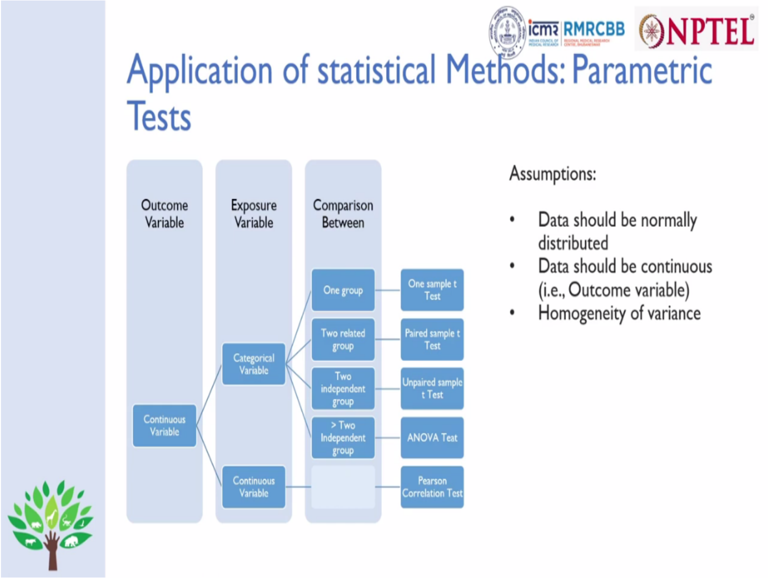
Continuous Variable

Discrete/ Categorical Variable

Continuous Variable

So, these are what I have discussed now the variables it can be dependent variable or outcome variable and independent variable or exposure variable. So, dependent variable can be a discrete variable or categorical variable or it can be a continuous variable. Similarly, the independent variable also can be the categorical variable or the continuous variable.

(Refer Slide Time: 21:48)



Application of statistical Methods: Parametric Tests

Outcome Variable

Exposure Variable

Comparison Between

Continuous Variable

Categorical Variable

Continuous Variable

One group — One sample t Test

Two related group — Paired sample t Test

Two independent group — Unpaired sample t Test

> Two Independent group — ANOVA Teat

Pearson Correlation Test

Assumptions:

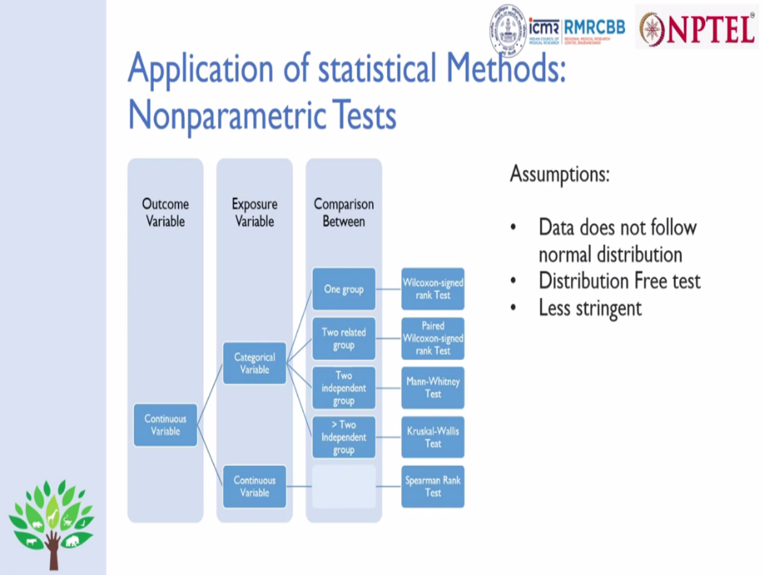
- Data should be normally distributed
- Data should be continuous (i.e., Outcome variable)
- Homogeneity of variance

So, now look at this picture so the outcome variable if it is a continuous variable and the exposure variable is a categorical variable or continuous variable. So, based on the comparison how many groups you are comparing. Suppose, there is one group we are comparing then we have to apply one sample t test. If there are two related groups that means before or after we are looking at the values.

So, then it will be a paired sample t test. If there are two independent groups or then we will use unpaired sample t test. If there are more than two independent group we can use ANOVA test. If the exposure variable is also a continuous variable and outcome variable is a continuous variable we can use the Pearson correlation test. So, all these things what I am telling these are parametric tests.

And we are assuming that the data is normally distributed and data should be continuous that is the outcome variable and homogeneity of the variance.

(Refer Slide Time: 23:00)



Application of statistical Methods: Nonparametric Tests

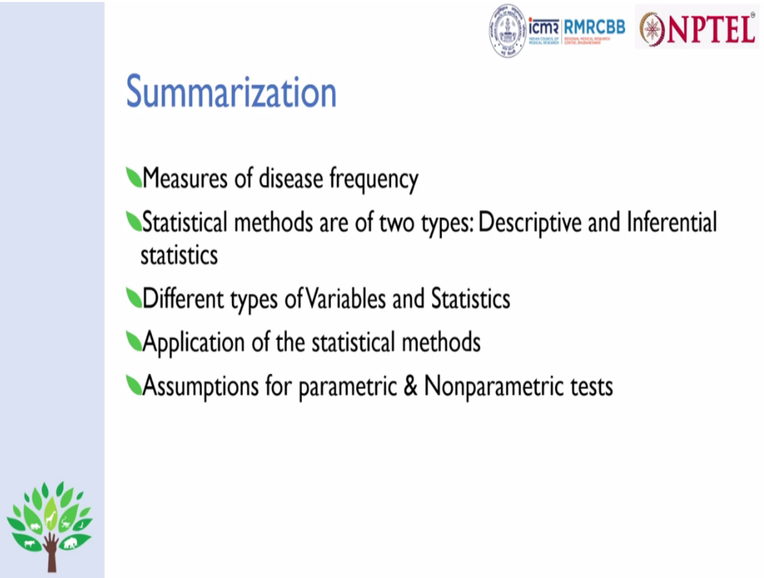
Assumptions:

- Data does not follow normal distribution
- Distribution Free test
- Less stringent

Next we can see there are different tests if the data is non normal that means data does not follow the normal distributions. So, in that case these are the tests. So, if there is one group then we can use Wilcoxon signed rank test that is two related groups we use pair Wilcoxon signed rank test if two independent groups are there we can use Mann-Whitney test and more than two independent groups we can use Kruskal Wallis test.

And if both the variables are continuous variable they are outcome variable and exposure variable, we use spearman rank test.

(Refer Slide Time: 23:37)



Summarization

- Measures of disease frequency
- Statistical methods are of two types: Descriptive and Inferential statistics
- Different types of Variables and Statistics
- Application of the statistical methods
- Assumptions for parametric & Nonparametric tests

Let us summarize the today's session. So, today we have discussed measures of disease frequency with incidence and prevalence also we have seen in this session there are different type of incidences that is incidence rate and incidence cumulative incidence and point prevalence and period prevalence. Then actually we have discussed there are different statistical method, the descriptive statistical method and the inferential statistical method.

Then different types of variables and statistics, application of the statistical method based on the variables and assumption of parametric and non-parametric tests. With this, I will stop this session.