

“THE PRINCIPAL STEPS IN A STATISTICAL ENQUIRY”



CONTENTS

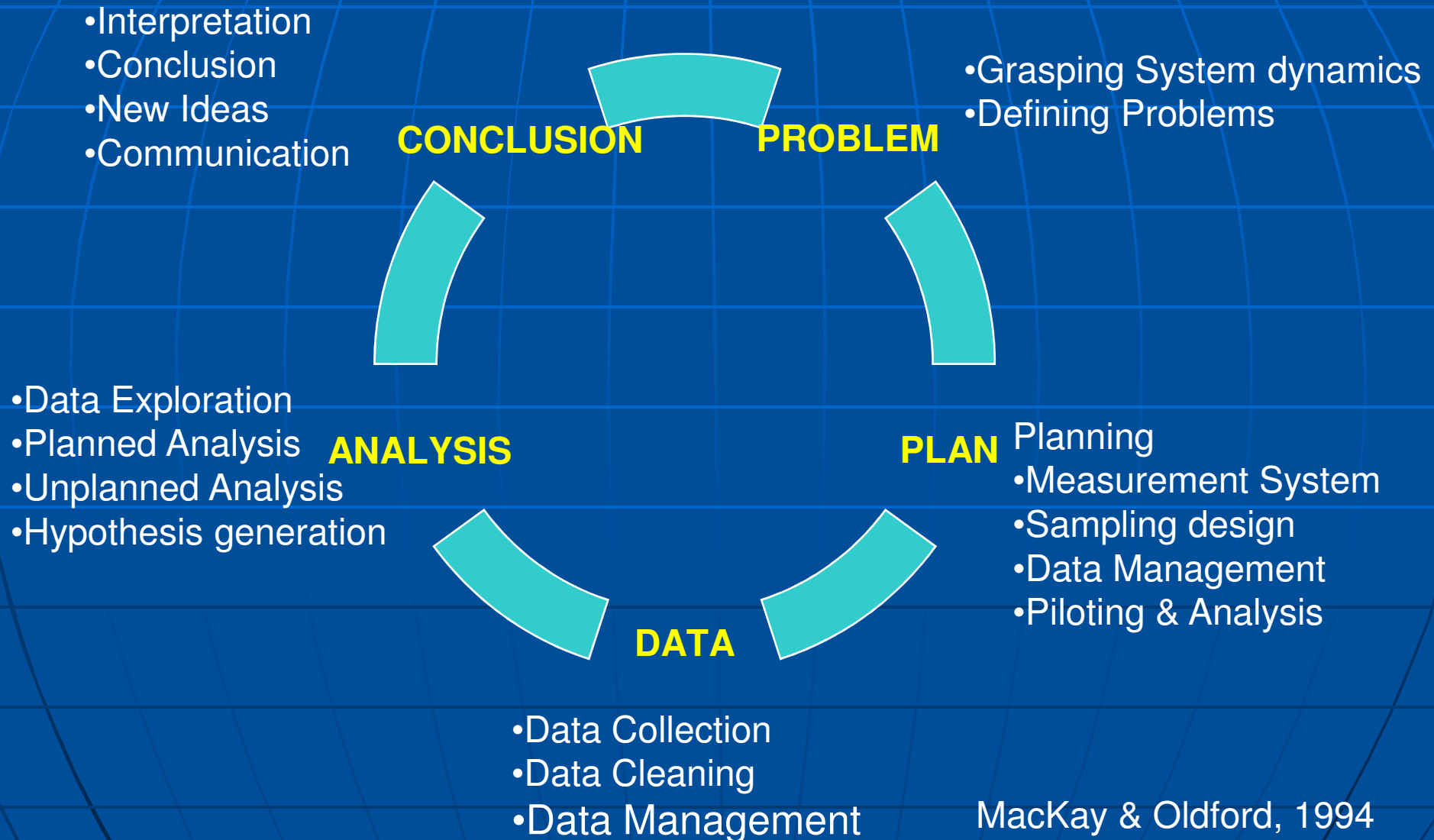
- Statistical Thinking in Empirical Analysis
- Principal Steps in a Statistical Enquiry

OBJECTIVES

- The trainees will adopt Statistical Thinking to problem solving in Medical Research.
- They will be able to plan Research in a scientific way.

STATISTICAL THINKING

DIMENSION – 1: Investigative Cycle (PPDAC Model)



DIMENSION 2:Types of thinking

GENERAL TYPES

- **Strategic**
 - Planning, anticipating problems
 - Awareness of practical constraints
- **Seeking Explanations**
- **Modeling**
 - Construction followed by use
- **Applying Techniques**
 - following precedents
 - recognition and use of archetypes
 - Use of problem solving tools

TYPES FUNDAMENTAL TO STATISTICAL THINKING

- ❑ Recognition of need for data

- ❑ Transnumeration

(Changing representation to engender understanding)

- ❑ capturing “measures” from real system

- ❑ Changing data representation

- ❑ Communicating messages in data

- ❑ Consideration of variation

- ❑ Noticing and acknowledging

- ❑ Measuring and modeling for the purpose of prediction, explanation, or control

- ❑ Explaining and dealing with

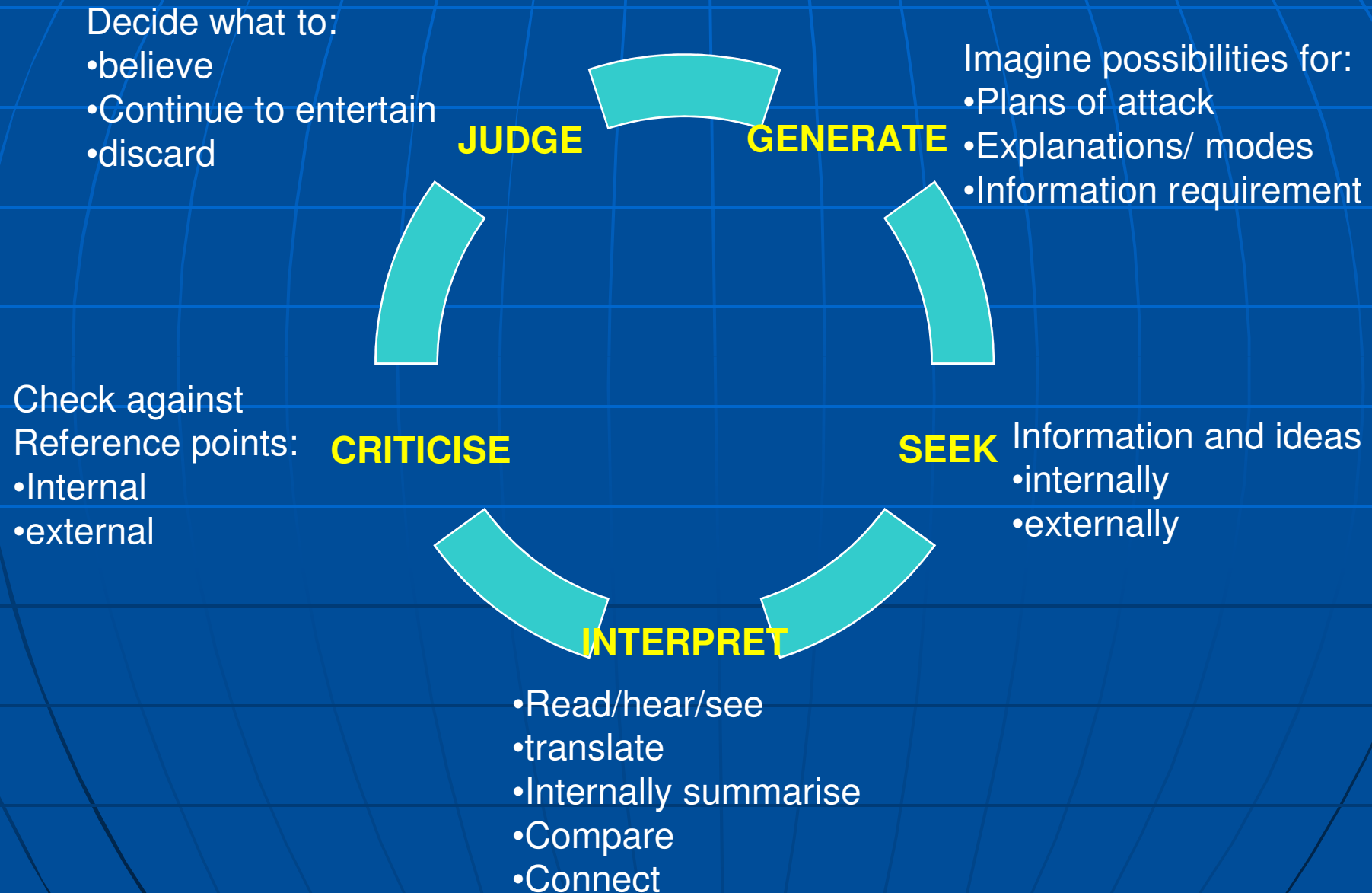
- ❑ Investigative strategies

- ❑ Reasoning with statistical model

- ❑ Integrating the statistical and contextual

- ❑ Information, knowledge, conceptions

DIMENSION – 3: The Interrogative Cycle



DIMENSION 4: Disposition (Arrangement / Person's natural Qualities of mind and Character)

- Scepticism
- Imagination
- Curiosity and awareness
 - Observant, noticing
- Openness
 - To ideas that challenge preconceptions
- A propensity to seek deeper meaning
- Being logical
- Engagement
- Perseverance

Pertinent methodological questions

1. What are the objectives?
2. What is the coverage?
3. What type of information required?
4. What are the methods to collect data?
5. Whether census or sampling?
6. If a sample survey- the design?

Some pertinent methodological questions

7. What are the potential sources of errors and possible precautions?
8. What are the steps to conduct the field work efficiently?
9. What are the methods to process and analyses the collected data?
10. What are the points to be kept in mind while writing the report and presenting the results of the survey?
11. What are the problems of accuracy, errors and approximations?

A STATISTICAL ENQUIRY

- It is with a Predefined purpose and dealing with collection of data in a systematic manner.
- Data collected from such enquiry are called statistical data.
- It may be Descriptive Field Surveys or Analytical Experimental study under controlled conditions.
- Descriptive Field Survey:
 - Cross Sectional Study/ Prevalence Study: Simplest form of an observational Study.
 - Longitudinal study: Observations are repeated in the same population over a prolonged period of time by means of follow up examinations.
- Analytical studies: Case Control Study and Cohort Study

Planning a Statistical Enquiry

1. What are the objectives of the survey?

- Objectives - clear and unambiguous.
- Possible uses of final results to be expected and the desired degree of accuracy be discussed at the beginning.
- Lack of clear cut objectives might undermine the value of the final results.
- and the manipulation of data at the end would not solve the issue in questions and eradicate the inherent defects.

2. Population to be covered:

- **What is the population to be studied?**

- **Target population**
- **Sampled population**

- **What are the units from which data is to be collected?**

3. Data to be Collected

- ☐ **Should be decided keeping in view the objectives of the study.**
- ☐ **Relevant data should not be omitted.**
- ☐ **Irrelevant data should not be collected.**

Primary & Secondary data

- **Primary data:** Data Collected by the investigator through direct observation, mailed questionnaire or interview method from the field or conducting experiments.
- **Secondary data:** Ready made data collected from published or unpublished form in Hospital Records
 - Involves less time and money
 - It may not always be adequate and representative enough to serve the required purposes and may also be outdated.
 - Further, secondary data may have unknown degree of accuracy

- The **primary data** are usually in a raw and bulky form and do not possess the simple look of the secondary data, which are already in processed form. This when processed and put in an appropriate classified and tabulated form by a recognized body, becomes secondary data for future users.

4. Designing questionnaire / schedule

- ❖ Questionnaire: To be filled in by the respondent.
- ❖ Schedule: - To be filled in by the interviewer

Designing questionnaire / schedule

- ❖ **This requires:-**

- ❖ Skill,

- ❖ special techniques,

- ❖ Familiarity with the subject matter

- **Question content:**

- Respondents are likely to have knowledge to answer them.

- Large number of questions to the respondents should be avoided.

- The questions should be such that one is able to extract brief answers.

- The questionnaire should have some cross-questions to check the authenticity of the information supplied

Question wording

- The question should be specific, simple and unambiguous. Vague, presuming and hypothetical questions should be avoided. If necessary, the questionnaire should contain some leading questions. The questions should be such that it is capable of being answered without prejudice.

Open and pre-coded questions

- In an open question the respondent has the freedom in choice of answers. But in case of pre-coded or multiple choice questions the respondent is asked to choose one from the limited number of given answers. Open questions are, often, found difficult to be coded. Pre-coded questions restrict the freedom of thought of the respondents.

Question order

- Ordering of the questions is an important part of the construction of questionnaire. The ordering systematizes the thought process of the interviewer and as a result reduces the refusal rate of the survey based on interviewing.
- ❖ **Suitable instruction for filling-up the same is a must.**

5. Approach to data collection

- **Complete enumeration (Census)**
- **or Sample study?**

6. Selection of a proper sampling design

- Sampling units
- Reference unit
- Sampling frame:-List of all the sampling units in the population under study?
- Minimum sample size required
- Sampling design

Minimum Sample size

- This will depend upon
 - Aim, nature, scope of the study, variability in the population and on the expected results.
- Standard procedure for different situations
 1. One sample situation
 - **Estimating a population proportion with specified absolute/relative precision**

2. Two sample situation

- **Estimating the difference between two population proportions**
- **Hypothesis tests for two population proportions Case-control studies**

3. Case-control studies

- **Estimating an odds ratio**
- **Hypothesis tests for an odds ratio**

4. Cohort studies

- **Estimating a relative risk**
- **Hypothesis test for a relative risk**

5. Lot quality assurance sampling

- Accepting a population prevalence as not exceeding a specified value
- Decision rule for “rejecting a lot”

6. Incidence-rate studies

- Estimating an incidence rate
- Hypothesis tests for an incidence rate
- Hypothesis tests for two incidence rates in follow-up (cohort) studies

Sampling Design

- Sampling design will depend upon
 - Objectives, scope and coverage of the study, nature and size of the population, hypothesis to be tested
- Different scheme of sampling
 - Probability sampling
 - Non-probability sampling

7. Different methods of data collection

- Personal observations,
- Mailed questionnaires,
- Interviewing,
- Documental sources, etc.
- A choice to be made depending upon: the type of data to be collected and type of population to be covered.

■ **Mailed Survey**

- Less costly
- Non-response may be high
- Practicable for educated respondents

■ **Interview Method**

- More cost
- Less non-response
- Practicable for educated and uneducated both
- Concept and definition can be explained to the respondents.

■ Observation Method

- Specify the method of measurement
- Specify unit of measurement
- Provide measuring equipments and instruments

8. Dealing with non-response

- Due to non-availability of the respondent
- Respondent may fail to give the data when contacted.
- Respondent may refuse to give information.
- Non response tends to change the results.
- Procedure to be devised to deal with it.
 - **Call-back procedures,**
 - **Substitution**

9. Errors

- Any statistical survey is not free from errors:
- Errors may be on account of
 - **Sampling: Which can be kept under control by following a suitable sampling design**
 - Faulty selection of the sample
 - Substitution
 - Faulty demarcation of the sampling unit
 - Improper choice of statistics for estimating the population parameter.

9. Errors- Contd.

- Non-sampling errors
 - **Faulty Planning or Definitions**
 - **Response Errors**
 - May be accidental, Prestige bias, Self-interest, Bias due to interview, Failure of respondent's memory
 - **Non-response bias:** Full information not obtained of all sample units
 - **Error in coverage:** Objectives not clearly stated
 - **Compiling Errors:** Various operations in data processing
 - **Publication errors:** Committed during presentation and printing of tabulated results

9. Organisation of field work

- **Reliable field work is a precondition for the success of the statistical enquiry.**
- **Organizational set-up**
 - **Selection of Team leader**
 - **Selection of Investigators**
 - **System of supervision**
 - **Logistic arrangements: Stay, Transport & Communication, printing of schedules, supply of appropriate stationeries**

9. Organisation of field work

■ Training

- Orientation of the Investigators with the;**
 - instruments,**
 - Levels and units of data collection**
 - concepts and definitions of the field of information**
- How to deal with non-response**

10. Pre-testing and Pilot survey

Pre-testing:

Trying out the questionnaire on a small scale.

The pre-testing of a questionnaire brings out the drawbacks of the questionnaire as to question order, content and wording and enables the survey specialist to reach a decision whether to add certain more questions or delete some already incorporated. It studies possible reactions and difficulties of the respondents to answer certain question and remodel accordingly.

Pilot Survey

- ❑ Pilot survey is like theatrical dress rehearsal
- ❑ This is to try out the entire survey on a limited scale to know:-
 - ❑ the nature of the population
 - ❑ efficacy of the sampling design
 - ❑ deficiency in the planning
- ❑ This enables the researchers to correct deficiency before the main survey

11. Processing of the Data

- ❑ Data Validation or Editing of data
- ❑ Coding-To put the results in quantitative form by assigning numeric codes
- ❑ Classification and tabulation of data
- ❑ Presentation of data
- ❑ Analysis of data

12. Reporting and Conclusions

- ❑ Finally, the analysis should be interpreted in the form of a report incorporating:-
 - ❑ Detailed statement of the different stages of the survey,
 - ❑ technical aspects of the design, viz., the types of the estimators used along with the amount of error to be expected in the most important estimates, and
- ❑ Presentation of the results.

13.Information gained for future surveys

- ❑ Any completed survey helps to understand the nature of the data in terms of means (average), standard deviations (variability), cost involved in obtaining the data.
- ❑ Any completed survey provides lesson for designing future surveys in recognizing and rectifying the mistakes committed in the execution of the survey.
- ❑ Thereby it is a learning point for improved planning and execution of the future surveys.

REFERENCES

- Biostatistics – A foundation for Analysis in the Health Sciences: Wayne W. Daniel, Seventh Edition, Wiley Students Analysis.
- A First Course in Statistics with Application: A.K. P. C Swain, Kalyani Publishers.
- Fundamentals of Applied Statistics, SC Gupta and V. K. Kapoor
- Sample Size determination in Health Studies: A Practical Manual, SK Lwanga and S Lemeshow, WHO,
- Parks Text book of Preventive and social medicines

The background is a solid blue color with a subtle grid pattern of thin, lighter blue lines. The grid consists of both horizontal and vertical lines, creating a series of small squares across the entire image.

THANK YOU