Chapter 1: General Overview of Biostatistics

1. What Is Statistics?

Statistics is the science of drawing conclusions about random phenomena using information collected from a sample. A sample is a smaller group selected from a larger population. Statistical methods help researchers organize data, describe patterns, and decide what the data may tell us about the population.

There are two broad areas of statistics.

Mathematical statistics develops statistical methods and relies on advanced mathematics.

Applied statistics uses statistical methods to answer questions in specific fields, such as economics, psychology, public health, medicine, and biology.

2. What Is Biostatistics?

Biostatistics is a branch of applied statistics that uses statistical methods to study medical and biological questions. It helps researchers design studies, collect and check data, summarize results, and evaluate evidence. A method that works well for one study may not be suitable for another, so biostatisticians may need to develop or adapt methods for particular research problems.

3. Biostatistics in Medical Research: An Example

Imagine that a patient's blood pressure is measured using an automatic machine and a manual blood pressure cuff. The readings may differ. Researchers may ask: Are the two methods comparable, or is the difference more than we would expect by chance? Biostatistics provides a systematic way to investigate this question.

4. Planning and Conducting a Study

A study must be planned carefully before data are collected. In the blood-pressure example, researchers selected four machines and planned to test them with participants at each location. The order of automatic and manual measurements was randomized to reduce the possibility that the order itself would affect the results. Researchers also recorded information that might influence readings, including age, sex, history of hypertension, and body size.

Good data management is essential. Each participant should receive a unique identification number. Data should be recorded in a consistent format, entered carefully, and checked for errors. Researchers can use computer programs to identify values outside expected ranges; unusual values should be checked rather than automatically deleted. Missing or invalid measurements can occur, so they should be anticipated during study planning.

5. The Main Steps in a Statistical Study

Data collection: obtain measurements or responses from participants.

Data entry: enter the information into a computer in a consistent format.

Data editing: check records for missing and inconsistent values.

Data analysis: summarize the data and apply appropriate statistical methods.

Publication: report the important findings in a scientific manuscript.

6. Descriptive Statistics

Descriptive statistics organize and summarize the data that were collected. Results may be presented numerically—for example, in tables or frequency distributions—or graphically, using charts and other visual displays. The type of summary depends on the kind of data.

Continuous data: measurements that can take many possible values, such as blood pressure.

Discrete data: data with separate possible values. For example, the number of visits to a doctor.

7. Inferential Statistics

Inferential statistics uses sample data to draw conclusions about a larger population. Researchers must consider whether an observed difference is likely to reflect a real difference or could reasonably have occurred by chance. They use a probability model to evaluate how compatible the observed data are with the assumption that there is no real difference.

8. Statistical Software

Statistical software contains programs for summarizing data and carrying out statistical tests. For example, SAS, SPSS, Stata, MINITAB, and Excel.

Key Terms to Remember

Population: the whole group of interest.

Sample: the part of the population actually studied.

Descriptive statistics: methods for organizing and summarizing data.
Inferential statistics: methods for using sample data to draw conclusions about a population.

Quick Review

- What is the difference between mathematical statistics and applied statistics?
- What is biostatistics?
- Why should researchers randomize the order of measurements?
- Give one example each of continuous data and discrete data.