📘 Introduction

Introduction

- Statistics estimate population characteristics from small, random samples.
- **Key terms:**
    - *Population* — entire group under study
    - *Parameter* — true value in population
    - *Sample* — subset surveyed
    - *Statistic* — estimate based on the sample
- Small, random samples can accurately represent large populations.
- Statistics make it possible to study large groups efficiently.

📘 Simple Random Sampling and Stratified Random Sampling

Simple-random-sampling-and-stra…

- Sampling only from convenient sources (e.g., one's hometown) introduces **bias**.
- Common biases:
    - *Selection bias* — certain groups favored
    - *Non-response bias* — some groups respond less
    - *Voluntary response bias* — only motivated participants respond
- **Simple random sampling**: every individual has an equal chance of being chosen.
- **Stratified random sampling**: population divided into subgroups (strata); samples taken from each — more precise but more complex.

📘 Bias and Chance Error

Bias-and-chance-error

- Random sampling introduces **chance error**, causing estimates to differ from true parameters.
- Each estimate = parameter + bias + chance error.
- Larger samples reduce **chance error**, which can be quantified and controlled.
- **Bias** (systematic error) does not shrink with sample size — it must be avoided through good design.
- Random sampling minimizes bias and supports reliable inference from samples to populations.
- Key types of bias: selection, non-response, voluntary response.
- Understanding and managing bias and chance error are essential for valid conclusions.

📘 Observation vs. Experiment: Confounding and the Placebo Effect

Observation-vs-experiment-confo…

- **Observational studies** show associations (e.g., diet and disease) but cannot prove causation due to **confounding variables**.
- **Experiments** use random assignment to treatment and control groups to test causality.
- **Control groups** receive **placebos** to balance psychological effects (the **placebo effect**).
- **Double-blind experiments** (neither subject nor evaluator knows assignment) reduce bias.
- The placebo effect highlights psychological influences in medical outcomes.

📘 The Logic of Randomized Controlled Experiments

The-logic-of-randomized-control…

- Good experiments require **random assignment** to ensure confounders are evenly distributed.
- Uneven results may still occur by **chance**, whose effects can be statistically estimated.
- Random sampling uses chance to fairly select participants and reduce bias.
- Larger samples decrease random error but do not fix systematic bias.
- **Bias** must be eliminated through design; **chance error** can be measured and managed.
- Randomized controlled trials (RCTs) allow reliable inference about causal effects in populations.

✅ Overall Summary:

- Statistics rely on random sampling to estimate population parameters accurately.
- Proper design minimizes **bias** and manages **chance error**.
- **Experiments** (especially randomized and double-blind ones) establish causation, while **observational studies** reveal correlations only.
- Randomization and appropriate sampling methods are the foundation of reliable, unbiased conclusions in research.