APPLICATION OF STATISTICS IN HEALTH AND USES OF COMPUTER IN STATISTICS

INTRODUCTION

Public health professionals use several appropriate tools and techniques tor evaluating the health status of the community. Traditionally, these tools and techniques have consisted of the community's vital health statistics, which include the counts of birth, death, illness, and various rates and ratios that may be computed from them. Public health professionals and epidemiologist use vital statistics and epidemiological tools in determining the prevalence of a given condition at a point in time and incidence at which new conditions arise in the population.

RATES, RATIOS, AND TRENDS

Rate: It is a proportion with defined denominator termed as population. The **numerator is Occurrence of an event, which is counted in the same population**.

The fraction is generally multiplied by a constant, which for most of the rates is 1000. In other words the term rate usually is reserved to refer to those calculations that involve the frequency of the occurrence of some event. A rate is expressed in the form:

$$\left\{ \frac{a}{a+b} \right\} k$$

Where,

a = The frequency with which an event has occurred during some specific period of time.

a +b = The number of persons exposed to the risk of the event during the same period of time.

k = Numbers such as 100, 1000, 10,000 or 100,000, where most of the rates are per 1000.

Ratio: Ratio is a fraction of the form.

The ratio is expressed in the form: $\left\{ c \div d \right\} k$

Where k is some base as already defined, and both c and d refer to the frequency of occurrence of some event or item. In the case or a ratio, as opposed to a rate, the numerator is not a component part of the denominator.

Trend: It is defined as a measurable change that occurs in behaviour developing among a population of individuals.

VITAL HEALTH STATISTICS

It is the analysis of the data that provides a description of the vital or very important events occurring in any community. Vital events consist of various events occurring during the lifetime of an individual. These include birth, death, sickness, marriage, divorce, adoption, etc.

Vital statistics is somewhat simílar to census, where there is a huge difference between the two terms. Census provides a picture and characteristics of the population at one moment of time, whereas vítal statistics provides information about the chang**es occurring** continuously in the instantaneous picture.

Definitions

Vital statistics are conventionally numerical records of marriage, births, sickness and deaths by which the health and growth of a community may be studied.

-B. Benjamin

Vital statistics is that branch of biometry which deals with data and laws of human mortality, morbidity, and demography. -**Arthur Newshobne**

Uses of Vital Statistics

Vital statistics is very much useful for the development of nation, as major planning is dependent on these statistics. These are important for an individual, for various agencies, for research purposes, and for administration.

1. **Uses for individuals**: Different types of records like births, death, marriages, and divorce etc. are very much important to an individual.

2. Uses for operating agencies: Government agencies need records of birth, deaths, and marriages to prepare various types of plan. These agencies make various policies on the basis of the records.

3. Uses in research: Vital statistics play a pivotal role in demographic research. It is again helpful in making various plans and policies for the well-being of the society.

4. Uses in public administration: Statistics and vital statistics play an important role in public administration, which are basic methods underlying all official programmes of economic and social development.

5. International uses of vital statistics: Vital statistics are used by international agencies to study the population of a particular area. For example, WHO can use the vital statistics of various countries to see the population variations in terms of birth rate, death rate, or any other vital event.

Methods of Obtaining Vital Statistics

There are various methods for obtaining vital statistics. There is registration method, census enumeration, and analytical method, as discussed below:

1. **Registration method**: It can be defined as the compulsory recording of occurrence and characteristics of vital events; primarily these are used as legal documents and secondarily these are used as a source of statistics. Continuous permanent recording of vital events can be done properly by the use of legislation, which makes registration compulsory.

2. **Census enumeration**: Census recording is done after every 10 years in most of the countries. A census can be defined as an enumeration at a specified time of individual's inhabiting a specified area, during which particulars are collected regarding age, sex, marital status, occupation, religion, etc.

3. **Analytical method (estimation of vital rates using census data)**: This is a mathematical method based on the analysis of two consecutive censuses of population.

Measures of Vital Statistics

The measures of vital statistics may be classified in following three broad categories

MEASURES OF FERTILITY

The various measures of fertility are crude birth rate, specific fertility rate, general fertility rate, and total fertility rate.

1. **Crude birth rate**: It is the simplest method of measuring fertility. It acts as an index of the Relative speed at which additions are being made to the population through childbirth.

$$\text{Crude birth rate} = \frac{\text{Total number of live births during a year}}{\text{Total population as of July 1}} \times 1000$$

Specific fertility rate: Specific fertility arises from the fact that fertility depends upon many factors like age, marriage, etc. If we calculate the fertility rate on the basis of age, it is called as specific fertility rate (SFR).

$$\text{Specific fertility rate} = \frac{\text{Number of live births in a specific age group of female a given area in 1 year}}{\text{Mid-year female population of the specified age group in the given area in 1 year}} \times 1000$$

3. **General fertility rate (GFR):** It is defined as number of children born per 1000 females in a childbearing or reproductive age group.

$$\text{GFR} = \frac{\text{Number of live births in a given area in 1 year}}{\text{Mid-year female population of age 15-49 in that area}} \times 1000$$

4. **Total fertility rate**: To calculate the total fertility rate (TFR) we need to measure specific fertility rates and then add them. So, the total fertility rate is sum of age-specific fertility rates from a given age to the last point of childbearing age of a female.

11

We can calculate the specific fertility rate for group 'x' years and under (x + 5) as:-

$$\frac{\text{Annual births to females aged x and under}(x + 5)}{\text{Mean number of females aged x and under }(x + 5)} \times 1000$$

Total fertility rate = 2SFR × t

t = Magnitude of the age.

Reproduction rates: The reproduction rates are of two types as discussed below.

Gross reproduction rate (GRR): Gross reproduction rate assures the rate at which a newborn female would add to the total female population, if they will be alive and experience the age-specific fertility rate till the end of the childbearing period.

$$\frac{\text{Number of female births}}{\text{Total number of births}} \times \text{Total fertility rate}$$

Net reproduction rate: Net reproduction rate considers both fertility and mortality rate. So, it can be calculated as multiplying the female-specific fertility rate of each age by the population of female survivors to that age and adding up the products.

$$\frac{\text{Number of female births x survival rate}}{100}$$

MEASURES OF MORTALITY

The various measures of fertility are crude death rate, specific death rate, standardized death rate, infant mortality rate, neonatal mortality rate, maternal mortality rate, are as follows:

1. **Crude death rate:** It is defined as number of deaths occurring in a given population over 1 year.

Annual crude death rate =

$$= \frac{\text{Annual deaths}}{\text{Annual mean population}} \times 1000$$

2. **Specific death rate:** It is defined as the number of deaths occurring in a given population in a specific age group over 1 year.

$$\frac{\text{Number of deaths in a specific age group in a given area}}{\text{Mid-year population of that age group in the Given area}} \times 1000$$

3. **Standard death rates**: We can obtain standardized death rates by two means, i.e. direct method and indirect method.

Direct method: It gives the death rate of some standard population if it had the mortality of the given community or the death rate that occurred in the community if its population was distributed as that of the standard.

Indirect method: In this method the crude death rate is adjusted by applying a faster measurement of relative mortality proneness of the population of the community.

4. **Infant mortality rate:** It is defined as number of deaths under 1 year of age among a give population in a particular area.

$$\frac{\text{Number of deaths under 1 year of age}}{\text{Number of live births}} \times 1000$$

5. **Neonatal mortality rate**: It is defined as number of deaths under 1 month of age (28 day among a given population in a particular area.

$$\frac{\text{Number of deaths under 1 month}}{\text{Number of live births}} \times 1000$$

6. **Fetal death rate:** It is defined as

$$\frac{\text{Total number of fetal death during a year}}{\text{Total deliveries during a year}} \times k$$

7. **Maternal mortality rate:** It is defined as number of deaths occurring due to puerperal caused among a given population

$$\frac{\text{Deaths from all puerperal causes during a year}}{\text{Total live births during the year}} \times k$$

k is taken as 100 or 1000; however, it is more convenient to use k = 100,000 to determine maternal mortality rate.

8. **Perrinatal mortality rate:** Since fetal deaths occurring late in pregnancy and neonatal deaths frequently have the same underlying causes, it has been suggested that the two be combined to obtain what is known as the perinatal mortality rate. This rate is computed as:

$$\frac{\text{(Number of fetal deaths of 28 weeks or more)} + \text{(Infants deaths under 7 days)}}{\text{(Number of fetal dea ths of 28 weeks or more)} + \text{(Number of live births)}} \times 1000$$

9. **Fetal death ratio**: This ratio is defined as

$$\frac{\text{Total number of fetal deaths during a year}}{\text{Total number of live births during the year}} \times k$$

where k is taken as 100 or 1000.

10. **Cause of death ratio**: This ratio is defined as

$$\frac{\text{Number of deaths due to a specific disease during a year}}{\text{Total number of deaths due to all ca uses during the year}} \times 100$$

This index is used to measure the relative importance of a given cause of death. It should be used with caution in comparing one community with another. A higher cause- of-death ratio in one community than that in another may be because the first community has a low mortality from other cause.

11. **Proportional mortality ratio**: This index has been suggested as a single measure for comparing the overall health conditions of different communities. The specific class is usually an age group such as 50 years and over, or a cause of death category such as accident. It is defined as:

$$\frac{\text{Number of deaths in a particular subgroup}}{\text{Total number of deaths}} \times 100$$

MEASURES OF MORBIDITY

Morbidity, illness, sickness, or morbid condition is defined as any deviation from a state of physical or mental well-being. For a given population morbidity can be measured in terms new cases, i.e. incidence or in terms of combination of new and old cases, i.e. prevalence.

I. **Morbidity incidence rates**: Incidence rates give us information on often an event occurs in a population over a period of time such as a week, a month, a year, etc.

$$\frac{\text{Total number of new persons of illness during a given period of time}}{\text{Population exposed to risk in that time}} \times 1000$$

II. **Morbidity prevalence rates:** It is used to find out how common an event in a population is. It is basically of two types.
- Period prevalence
- Point prevalence

Period prevalence=

$$\frac{\text{Total number of new and old cases found during a specified period}}{\text{Population exposed to risk at the same period}} \times 1000$$

Period prevalence =

$$\frac{\text{Total number of new and old cases found at a particular point of time}}{\text{Population exposed to risk at the same point of time}} \times 1000$$

III. **Duration of illness**: It can be calculated as

(a) Days of illness per person is exposed to

$$\frac{\text{Total days of illness of all sick persons}}{\text{Total population exposed}}$$

(b) Days per illness per ill person:

$$\frac{\text{Total days of illness of all sick persons}}{\text{Number of persons who fell ill}}$$

(c) Days of illness per spell:

$$\frac{\text{Total days of illness of all sick persons}}{\text{Number of spells of illness}}$$

IV. **Case-fatality ratio**: This ratio is useful in determining how well the treatment program for a certain disease is successful. It is defined as

$$\frac{\text{Total number of deaths due to a disease}}{\text{Total number of cases due to the disease}} \times 100$$

The period of time covered is arbitrary depending on the nature of the disease and may cover several years for an endemic disease. Note that this ratio can be interpreted as the probability of dying following contraction of the disease in question and as such reveals the seriousness of the disease.

Immaturity ratio: This ratio is defined as

$$\frac{\text{Number of live births under 2500 grams during a year}}{\text{Total number of live births during the year}} \times 100$$

V. **Secondary attack rate:** This rate measures the occurrence of a number of a contagious disease among susceptible persons who have been exposed to a primary case and is defined as

Number of additional cases among contacts of a primary

case within the maximum incubation period   x 100

_____

Total number of susceptible contacts

This rate is used to measure the spread of infection and is usually applied to a closed group such as a household or classroom, where it can reasonably be assumed that all members were, indeed, contacts.

STATISTICAL PACKAGES AND APPLICATIONS

Traditionally, in the field of healthcare biostatistics, data has been handled manually; it involves laborious work, needing lots of paperwork. The results often become available too late and contain different types of errors. The advent of computers and their popularity in almost every sphere of life have significantly changed the way biostatistics is now applied to health sciences. Computer hardware and appropriate software have optimized the various Processes involved in the application of biostatistics by making the data analysis and storage easier, faster, and errorless. On the other hand, it has also improved the comprehension of data through the use of sophisticated graphic presentation techniques.

Several computer packages are available to be used for the statistical analysis: among them commonly used packages are Microsoft Excel, SPSS, SAS, Minitab, Stata, Systat, NCSS, etc.

- ❖ **Microsoft Excel:** Microsoft Excel is a very popular and useful spreadsheet program that can be used for data entry. It has the capacity to generate random numbers and can be used for the computation of many standard statistical applications like computation of mean, range, standard deviation,

etc. It can also be used for relatively sophisticated work like computing regression and ANOVA.

❖ **SPSS:** Originally known as Statistical Package for Social Sciences, SPSS was developed in 1960 at Stanford University to help solve problems in the social sciences. SPSS now stands for Statistical Product and Services Solutions and is among the most comprehensive and popular statistical packages. In addition to its provisioning for the standard procedures needed in descriptive statistics, it can also be provisioned for regression analysis and analysis of variance (ANOVA). It is also capable of multivariate analysis involving sophisticated techniques, namely cluster analysis, time series analysis, etc. The product is also capable of determining sample size and estimating power of a statistical procedure. To know more about SPSS, visit the website www.spss.com.

❖ **SAS:** The Statistical Analysis System (SAS) is very comprehensive software developed by North Carolina State University. This software is divided into many modules and its licensing is flexible, based upon the need for functions. This system contains a very large variety of statistical methods and is the software of choice for much major business including the entire pharmaceutical industry. SAS has also developed a PC SAS, which compatible with the personal computer and has a user-friendly Windows interface. more useful components of SAS are Base SAS, SAS/STAT, SAS/GRAPH, etc. An advan tage of SAS is its capability to transport data files in various formats and convert the SAS data set without much effort. To learn more visit their website www.asa.com

❖ **Minitab:** This is another statistical package designed to facilitate the teaching of statistical methods by using computer. The Minitab is also very user-friendly product with designed documentation facilities being widely used in educational institutions. To learn more about minitab visit their website www.minitab.com.