QSF-3

~~Date - 18.03.22~~
Date - 18.05.22

~~Municipal Security, Stations~~

Basic concepts of the terms "Health" and "Morbidity" and other Related Terms

Health is a state of complete physical mental, and social well being and not merely the absence of disease or infirmity". The term 'morbidity" or 'sickness" means departure from the above ideal healthy condition.

Disease and Injury

It should be noted that a "morbid condition" appears due to an attack of one or more "diseases" or "injury". The term disease denotes a disturbance in the normal physical and/or mental health. This disturbance may not be apparent but discovered only after special investigation. The term injury is a condition produced by an external cause such as violence, accidents poisoning and misadventure.

Need for and Importance of Morbidity Statistics

The need for and importance of morbidity statistics lies in general in providing information on the following:

a) How many people suffer from particular diseases, how often and for how long?

b) What types of demands these diseases make on the medical and public resources; and what financial loss they cause?

c) How fatal the different diseases are?

d) To what extent are people prevented by these diseases from carrying out their normal activities?

e) To what extent the diseases concentrated in specific group of the population. e.g. according to age, sex, marital status, ethnic group, occupation or place of residence. etc?

f) How far do the above factors vary over time?

g) What is the effect of medical care and health services on the control of disease incidence?

Sources of Morbidity Statistics

The availability of morbidity statistics is mainly restricted to the developed countries. However, among the various sources of morbidity statistics which are available at present in different countries, the following are of importance:

i) Notifiable disease records.

ii) Hospital records of in-patients together with records of attendance of out patients in hospitals, clinics and dispensaries.

iii) Cases seen by the staff of Primary Health Care centres, clinics, school medical services by factory medical staff and other medical staff responsible for special groups of population.

iv) Special sickness surveys covering the whole population or samples of it.

v) Statistics on illness collected in connection with life or sickness insurance.

vi) Medical care statistics, that is, statistics collected under community medical care programmes.

vii) National or local registers on cancer, tuberculosis.

viii) Statistics of causes of death in relation to knowledge of duration of illness, if recorded.

ix) Records from Industrial sickness benefit association.

x) Records from recruitment to the armed forces on goal records.

xi) Screening survey and supervision of some diseases e.g. tuberculosis, sexual diseases and HIV/AIDS infections.

xii) Absenteeism and sickness records in educational institutions, civil service examinations and industrial concerns.

xiii) National family Health Surveys phase I and II conducted in all Indian states and UTs during 1992-93 and 1998-99 also provide limited but valuable information on morbidity statistics.

Uses of Morbidity Statistics

Given below are some specific areas having a very good potential to use morbidity statistics at different levels.

i) Required for the purposes of medical administration.

ii) To indicate the pop^{n}s requirements for medical care at home and hospital, etc.

iii) Required by public health agencies for the control of diseases and epidemics, and for the location, design

and administration of public health and medical facilities and services, including rehabilitation programmes, and for estimating health manpower needs.

iv) Required by various industrial sectors, as sickness prevents employees associated with various industries to carry out their normal duties: the armed forces has a similar concern with respect to its manpower resources.

v) Pharmaceutical houses and manufacturers of medical, surgical and hospital supplies assess their markets from morbidity statistics.

vi) Many research activities are guided by morbidity statistics, particularly when the data are sufficiently detailed to establish causal relationship during the course of the disease.

In deciding upon the appropriate rates of morbidity, the illness that exist in a population during a given time period may first be classified as follows:

i) Illness beginning during the period and ending during the period.

ii) Illnesses beginning during the period and still existing at the end of the period.

iii) Illnesses existing before the beginning of the period and ending during the period.

iv) Illnesses existing before the beginning of the period and existing at the end of the period.

Incidence Rate

The crude incidence rate (C.I.R.) is defined by the number of new cases if illness or injury (I) in a community on an annual basis per 1,000 average number of persons (P) living in the community during the reference period.

Symbolically, C.I.R. may be defined as follows:

$$C.I.R = \frac{I}{P} \times 1000$$

When a population is exposed to the risk of a disease for a limited period of time only, the incidence rate may be termed as, the 'attack rate'.

In order to study the spread of infection in familial exposure, a morbidity investigation may be conducted among families brought under observation as a result of the presence of the diseases in one of its members: this member may be termed as the "index person". The number of new illnesses with the same disease may then be related to the number of persons exposed, excluding the index in each instance.

Thus,

$$\text{Secondary attack rate} = \frac{\text{The number of new illnesses with the same disease of the index person}}{\text{Number of persons exposed excluding the index person}} \times 100$$

Prevalence Rate: Prevalence of a disease indicates how common the disease is in a population. It is of two types: Period prevalence rate and Point prevalence rate.

Period prevalence rate: It is used to measure the frequency of illness in existence during a defined period. It includes all the cases in the defined period old and new cases occurring during the same period.

The crude period prevalence (C.P.R) is defined as the number of illness (e) existing at any time within a specified period per 1000 average number of persons (P) exposed to risk during that period.

Point prevalence rate: The term is used to measure the number of cases of illness, new and old, existing at a particular point of time, such as 10.00 a.m. on Monday, the 25th October, 1999 and related to the number of persons exposed to the risk at that point of time. More specifically, the numerator includes all cases (spell of illness on persons affected by the disease) at the given moment, irrespective of the length of time which has elapsed from the beginning of the illness to the time when point prevalence is measured. The denominator is the total persons within which the disease is ascertained.

Duration of illness: There is another measure for morbidity which is calculated as follows:

a) Average duration of illness per person = $\frac{\text{Total days of illness of all sick persons}}{\text{Total population exposed}}$

b) Average duration of illness per person sick = $\frac{\text{Total days of illness of all sick persons}}{\text{Number of persons fell ill}}$

c) Average duration of illness per spell = $\frac{\text{Total days of illness of all sick persons}}{\text{Number of spell of illness}}$

Case Fatality Rate (C.F.R): If I_i is the number of new cases of illness or injury for cause (i) during a specified period, and D_i is the number of deaths from cause (i) within the same period and community, then the case-fatality rate (C.F.R) is defined by

$$C.F.R = \frac{D_i}{I_i} \times 1000$$

The C.F.R is intended to measure the risk of death from a specific condition among those suffering from it

Cancer cure rate = $\frac{\text{No. of cases still alive and without evidence of the disease five years after treatment}}{\text{No. of cases treated during the reference period under study.}}$

Interrelationship between measures — Date - 23/05/22

Let, P and I denote the prevalance and incidence rate respectively of a disease in a community and δ represent the average duration of the disease from the onset to termination, measured in the same time units used to specify the incedence. It is found that P is directly proportional to the product I and δ.

Since, P varies proportionally to the product I and δ, a change in the prevalance rate on disease from the period to another may arise due to the changes in incidence or duration or both. In an ideal circumstances where Incidence

and duration present to normal cont over time, the disease is said to be stable and then the prevelance (P.) equals to the product of incedence (I) and duration (D). Symbolically ~~P equals to~~

$$P = I \times D$$

Another relation that exists if the disease is stable on nearly 50 - is between incidence (I) and mortality (M) from which the case fatality rate (F) may be estimated -

$$F = \frac{M}{I}$$

Intraclass Correlation

Date - 25/05/22

$$\sigma_x^2 = \sigma_y^2 = \frac{\Sigma(k_i - 1)\Sigma(x_{ij} - \bar{x})^2}{N}$$

$$cov(x_i, x_j) = \frac{1}{N}\sum_i \left\{ \sum_{j \neq l} (x_{ij} - \bar{x})(x_{il} - \bar{x}) \right\}$$

$$= \frac{1}{N}\sum_i \left\{ \sum_{j=1}^{k_i} \sum_{l=1}^{k_i} (x_{ij} - \bar{x})(x_{il} - \bar{x}) - \sum_{j=1}^{k_i} (x_{ij} - \bar{x})^2 \right\}$$

$$= \sum_i \left\{ \sum_{j=l=1}^{k_i} \sum^{k_i} (x_{ij} - \bar{x})(x_{il} - \bar{x}) \right\} -$$

$$= \sum_i \left\{ \sum_{j=1}^{k_i} (x_{ij} - \bar{x}) \sum_{l=1}^{k_i} (x_{il} - \bar{x}) \right\} -$$

$$= \sum_i \left\{ k_i(\bar{x}_i - \bar{x})\, k_i(\bar{x}_i - \bar{x}) \right\} -$$

$$cov(x_i, x_j) = \frac{1}{N}\left\{ \Sigma k_i^2 (\bar{x}_i - \bar{x})^2 - \sum_i \sum_j (x_{ij} - \bar{x})^2 \right\}$$

$$r = \frac{\Sigma k_i^2 (\bar{x}_i - \bar{x})^2 - \Sigma\Sigma (x_{ij} - \bar{x})^2}{\sum_i (k_i - 1) \sum_j (x_{ij} - \bar{x})^2}$$

$k_i = k$

$$r = \frac{\Sigma k^2 (\bar{x}_i - \bar{x})^2 - \sum_i \sum_j (x_{ij} - \bar{x})^2}{\sum_i \sum_j (k-1)(x_{ij} - \bar{x})^2}$$

$$= \frac{k^2 \Sigma (\bar{x}_i - \bar{x})^2 - \Sigma\Sigma (x_{ij} - \bar{x})^2}{(k-1) \Sigma\Sigma (x_{ij} - \bar{x})^2}$$

Let, $\sigma_m^2 = \frac{1}{n}\sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2$

$n\sigma_m^2 = \sum (\bar{x}_i - \bar{x})^2$

$\sigma^2 = \frac{1}{nk} \sum_i \sum_j (x_{ij} - \bar{x})^2$

i. $nk\sigma^2 = \sum\sum (x_{ij} - \bar{x})^2$

$r = \frac{nk^2 \sigma_m^2 - nk\sigma^2}{(k-1)\, nk\sigma^2}$

$= \frac{nk^2 \sigma_m^2}{(k-1)nk\sigma^2} - \frac{1}{k-1}$

$= \frac{1}{k-1} \left\{ \frac{k\sigma_m^2}{\sigma^2} - 1 \right\}$

$= $

$r = \frac{1}{k-1} \left\{ \frac{k\sigma_m^2}{\sigma^2} - 1 \right\}$

i. $r(k-1) + 1 = \frac{k\sigma_m^2}{\sigma^2} \geq 0$

$r(k-1) + 1 \geq 0$

i. $r \geq -\frac{1}{k-1}$ ——— (i)

$1 + r(k-1) \leq k$

$r(k-1) \leq k - 1$

$r \leq 1$ ——— (ii)

$$\boxed{-\frac{1}{k-1} \leq r \leq 1}$$

Q. In 4 families each containing 8 persons, the chest measurements of persons are given below. Calculate the intraclass correlation coefficient.

Family 1	Family 2	Family 3	Family 4
43	33	56	34
46	34	52	37
48	37	50	38
42	39	51	40
50	42	54	40
45	35	52	41
45	37	39	44
49	41	52	44

CLINICAL RESEARCH & CLINICAL TRIALS-I

Define Clinical Research:

Clinical research is a branch of healthcare science that determines the safety and effectiveness of medications, devices, diagnostic products and treatment, diagnosis or for relieving symptoms of a disease. Clinical research is different from clinical practice. In clinical practice established treatments are used, while in clinical research evidence is collected to establish a treatment.

Define Clinical Trials:

Clinical trials experiments or observations done in clinical research. Such prospective biomedical or behavioral research studies on human participants are designed to answer specific questions about biomedical or behavioral interventions, including new treatments and known interventions that warrant further study and comparison. Clinical trials generate data on dosage, safety and efficacy. They are conducted only after they have received health authority/ethics committee approval in the country where approval of the therapy is sought. These authorities are responsible for vetting the risk/benefit ratio of the trial - their approval does not mean the therapy is 'safe' or effective, only that the trial may be conducted.

Why we need the clinical trials?

Clinical trials show us what works in medicine and health care. They are the best way to learn what works in treating diseases like cancer. Clinical trials are designed to answer some important questions:

- Does the new treatment work in people? If it does, doctors will also look at how well it works. Is it better than treatment now being used? If

it's not better, is it as good and cause fewer side effects? Or does it work in some people who aren't helped by current treatments?

- Is the new treatment safe? No treatment or procedure - even one already in common use - is without risk. But do the benefits of the new treatment outweight the risks.
- Is this treatment better than the standard treatment given for this disease? Clinical trials help show if a new drug or treatment, or a new treatment combination, works better than what is now used.

Phases of Clinical Trials:

The different phases of clinical trials are:

Phase 1: The purpose of phase 1 is to ensure that the treatment is safe in human and to determin how and where it distributes within the body. This testing normally takes place with a small group of healthy volunteers. At the end of phase 1, the results are collected, analyzed, and submitted to the authority for permission to proceed to phase 2 Clinical Trials. However, if the results show that the treatment was associated with one or more serious adverse events, then the authority may not give ~~pee~~ permission to proceed to phase 2.

Phase 2: The purpose of a phase 2 clinical Trial is to determine the right dosage and effectiveness in treating that particular disease. This testing normally takes place with a larger number of volunteers who have the disease. There are many different ways that a trial sponsor can conduct their trial, but the plan normally involves assigning participants to different treatment groups, where ~~each~~ group can receive different doses or delivery of the treatment. Normally, there is a "control group" that receives either the current

standard of care, if another type of treatment is already available on the market for the disease, on a "placebo" treatment, such as a sugar pill or harmless injection that does not contain the treatment. The health of the group(s) of patients who received the different types of treatment is compared to the control groups. However, if the results show that the treatment did not work better than the current standard of care or even caused acceleration of the disease or other unexpected serious adverse events, the authority may not give permission to proceed to phase 3.

Phase 3: A phase 3 clinical Trial involves a much larger group of volunteers and primarily focuses on determining whether the treatment would be safe and effective for a wide variety of people. The plan normally involves assigning participants to treatment or control groups. There can be more than one treatment group, especially if the treatment involves a combination of drugs or different components. Again, there is a control group that receives either the current standard of care regimen or a placebo treatment. After completion of phase 3 clinical Trials, the health of the patients who received the different types of treatment are compared to the control groups. If the results show that the treatment did not work better than the current standard of care or even caused acceleration of the disease or other unexpected serious adverse events, the authority may not give permission to proceed to apply for a New Drug Application (NDA).

Phase 4: After approval by the authority and manufacturing of the drug on a large-scale by the sponsor, the process enters what is called phase 4 clinical Trial/Post-market surveillance/Report Adverse Events. For at least the entire time a treatment* is on the market, the authority monitors for public safety and potentially serious adverse events.

Kappa Statistics

Date - 1/6/22

The reliability of clinicals ratings is an important consideration is areas such as diagnosis and the interpretation of examination find age. Often these ratings lie on a nominal on and ordinal scale. For such data the kappa coefficient is an appropriate measure of reliability.

A common example of a situation in which a researchan may want 2 asses agreement on a nominal scale is to determine the presence on absence of some desease on condition. These agreement could be determine in situation in which two researchers on clinical have used. The same examination tool on different tools to determine the diagnosis. One way of gauging the agreement between two clinicals is to calculate over all pensentage of agreement on effective pensentage of agreement. All though these calculations provide a measure of agreements, nither takes into account that would be expected purely by chance. If ~~clinicators~~ clinician, agree purely by chance, they are not really "agreeing" at all:

Only agreement beyond that expected by chance can be considered "True" agreement. Kappa is such a measure of true agreement. It indicates the proportion of agreement beyond that expected by chance, ie., the achived beyond chance agreement as a proportion of the possible beyond chance agreement. It takes the form

$$K = \frac{\text{Observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}}$$

In terms of simble this is,

$$K = \frac{P_o - P_e}{1 - P_e}$$

, where, P_o is the proportion of observed agreements and P_e is the proportion agreement of chance

		Clinician 2		
		Patient	not patient	
Clinician 1	Patient	22 $\frac{25\times24}{39}$ (16)	2	24
	not patient	4	11	15
		26	13 $\frac{13\times15}{39}$ (5)	39

$P_e = \frac{16+5}{39}$, $P_o = \frac{33}{39}$, $K = \frac{0.84-0.53}{1-0.53}$

$= 0.53$ $= 0.84$ $= 0.45$

The simplest use of kappa is for the situation in which two clinicians each provide a single rating of the same patient, or where a clinician provides two ratings of the same patient, representing inter-rater and intra-rater reliability respectively. Kappa also can be adopted for more than one rating per patient from each of two clinicians or for situations where more than two clinician rate each patient or where each clinician may not rate every patient [the range of possible values of kappa is from -1 to +1. Though it usually falls between 0 and 1] Unity represents perfect agreement, indicating that the rate w agree in their classification of every case. Zero indicates agreement no better than that expected by chance, as if the raters have simply guessd every rating. A -ve kappa would indicate agreement was then that expected by chance. However these rarely occurs in clinical contexts.

[Landis and Koch have proposed the following as standards for strength of agreement for the kappa coefficient.

≤0 : poor

0.012 - 0.20 : slight

0.21 - 0.40 : fair

0.41 - 0.60 : moderate

0.61 – 0.80 : substancial
0.81 – 1 : Almost perfect]

Screening — Date – 06/06/22

The seaneh for unnecognize desease on defect by means of napidly apply test, examinations or othen pnocedunes in appenently healthy individuals is known as seneening.

Seneening defens from Penlodic health examination in the following nespects —

i) Capable of wide application
ii) Relatively inexpensive
iii) Requines little physician time. Infect the physician is not nequined to administen the test, but only to intenpnet.

Diffenence between Seneening and Diagnosis

A seneening test is not indended to be a diagnostic test. It is only an finicial examinations. Those who one found to have +ve test nesults one nefened to a physician for funthen diagnostic workup and treatment.

Screening	Diagnostic test
a) Done on apparently healthy.	a) Done on those with indication or sick.
b) Applied to groups.	b) Applied to single patients, all diseases are considered.
c) Test results are arbitrary and final.	c) Diagnosis is not final, but modified in light of new avidance, diagnosis is the sum of all evidence.
d) Based on one criteria or cut-of point.	d) Based on evaluation of a no. of symptoms, sings and laboratory findings.
e) Less aquarate.	e) More aquarate.
f) Less expensive.	f) More expensive.
g) Not a basis for treatment.	g) Used as a basis for treatment.
h) The initiative comes from the investigator or agency providing care.	h) The initiative comes from a patient with a complane.

Uses of Screening :

Four main uses have been described.

a) Case detection – These is also known as prescriptive screening. It is defined as the presumptive identification of unrecognize desease, which doesn't arrise from a patients request.

b) Control of disease – These is also known as prospective screening. People are examined for the benifit of others. For example. screening of immegrants from infectious diseases such as tuberculosis to protect the home population.

c) Research purpose, d) Educational Opurtunities

c) Research Purpuses - Screening may sometimes be perform research purpose. For example there are many chronic diseases whose natural history is not fully known (like on cancer, hypertention). Screening may aid in obtaining more basic knowledge about the natural history of such diseases.

d) Educational Opuntunities - Screening programs provide opuntunities for creating public awarness and for educating health professionally.

Types of Screening :-

Three types of screening have been described -

a) Mass Screening - mass screening simply means the screening of a whole prepation or a subgroup. It is offered to all, inrespective of the particular risk individual may run of contracting the disease in ponion.

b) High risk or selective Screening - Screening will be most productive if applied selectively to high risk groups, the groups defined on the basis of epidemiological research. For ex. since measels occurs more often in children then the adult, screening for misels in the the children would increase the yield of new cases.

c) Multifacle Screening - It has been define as the application of two or more screening test in combination to a large no. of people at one time then to canny out separate screening tests for single diseases. The procedure may also include a health question yeair, clinical examination and a range of measurements and investigations.

Evaluation of a Screening Test :-

The following measures are used to evaluate a screening test.

1) Sensitivity - The term sensitivity was introduced by Yerushalmy in 1940's as a statistical index of diagnostic aquaracy. It has been defined as the ability of a test to identify correctly all those who have the disease. A 90% sensitivity means that 90% of the disease people screan by the test will give a true +ve results and the remaining 10% a false -ve result. Sensitivity is the prob. that the test will indecate the disease will among those with the disease.

	Diagnosis		
	Diseased	Not diseased	
Screening +ve	a (True)	b (false)	(a+b)
Screening -ve	c (false)	d true)	(c+d)
	(a+c)	(b+d)	a+b+c+d = N

$$\left\{ \frac{a}{a+c} \times 100 \right\}$$

2) Specificity - It is defined as the ability of a test two identify correctly those who donot have the disease, ie. true negative. A 90% of the specificity means that 90% of the non-disease persons will ~~have~~ give true negative result. 10% of non-disease people screan by the test will be wrongly classified as diseased when their not.

Predictive value of a +ve test:
It is defined as the percentage of ~~true +ve among the total +ve in screening~~ "diseased" identified by diagnosis among the ~~disease~~ positive (case) identified by screening.

Therefore,

Predictive value of a +ve test = $\frac{a}{a+b} \times 100$

Predictive value of a -ve Test : It is defined as the percentage of not diseased identified by diagnosis among the -ve identified by screening

∴ Predictive value of -ve test = $\frac{d}{c+d} \times 100$

① Calculate the prevelence of disease, sensitivity, specificity, positive predictive value, and -ve predictive value for the following data:

		Truth: Disease	Non-Disease	
Test Result	Positive	10	40	50
	Negative	5	45	50
		15	85	100

$\frac{10}{15} \times \frac{2}{100} = \frac{2}{3}$

+ve predictive value = $\frac{10}{15} \times 100$

$= \frac{2}{3}$

-ve predictive value = $\frac{45}{85} \times \frac{20}{100} = \frac{900}{17}$

Sensitivity = $\frac{10}{15} \times 100$

Specificity = $\frac{45}{85} \times 100$

prevalance = $\frac{15}{100} \times 100 = 15\%$

②

	20	33
	10	37

Unit-2 Date - 13/06/22

Epidemiology :-

The epidemiology is defined as the study of the distribution and determinants of health related states on events in specified popn and the application of this study to control of health problems.

Aims and uses of Epidemiology :-

Acording to the International Epidemilogical Association aims of epidemilogy are -

① To describe the distribution and magnitude of disease problem in human populations ② To identify aetiological factors (risk factors) in the pathogenesis of disease.

③ To provide the data essential to the planing, implimentation and evaluation of survices for the prevention control and treatment of disease and to the setting up of probities among those survices.

Uses of Epidemilogical :-

Uses of epidemilogy are -

① To study the natural history of disease, this ability, injury and death - It helps to study the rise and fall of diseases and changes in their character. Remember that the occurance of disease in the community is never stake. It keeps on changing over a period of time effective new suseptable popn. Epidemiology has played a major role in the understanding of the occurance of new are diseases and in the control of other diseases.

② To make a community diagnosis - refers to the identification of health problems, related factors and there quantification in the popn.

④ To asses risk factors of diseases in the popn - There are several contributing factors of occurance of diseases apart from causal agents. Sometime it may be difficult to identify the causal agents we may only come across associated with the disease in the community. Such factors are consider to be risk factors. As an example the risk factors of hypertension Obesity, family history of hypertension, uncontrol lifestyle, smoking, alcoholism, high salt intect, high intect of saturated facts etc.

⑤ To asses, evaluate and conduct research on the health programs - There are several heath programs lanched by the gov. with the aim of controling, preventing for giving treatment to the effected persons in the popn. It in nessesary understand the progress and effectiveness of these health programs for further modifications in the program, if needed, and also to understand how well do public health and health sunvices meets the probles and needs of the popn. (Epidemiological methods are applied in assessing and evaluating the outcome, effectiveness of health programes in the popn. Studies using epidemiological approaches are also used for converting research on the health problems and bringing out appropriate solutions to tackle the problems).

⑤ To complete a clinical picture - Understanding of the clinical sprectrum of diseases is often the result of reports of diseases which are presented in the hospitals or clinics. Why It is easier for the physician to study details of the disease in hospital setup, it is limited

by a large no. of diseases which go unreported to the doctor in the hospital. Epidemilogical approaches in the community health in overcoming this gap to a complete clinical picture of the disease.

⑤ To Identify Syndroms - Syndroms are a group of clinical features, characteristicly related to a particular disease. Syndrom identification is done by describing the distribution, association and disassociation of clinical phenomena in the population using appropriate emidemological method.

⑥ To search for causes of health and disease by studing the incidence in different pop^n groups, in terms of inheritance, behaviour and environment.

Epidemilogical models of causation of disease :- When we talk about the occurance of a disease, we need to understand that it involves certain causitive factors, effected person and the circumstances under which the disease occures. In an outbreak of the disease several factor often play a role. As an example when Kolera outbreak occures, there is a causitive factors ie. a type of bacteria called as Vibrio Cholerae : transmetting midium which is usually contaminated water or food items, and occasion where a large no. of people gather together, take foods or drinks from a common source which may be contaminated by the bacteria, the quantam of bacteria

ingested and the condition of immune system of the person effected.
These factors can be interrelated to each other and absence of one or more factors may not lead to the occurance of colera. When we talk about this factors we try to chatagonies these factors into 3 brought groups. Namely - ① Agent, ② Host. ③ Environment. In this example agent is vibrio colery, host is human being and environment factors are water supply and sanitation etc.

Interrelationship of Agent, Host and Environment

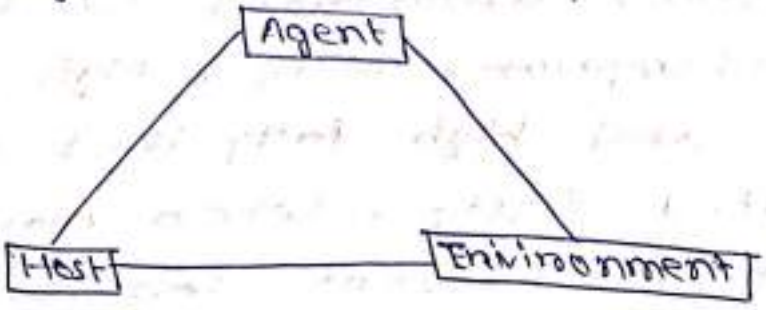
Agent - Agent is usually consider to be the causal factor. A single factor must be present for an infectios disease occur. It may be the microbes which may be baltenia, fungi, virus. In occupation related disease lead is the agent of on lead poisoning, asbestos for asbestosis etc. In other disease conditions disability, injury, death situation the agent may be chemical such as a poison, physical factor such as cold, head, radiation, nutritional deficiency, animal poison such as that of a snack etc. However all disease conditions can't be attributed to a single causal agent, such as ~~and~~ heart attack which has several factors ie family history, uncontroled lyfestyle, smoking, alcohol consumption, ovicity, high blood pressure and high fatty diet consisting of saturated facts. When more than a single cause is present for occurance a disease this is called multiple causation.

Host - The presence of an agent alone doesn't leave to disease condition. The agent should enter the body of another living be such as a human or animal to cause the disease or harbour the disease agent. Thus a host harbours the disease. The host may or may not get the disease. It again depends upon several factors such as immunity level, genetic makeup and the state of health.

Environment - The external conditions or surroundings of the human or animal favouring transmition of the disease agent are considerd to be environmental factors. In the case of Colera it may include a poor sanitation leading to contamination of the water supply with disease causing microbs, in case of dengu feaber, malenia feaber - Stagnant water ten mosquito bhiding and so on. The environment includes 3 components -

① Biological environment,
② social environment,
③ Physical environment.

The relationship between the agent, host and environment is shown in the form of a triangle.



Epidemilogy Triangle

Concepts of Disease Transmition :-

Fomites - Fomites are in animate objects which play a role in disease transmition. This could be toul, drinking glasses, pen, pencil, doon handles etc which can help in transmition of infpection by mean contaminated with infectious agent and then touched and claim in close contact with another person.

Vecton - A vecton is an anthropod such as mosquito, flight, nat etc which is living non-human carrier of disease that transports on surves the process of disease transmi-tion. The vecton carries the agent from an infected persons on animals through bite, body fluids, waste products on by contaminating foods and spreads to another person.

Reservoin - A reservoin is human, animal, plant, soil on in animate organic mater (food) in which the infectious organisams leave and multiply. Human being can surve both as a reservoin on a host.

Cannien - A cannien harvours, contain or spreed and infectious organisums. A person on animal may appear to be normal despide harbouring the infectious organism but he on it can transmit the infectious agent to another persons on animal. Through contact a tranimition media such as water, food etc and cause disease. Depending on the stage of disease state cannien can be classified as

① Incubetony cannien : A person who shades the infectious agent during eanly phase

of the disease when the infectious disease as enter the body and the body has not started showing syndroms and signs of the disease.

② Convalescent carrier :- And individual who has been exposed to a humous disease causing organism during the recovery phase of the course of the illness, but its still infectious it considered to be convalescent carrier.

③ Healthy carrier : An individual who has been exposed to and harboring a ~~cont~~ disease causing organism, but has not become it an has not shown any symptoms and signs of the illness at any point of time is called as a healthy carrier.

Modes of Disease Transmition :-

The modes of disease transmition can be either direct or indirect.

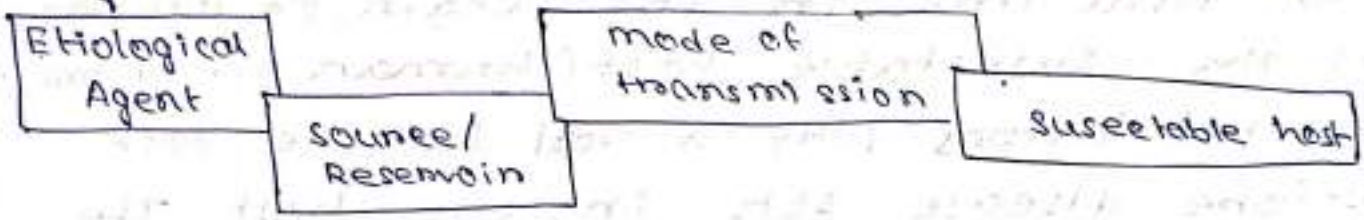
Direct Transmition - It is also refend to as person to person transmition. In this there is direct and imidiate transfer on the micro organism or the agent from a host or reservoir to a susceptible host. Direct transfer can happen through: ① direct contract, ② Droplets, ③ Contact with soil, ④ I.E. Einoculation through skin or mucosa, ⑤ Trans placental (through placenta from mother to the fetus).

Indirect Transmition - This is an condition when the disease causing agent are

transfer or carried by some intermidiate items, means, processes or organisms to the susectable host resulting in disease indirect transmition can occur through several means: ① Air borne transmition, ② water born transmition, ③ vehicle born transmition ④ Food born transmition, ⑤ vector born transmition.

Chain of Infection :-

Communicable disease transmision follows a medical modle shown in the following figure:



Chain of infection

This is tunned as chain of infection. Transmition of disease occurs when the infectious disease agent leaves the reservoir through a portal of exit, and is spread by one or more modes of transmition to the susceptable host. The infectious agent enters the host through a portal of entry. The etiological agent ~~of~~ / Infectious agent can be microbs such as bacteria, virus, worms, chemicals, other plant or animal related substances having the potential to cause diseases, disability or death. The source is the person, animal, object or substance from which and infectious agent passes or is disseminated to the host. The reservoirs are

mostly human, animal on non living things such as soil, food, decaing organic matter or substance where in the infectious agent lives, multiplies or reproduce. In hookworm infestation the source of infection is the soil, contaminated with infective larva and man is the reservoir of infection. Wants the aitiological agents needs the reservoir, it is past on to the host by mode of transmition either by direct (person to person) or indirect transmition. The final link in the chain of infection is the susceptable host (human or animal). several factors play a rool in the host before disease sets in the host. The host have defensive machanism against infactive agent in the form of deferent protection sceme of the body. If the infectious agents overcome these protective mechanism enter the body of the host, it is lightly that the host will fall in.

Definations Used in Disease Epidemiology :-

Endemic — The constant presents of a disease or infection within a specified geographic area or popn group is considered to be endemic for that perticulan area. For example upper respiratory infection, diarrhoeal disease are endemic in India.

Epidemic — When a large no. of persons are effected out of proportions to the routine occurance of diseases in a specified population then it is turned as epidemic. In other words epidemic is

'Occurance of unusual frequency disease above the endemic or expected frequency of occurance' in a specified popn place and time. For example epidemics of infections such as measeals, hepatytics, chicken pox, Colera etc. keeps on occuring of and on in specified populations.

Sporadic - When the disease occures in scatered popn inregularly hapazardly from time to time and generally infrequently then it is consider to be sporadic. There is no specific pattern in occurance on disease in the popn with respect to time and place. The cases are usually few in number.

Pendemic - An epidemic effecting large popn groups extending to a wide geographycal areas such as a nations, continant or the world is consider to be pendemic. For example is Corona in 2020, spanish flue 1918.

Described Epidemology

Described epidemology are usually the 1st phase of an epidemilogical investigation. This studies are concern with observing the distribution of disease on heath related characteristics of human popn. The data is collected about :

① personal characteristic such as age, genden. manital status, occupation, education, income, social class, dietecny pattern habits.

② Player dis^n of cases i.e. ares of high concentration. low concentration and spoting of cases in the map.

③ Time dis^n/ trends such as year season, month, week, data and hour of concept of the disease. Such information gives crues to possible associated factors such as age specific disease for crample-misels, diptheria in eanly child hood, canten; dietry patenn with odicity, seasonal variation, periodic fluctuation etc. The data collected are analysed and presented in terms of % at rates and ratios. These refers to statistical information of a problem and given in the table for use to learn and compuse. Thus descriptive epidemology provides information for —

① making community diagnosis i.e. describing the nature of diseases of problems and measuring these problems and terms of incidence/ prevelance rate / ratios / mortality rate etc. by

age, gender, occupation, social class etc.

(2) Providing the clues to etiology of disease for further vigorous investigation and confrontation of the causes.

(3) Planing, organising and implimentation of health care services to deals with this problems.

Analytical Epidemology:

The descriptive studies yield etiological clues for various disease which helps in making a guese or formulation of hypothesis for further vigorus study ~~education~~ testing for example - "Siganate smoking (10-20) in a day causes lunge cancer in 10-15% of smokers after 20 years of exposer." These type of hypothesis are further study and tested by analytical studies to determine the association of cause with the effect. The analytical studies are of two types.

(1) Case control study and

(2) Cohord study.

From each of these study desines one can determined:

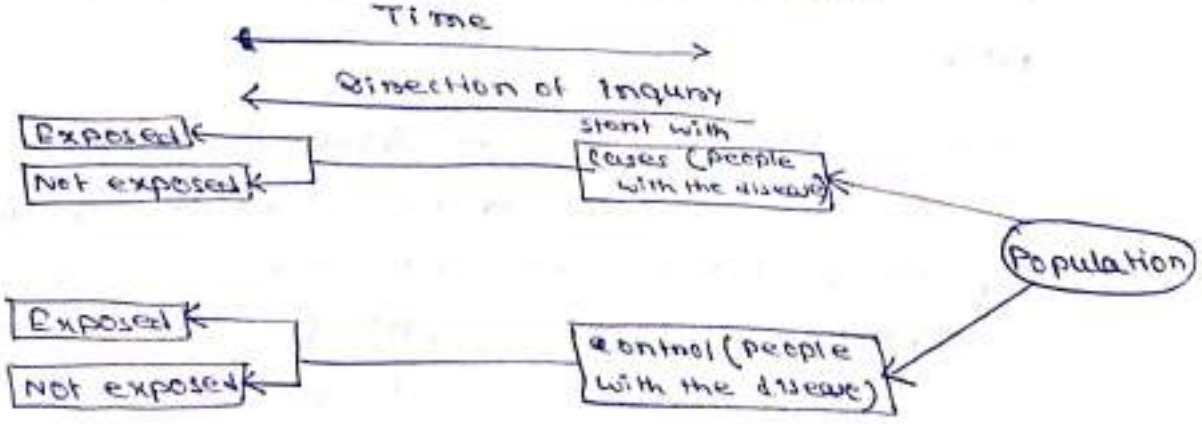
(1) A whether or not a statistical essociation exist between a disease and a suspected factor.

(2) If one exist, the strength of the association.

Case control study: In these method a group of people who have been diagnosed as having a particular problem (cases) are compared with a group of people who are similar in characteristics to that of cases but they are free from

the problem under study (controls). Here the approache use is retrospective i.e. the disease have already occurr and epidemiology goes that go back in time. He reviews the record interviews the cases and there family member. The data has collected about the suspected factors and analysed statistically to determine the extend of its association with the disease. These method therefore is called as retrospective method. These approach has held is identification of causative factors of many disease / problems etc. For example - Rubella in mothers during early pregnancy in the cause of congenital deformities in children, smoking associated with lunge cancer. Iodin deficiency associated with hypen thyroidism. These conclutions are based on repeated case control studies. Case control studies are easy to organise and are less expensive. Schematic of case control study is given below:

Design of a case control study —



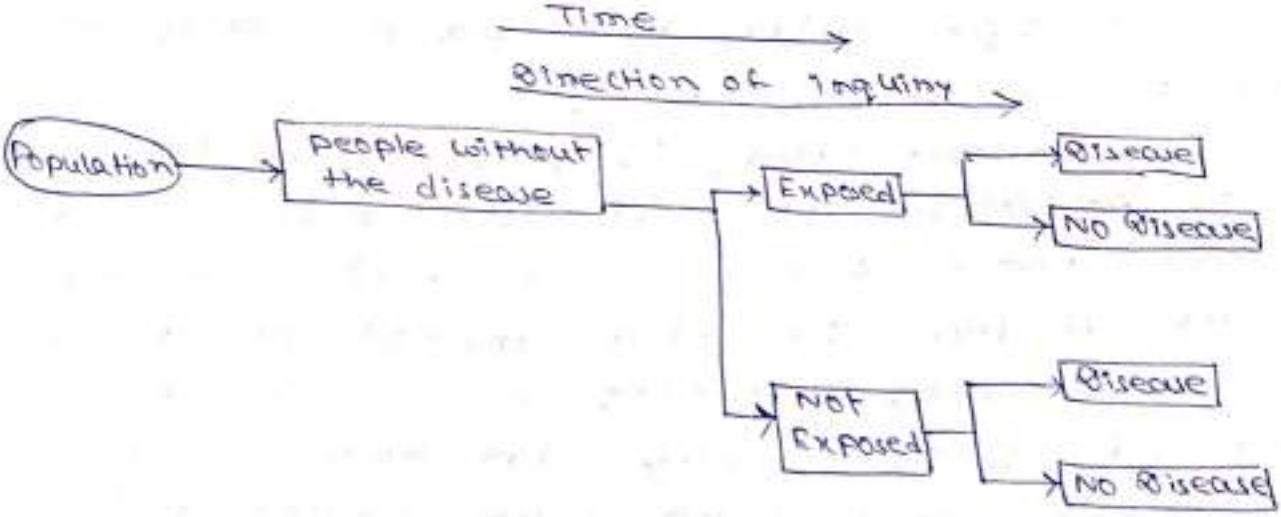
Cohort Study : A cohort is a specific group of people, at a certain time, sharing common characteristics or experience for example people bond on the same date or the same year (birth cohort), couple married in the same year (marriage cohort) a class of nursing students (experience cohort), people with the same occupation (occupational cohort) etc.

Cohort study is prospective in nature because the group under study is free from the disease but exposed to risk factor. In this method of study the epidemiology selects a cohort ie. a group of people say in the same age group and who are exposed to a certain risk factors say cigarette smoking (study group) and who are not exposed to the risk factors (control group). Both the groups are followed up for several years and observations are made with reference to frequency and distribution of the suspected disease (in these example lung cancer) over a period of time. The data is statistically analysed and comparisons are done between the incidence among smokers and non smokers to determine the association of risk factors to the disease. ~~Three type of cohort study~~ The prospective study is expensive and time consuming but it has its advantages over retrospective method. This method can health is studing the natural history of disease, estimating incidence rate, risk factor

under study to other disease or problem. For example cigarette smoking and high blood pressure, cardiovascular problems etc.

A schematic design of cohort studies is shown below:

Design of a cohort study —



Date- 06/07/22

Relative Risk :-

The estimation of disease risk associated exposen is obtain by an index known as Relative Risk or (RR) risk ratio. which is defined as the ratio between the incidence of disease among exposed person and incidence among non exposed. It is given by the formula. Relative Risk = Incidence among exposed / Incidence among non exposed.

Let us.

~~If~~ . consider the following hypothetical table:

		Disease		
		Yes	No	Total
Exposure	Yes	a	b	(a+b)
	No	c	d	(c+d)
	Total	(a+c)	(b+d)	N

$$R.R = \frac{(a/a+b)}{(c/c+d)}$$

Odds Ratio :-

From a case control study we can derive what is known as odds ratio (OR). which is a measure of the strength of the association between the risk factor and the outcome. Odds ratio is closely related to relative risk The derivation of odds ratio is based on 3 assumptions.

① The disease being investigated must be relatively rare.

② The cases must be representative of those with the disease.

③ The control must be representative of those without the disease.

Let us. consider the following hypothetical table :

		Disease		
		Yes	No	Total
Exposure	Yes	a	b	a+b
	No	c	d	c+d
	Total	(a+c)	(b+d)	

The odds ratio is defined by $\frac{(\frac{a}{b})}{(\frac{c}{d})} = \frac{ad}{bc}$

Attributable Risk :-

Attributable Risk (AR) : is the difference in incidence rate of disease (or death) between and exposed group and non exposed group. Attributable is often expressed as a person This is given by the formula of: Incidence of ~~disease rate among~~ exposed - incidence of disease rate among no exposed ".

~~rate among exposed x 100~~

$$\frac{\text{incidence of disease rate among the exposed} - \text{incidence of disease rate among no exposed}}{\text{incidence rate among exposed}} \times 100$$

Attributable risk indicates to what extend the disease under study can be attributed to the exposure.

$$\frac{\frac{33}{33+55} - \frac{2}{2+27}}{\frac{33}{88}} \times 100$$

a = 33
b = 55
c = 2
d = 27

= 81 %.

Population Attributable Risk :-

It is the incidence of the disease in the total population - the incidence of disease among those who were not exposed to the suspected causal factor. The concept of popn attributable risk is useful in that it provides an estimate of the amount by which the disease could be reduce in that population. If the suspected factor eliminated or ~~fac~~ modified. Let us consider a following hypothetical example: exposed to suspected factor & ~~Dis~~ (A) nonexposed to suspected causal factor (B) incidence of the total population (C) then popn AR $\times \frac{C-B}{C} \times 100$

Difference between case control study and cohort control study :-

Case control Study

① Proceeds from effect to cause.

② Stats with the disease.

③ Test wheather the supected cause occures more frequently in those with the disease then among those without the disease.

④ Usually the first approach to the testing of a hypothesis but also useful for exploratory studies.

⑤ Involves fewer no. of subjects.

⑥ Yields relatively quick results.

⑦ Suitable for the study for rare diseases.

⑧ Generally yields only estimate of RR and OR.

⑨ Can't yield information about disease other than that selected for study.

Cohort control Study

① Proceeds from cause to effect.

② Stats with people exposed to risk factor or suspected cause.

③ Test wheather disease occur more frequently in those exposed, then in those not similarly exposed.

④ Relative ~~of~~ testing of presigely formulated hypothesis.

⑤ Involves langer no. of subjects.

⑥ Long followup period often needed involving delayed results.

⑦ In appropriate when the disease or exposer under investigation is rare.

⑧ Yields incidence rate relative risk as well as attributable risk.

⑨ Can yield information about more than one disease outcome.

(10) Relatively inexpensive.	(10) Expensive

Blinding :-

Blinding can be done in three ways.

(1) Single Blind Trial - The trial is so plan that the participent in not awane wheather he belongs to the study to control group.

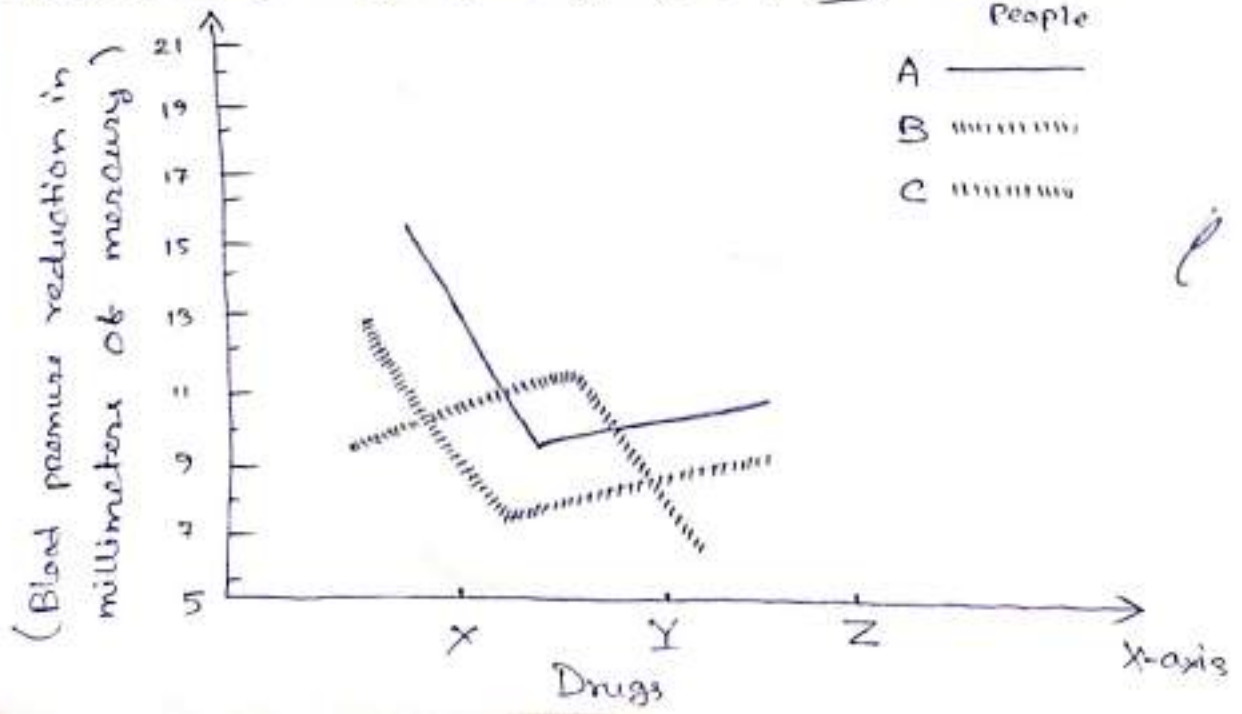
(2) Double Blind Trial - The trial is so plan that nither the doctor not the participent is awane of the group allocation and the treatment receive.

(3) Triple Blind Trial - This goes one step further the participent, the investigator and the person analysing the data all 'Blind'. Idealy offcourse triple blinding should be used but the double blinding is the most frequently use method when a blind trial is conducted.

Date - 20/07/22

Graphic method of studying interaction in a two way design:

Interaction can be studied in a two-way design with repeated measurements through graphic method also. For such a graph we shall select one of the factors to be used as the X-axis. Then we plot the average for all the samples on the graph and connect the averages for each variety of the other factor by a distinct mark. If the connecting lines do not cross over each other, then the graph indicates that there is no interaction, but if the lines do cross, they indicate definite interaction or inter-relation between the two factors) Let us draw such a graph for the data of Illustration 3 of this chapter to see whether there is any interaction between the two factors viz., the drugs and the groups of people. [Groups of the averages for amount of blood pressure reduction in millimeters of mercury for different drugs and different groups of people] Groups of People



The graph can alternatively be drawn by taking different group of people on x-axis and drawing lines for various drugs through the average.

The graph indicates that there is a significant interaction because the different connecting lines for groups of people do cross over each other. We find that A and B are affected very similarly, but C is affected differently. The highest reduction in blood pressure in case of C is with drug Y and the lowest reduction is with drug Z, whereas the highest reduction in blood pressure in case of A and B is with drug X and the lowest reduction is with drug Y. Thus, there is definite inter-relation between the drugs and the groups of people and one cannot make any strong statements about drugs unless he also qualifies his conclusions by stating which group of people he is dealing with. In such a situation, performing F-tests is meaningless. But if the lines do not cross over each other (and remain more or less identical), then there is no interaction or the interaction is not considered a significantly large value, in which case the researches should proceed to test the main effects, drugs and people in the given case, as stated earlier.

R^2 and Adjusted R^2 Coefficients

As in simple linear regression analysis, here also we can use coefficient of determination

R^2 to judge goodness of fitted model by the same formula :

$$R^2 = SSR/SST) = 1 - (SSE/SST)$$

where, SSR, SSE and SST are obtained using the same formula as discussed earlier. The only difference is in calculating predicting Y values which are given by $\hat{Y} = \hat{a} + \hat{b}_1 X_1 + \hat{b}_1 X_2 + \cdots + \hat{b}_K X_K$.

In example 14.8, $R^2 = 0.5215$. If we consider fitting a simple linear regression model taking Pie sales as response variable and price as explanatory variable, $R^2 = 0.1965$. If we fit a simple linear regression model taking Pie sales as response variable and Advertising Expenditure as explanatory variable, $R^2 = 0.3055$

Thus, we observe here R^2 is higher when there are more explanatory variables in the model. This is not just a coincidence, but a property of R^2. Whenever an explanatory variable is added to the model, the value of R^2 increases. This increase is regardless of the contribution of newly added explanatory variable. Thus the value of R^2 may be misleading in multiple linear regression models. Therefore, an adjusted value of R^2 is defined, which is called as adjusted R^2 and defined as

$$R^2_{adj} = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)}$$

Adjusted R^2 increases only if the

added explanatory variable contributes variable contributes in explaining the variation in response variable Y.

In example 14.8, Adjusted $R^2 = 0.4417$

Variable Elimination

There are mainly three techniques of variable elimination : (i) stepwise method, (ii) forward inclusion method, and (iii) backward elimination method.

Stepwise method

This method is based on the p-value of F-test obtained in ANOVA. The method is illustrated as below:

Suppose we have a response variable and 5 explanatory variables X_1, X_2, X_3, X_4, X_5.

Step 1 : Fit 5 simple linear regression models:

- Predicted $Y = a + b_1 X_1$
- Predicted $Y = a + b_2 X_2$
- Predicted $Y = a + b_3 X_3$
- Predicted $Y = a + b_4 X_4$
- Predicted $Y = a + b_5 X_5$

We obtain the p-value of F-test for all the five simple linear regression models mentioned above. Suppose that the p-value corresponding to the fourth model (having X_4) is the smallest (and below 0.05 as well), then the first explanatory variable included in the model is X_4.

Step 2 : Fit 4 two-variable linear regression models having X_4 as one explanatory variable:

- Predicted $Y = a + b_4 X_4 + b_1 X_1$
- Predicted $Y = a + b_4 X_4 + b_2 X_2$
- Predicted $Y = a + b_4 X_4 + b_3 X_3$
- Predicted $Y = a + b_4 X_4 + b_5 X_5$

We again obtain the p-value of F-test for all four models of this step. Suppose that the p-value corresponding to the third model (having X_4 and X_3) is the smallest (and below 0.05 as well), then the second explanatory variable included in the model is X_3.

Step 3 : Fit 3 three-variable linear regression models having X_4 and X_3 as two explanatory variable :

- Predicted $Y = a + b_3X_3 + b_4X_4 + b_1X_1$
- Predicted $Y = a + b_3X_3 + b_4X_4 + b_2X_2$
- Predicted $Y = a + b_3X_3 + b_4X_4 + b_5X_5$

Suppose p-value is not below 0.05 for any of these models. We stop here. The best model is the one with X_3 and X_4 only.

Q

What is agglomerative clustering? How it is different from divisive clustering?