

Statistics

| Introduction

Prof. Dr. Mostafa Elhosseini

<https://youtube.com/drmelhosseini>

| Agenda

- What is a statistic?
- What is statistics?
- Everyday Scenarios Involving Statistics
- Statistical Problem
- Observational - Experimental Studies
- Margin of error
- Subject
- Population and Sample



What Is Statistics?

Statistics is the art and science of:

- Designing studies
 - Collecting and analyzing data
 - Drawing conclusions and making informed decisions
-

الإحصاء هو فن وعلم يُعنى بتصميم الدراسات وجمع البيانات وتحليلها، بهدف استخلاص استنتاجات دقيقة واتخاذ قرارات مبنية على الأدلة. يُستخدم الإحصاء لفهم الأنماط، واكتشاف العلاقات، والتنبؤ بالنتائج في مختلف مجالات الحياة مثل الطب، والاقتصاد، والتعليم، والهندسة.



What is a Statistic?

- A statistic is just a single number calculated from data.
- Example: average, median, percentage



Examples:

Field	Example Statistic
 Sports	Number of points scored by each player in a game
 Economy	Median household income, Unemployment rate



In short:

- A statistic is a number.
- Statistics is the method of understanding and using such numbers meaningfully.



Everyday Scenarios Involving Statistics

- Statistics is all around us — helping us understand, evaluate, and make better decisions in daily life.



Media & Society



Are cell phones dangerous to your health?



Is there a bias against women in appointing managers?



How can you evaluate evidence about global warming?



Health & Wellness



How can you analyze whether a diet really works?



Are new treatments better than current ones for diseases?



Economics & Lifestyle



How can you predict the selling price of a house?



How do you measure employee performance?



What's the chance your tax return (إقرار ضريبي) will be audited?

| Everyday Scenarios

- Important to **understand** statistics **even if you will never use** it in your job
 - help you make better choices
 - political campaigns, and surveys
 - Does taking **aspirin** daily truly lessen the chance of having a **heart attack**

| Statistical Problem

- Statistical problem solving is an investigative process that involves four components:
 - **Formulate** a statistical question
 - **Collect** data
 - ↳ Information we gather
 - ↳ low-carbohydrate diet - BMI
 - **Analyze** data, and
 - **Interpret** results

| Statistical Problem

Field	Real-World Problem	Statistical Question	Data Collection	Data Analysis	Interpretation
Sport	Evaluating player fitness impact on match performance	Does better fitness level correlate with improved match performance?	Fitness scores, match statistics (goals, assists, minutes played)	Correlation analysis, scatterplots, regression	Strong correlation supports focusing training on fitness improvement
Politics	Measuring voter reaction to televised debates	Do televised debates influence voter preference?	Pre- and post-debate polling data, demographic info	Paired t-tests, bar charts, opinion trend analysis	A significant shift in polling may show debate influence
Marketing	Assessing social media campaign impact on product sales	Did the new social media campaign increase product sales?	Sales data before/after campaign, campaign reach metrics	A/B testing, line graphs, time-series analysis	A clear increase in sales post-campaign indicates marketing effectiveness
Health	Testing effectiveness of a new flu vaccine	Is the new flu vaccine more effective than the previous one?	Infection rates in vaccinated and unvaccinated groups	Chi-square test, proportion comparison	A lower infection rate in the vaccinated group supports vaccine effectiveness
Industry	Analyzing effect of machine maintenance on production defects	Does regular machine maintenance reduce defect rates in a factory?	Defect counts before and after maintenance; maintenance logs	Control charts, z-test for proportions	Fewer defects after maintenance justify scheduled upkeep

| Observational - Experimental Studies

- Cell phones emit electromagnetic radiation, and a cell phone's antenna is the main source of this energy. The closer the antenna is to the user's head, the greater the exposure to radiation. Several studies have explored the possibility of such risks
 - Why do results of different medical studies sometimes disagree?
 - ↳ They are not the same type of study
 - ↳ How can you design the study to be efficient?
 - ↳ Experimental and Observational Studies

| Margin of error

? Does regular aspirin intake reduce deaths from heart attacks?

— Aspirin or a placebo?

? Can you conclude that it's beneficial for people to take aspirin regularly?

Margin of Error

- The margin of error quantifies the uncertainty in study results.
- Studies with a high margin of error are less reliable.
- A low margin of error means greater confidence in the findings.

| Answering A statistical Question

- To answer a statistical question, we follow a structured process involving:
 - **Design** – Plan how to collect data
 - **Description** – Summarize and explore the data
 - **Decision** – Use data to make predictions or conclusions
- Example:
 - 💡 Do most voters support Candidate A?

| Design

- Design questions: How to conduct the experiment
 - Plan how to collect data
 - How to select people for the survey to ensure trustworthy results
 - Choose method: survey, experiment, observation
- Example:
 -  Conducting exit polls after voting
 -  Planning the methods for data collection to study the effects of Vitamin C
 -  How do you select people for your survey - predictions about future sales?

| Description

- Description: **Summarize** the raw data and present it in a useful format (e.g., **average**, **charts**, or graphs).
- It is more informative to use a few numbers or a graph to summarize the data
 - Organize and summarize the collected data
 - Use charts, graphs, averages, variability
- Examples
 -  Summarizing voter responses by region or age group
 -  Average amount of TV watched

| Inference

- Inference:

- Use statistical models to interpret results
- Generalize from sample to population
- Estimate error and uncertainty
- Make decisions or predictions based on the data.

- Examples:



Has there been **global warming** over the past decade?



Is having the **death penalty** as a possible punishment associated with a **reduction** in violent crime?



Does **student performance** in school depends on the amount of **money spent per student**, the **size of the classes**, or the **teachers' salaries**?

| Subject

- An individual unit or entity from which data is collected
 - A person (e.g., a patient in a clinical trial)
 - An object (e.g., a manufactured part in a quality control study)
 - An event (e.g., a crime incident in a crime rate study)
 - A unit of analysis (e.g., a school, company, or country)

| Subject - Variable

◆ A **variable** is a characteristic or measurement recorded for each subject.

Schools

 Per-student expenditure

 Average class size

 Average student test scores

Countries

 % of residents living in poverty

 Birth rate

 % unemployed

 % computer literate

Days – Internet Café

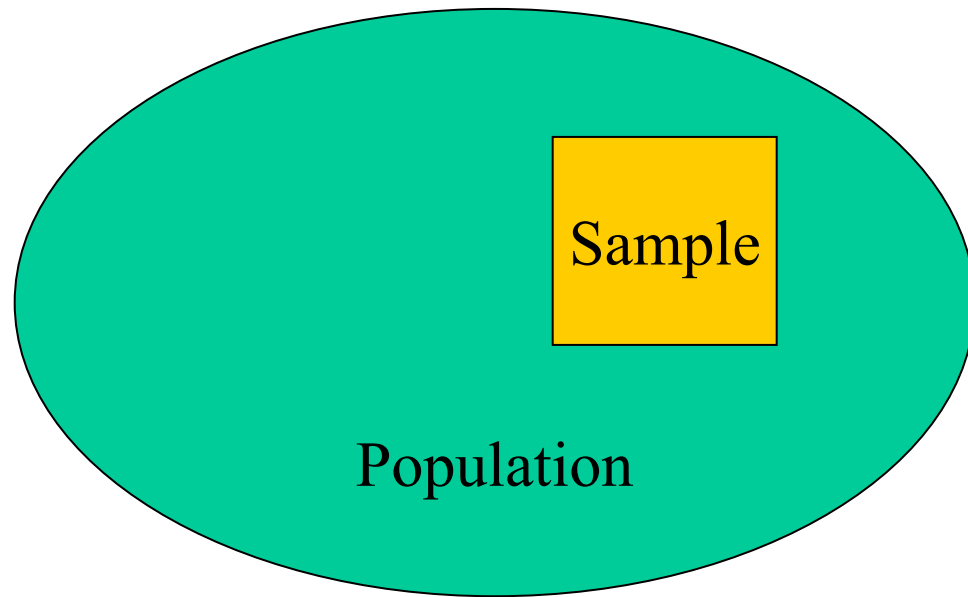
 Amount spent on coffee

 Amount spent on food

 Amount spent on Internet access

| Sample and Population

- Population: The total set of subjects in which we are interested
- Sample: Subset of the population for whom we have data



| Population and Sample

Population

- The entire group of subjects we want to learn about.
- Sometimes well-defined and countable.
 -  Example: Exit poll — All voters in an election.
- Sometimes hypothetical or not fully listed.
 -  Example: In a clinical trial on a cholesterol-lowering drug,
The population = all individuals with high cholesterol who could have enrolled.
 -  All patients who could be affected by a new gene therapy in the future

| Population and Sample

Sample

- A subset of the population that we actually observe or collect data from.
- Often selected randomly to represent the whole.
- Used to make inferences about the population.

Key Point

-  We observe samples, but our goal is to understand populations.