MA-120 Probability and Statistics

Nazar Khan
PUCIT
Lecture 1: Introduction

What is Statistics and Probability

- **Statistics** is the 'art' of
  - understanding the "real" world as it is and not "how we think" it is,
  - intelligently summarizing large amounts of data,
  - making numerical guesses for puzzling questions.
- **Probability** is the 'tool'
  - to work with statistics,
  - to make conclusions/predictions from statistics.

Statistics

- Even a lifeless calculator can give you statistics by plugging numbers into formulae.
- But the true meaning of those statistics requires careful thinking.
- One aim of this course is to make you think like a statistician, not like a calculator!

Probability

- One of the more important branches of mathematics.
- Can be a bit unintuitive.
- Has its own terminology.
- Every probability problem requires **thinking**.
  - Fortunately, there are some tricks.
- One aim of this course is to make you develop thinking skills that help solve probability problems!

Applications of Probability and Statistics

- Computer Networks
- Machine Learning
- Computer Vision, Image Processing, Graphics
- Algorithms
- Data Mining

Applications of Probability and Statistics

- Politics
- Economics
- Social Sciences
- Medicine
- Everything involves probability and statistics!

Applications of Probability and Statistics

- Every two days we create as much data as we did from the beginning of mankind till 2003.
- The **only** way to deal with such large amounts of data is to summarize it.
- Statistics is the method of summarization.

The Scientific Method

1. Define the question
2. Background research, observation
   – Have others tried to answer this earlier?
3. Formulate a hypothesis
   – If we do X, then Y will happen.
4. Design and run an experiment
5. Analyze the results
6. Communicate the results

• Experimental measurements are