Statistical methods in epidemiology

Statistical Methods in Epidemiology

Definition:

- Statistical methods in epidemiology involve the application of statistical principles and techniques to understand the distribution and determinants of diseases in populations.
- These methods help in designing studies, analyzing data, interpreting results, and making decisions for public health.

Core Elements:

A. Study Design

- Statistical methods begin with selecting an appropriate study design.
- **Common types include:**

1. **Descriptive Studies**
   - Focus on *who, what, when, and where* of disease occurrence
   - Use measures such as rates and proportions
   - Describe disease patterns in populations
2. **Analytical Studies**
   - Examine causes or risk factors of diseases
   - Include:
     - Cohort studies
     - Case-control studies
     - Cross-sectional studies
3. **Experimental Studies**
   - Involve active intervention by the researcher
   - Include randomized controlled trials (RCTs)
   - Participants are randomly assigned to intervention or control groups

B. Data Collection and Management

- Accurate and systematic data collection is essential and includes:

1. **Identification of variables:**
   - Exposures
   - Outcomes
   - Confounders

2. **Ensuring data quality:**
    - Minimizing bias
    - Handling missing data
3. **Use of databases and statistical tools:**
    - Epi Info
    - SAS

C. Descriptive Statistics

- Used to summarize and describe data:
- **Measures include:**
    - Mean
    - Median
    - Mode
    - Standard deviation
    - Prevalence and incidence rates
- Helps understand the basic characteristics of the data

D. Inferential Statistics

- Used to draw conclusions about a population from a sample:
- **Key techniques include:**
    - Hypothesis testing (null and alternative hypotheses)
    - Confidence intervals (range of expected true values)
    - p-values (probability that results occurred by chance)

E. Regression Analysis

- Used to assess relationships between variables:
    - **Linear regression** – for continuous outcomes
    - **Logistic regression** – for binary outcomes (e.g., disease vs. no disease)
    - **Cox proportional hazards model** – for survival or time-to-event data

F. Bias and Confounding

- **Statistical methods help to:**
    - Detect bias (systematic errors)
    - Control confounding through:
        - Stratification
        - Multivariable regression models

G. Modeling and Simulation

- **Used for:**
  - Disease prediction
  - Estimating disease burden
  - Simulating the impact of interventions
  - Applying statistical or mathematical models

H. Sensitivity Analysis

- Tests the robustness of results
- Examines how changes in assumptions or methods affect conclusions

Incidence and Prevalence

These are fundamental epidemiological measures used to describe the frequency of disease in a population.

A. Incidence

Definition:

Incidence measures the number of **new cases** of a disease that occur in a defined population over a specified period of time.

A. Incidence Proportion (Cumulative Incidence):

$$\text{Cumulative Incidence} = \frac{\text{Number of new cases during a specified time period}}{\text{Number of people at risk at the beginning of the period}}$$

- It is a proportion, ranging from 0 to 1 (or expressed as a percentage).
- Best used in closed populations (no loss to follow-up, no new entries).

B. Incidence Rate (Incidence Density):

$$\text{Incidence Rate} = \frac{\text{Number of new cases}}{\text{Total person-time at risk}}$$

- Person-time = sum of time each individual was at risk before developing the disease or being censored.
- Useful in dynamic populations or studies with variable follow-up.

B. Prevalence

Definition:

Prevalence measures the **total number of existing cases** (both new and old) of a disease in a population at a specific point in time (point prevalence) or over a period of time (period prevalence).

A. Point Prevalence:

$$\text{Point Prevalence} = \frac{\text{Number of existing cases at a specific time}}{\text{Total population at that time}}$$

B. Period Prevalence:

$$\text{Period Prevalence} = \frac{\text{Number of people who had the disease at any time during a period}}{\text{Average population during that period}}$$

Differences between Incidence and Prevalence:

Feature	Incidence	Prevalence
Counts	New cases only	All cases
Time	Over a period	At a point or during a period
Use	For studying disease etiology	For planning health services
Affected by	Disease risk	Both incidence and duration

Relationship:

$$\textbf{Prevalence} \approx \textbf{Incidence} \times \textbf{Average Duration of Disease}$$

Relative Risk (RR)

Definition:

- Relative risk is a ratio that compares the risk of an outcome (usually disease) in the exposed group to the risk in the unexposed group.

Formula:

$$\text{RR} = \frac{\text{Incidence in Exposed Group}}{\text{Incidence in Unexposed Group}}$$

Interpretation:

- **RR = 1:** No association; risk is the same in both groups.
- **RR > 1:** Positive association; exposure increases risk.
- **RR < 1:** Negative association; exposure may be protective.

Example (Cohort Study):

Group	Developed Disease	Did Not Develop Disease	Total
Exposed	a	b	a + b
Unexposed	c	d	c + d

First Hope

FREE PREVIEW

Click here to visit our website

PHARMACY PREMIUM NOTES

Complete subject and semesters are available on FirstHope app only.

Scan or click here

Click here to get the app >

Available on Google Play

JOIN THE FIRSTHOPE COMMUNITY

Join our Telegram channel

Click here to join Telegram >

Scan or click here

@FirstHopeNotes

Instagram

@firsthope.official

Click here to follow >

YouTube channel

FirstHope-Pharmacy / Live streams

Click here to watch >

WhatsApp channel

Follow FirstHope on WhatsApp

Click here to join >

FirstHope website

firsthope.co.in

Click here to visit >

B.Pharm / D.Pharm / M.Pharm / Pharm.D

Preview notes. Complete subjects and semesters in the FirstHope app.