Lecture 5
Introduction to Probability and Probability Distributions

Prof. Dr Saad Abed Madhi

1. What is Probability?

- **Probability** is a measure of how likely an event is to occur.
- It ranges from **0 to 1**:
  - 0 → impossible event
  - 1 → certain event
- **Formula (for an event A):**

$$P(A) = \frac{\text{Number of favorable outcomes}}{\text{Total number of possible outcomes}}$$

- **Example:**
- If 20 out of 100 patients recover from a treatment, the probability of recovery = 20/100 = 0.2

2. Types of Probability

- **Theoretical probability** – based on known outcomes (e.g., tossing a fair coin).
- **Empirical (observed) probability** – based on observed data (e.g., recovery rate in a study).
- **Subjective probability** – based on personal judgment or experience (e.g., likelihood of treatment success).

3. Random Variables

- A **random variable** is a quantity that can take different values based on chance.
- **Types:**
  - **Discrete:** Countable values (e.g., number of infected patients).
  - **Continuous:** Any value in an interval (e.g., blood pressure, cholesterol level).

4. Probability Distributions

- A **probability distribution** describes how probabilities are distributed over the possible values of a random variable.

5. Binomial Distribution

- Used for **discrete outcomes** with **two possible results** (success/failure) in **n independent trials**.
- **Parameters:**
  - n = number of trials
  - p = probability of success
  - q = 1 − p = probability of failure
- **Probability formula:**

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Mean and Variance:

$$\text{Mean} = \mu = np, \quad \text{Variance} = \sigma^2 = npq$$

- **Medical Example:**
- A clinical trial with 10 patients (n = 10), probability of recovery p = 0.7.
- Probability that exactly 7 recover:

$$P(X = 7) = \binom{10}{7} (0.7)^7 (0.3)^3$$

6. Poisson Distribution

- Used for discrete events that occur randomly over a fixed interval of time or space.
- Common for rare events.
- Parameter:
- λ = average number of events per interval
- Probability formula:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Mean and Variance:

$$\mu = \sigma^2 = \lambda$$

- **Medical Example:**
- On average, 2 patients visit ER per hour (λ = 2).
- Probability that exactly 3 patients visit in one hour:

$$P(X = 3) = \frac{e^{-2}2^3}{3!} = 0.180$$

7. Normal Distribution

- Continuous distribution; **bell-shaped curve**.
- Many biological measurements follow this distribution (e.g., blood pressure, height, cholesterol).
- **Parameters:**
- μ = mean
- σ = standard deviation
- **Probability density function (PDF):**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

- **Properties:**
- Symmetrical around mean μ
- 68% of data within $\pm 1\sigma$, 95% within $\pm 2\sigma$, 99.7% within $\pm 3\sigma$
- **Medical Example:**
- Adult systolic BP ~ $N(120, 15^2)$.
- Probability of SBP between 105 and 135?
- z-scores: $z1 = (105-120)/15 = -1$, $z2 = (135-120)/15 = 1$
- 68% probability.

8. Exponential Distribution

- Continuous distribution, often used for **time between events** (e.g., survival time, waiting time).
- **Parameter:**
- λ = rate of occurrence (1/mean)
- **Probability density function (PDF):**

$$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0$$

Mean and variance:

$$\text{Mean} = 1/\lambda, \quad \text{Variance} = 1/\lambda^2$$

- **Medical Example:**
- Time between arrivals of patients at a clinic is exponential with mean 5 minutes (λ = 1/5 = 0.2).
- Probability the next patient arrives within 3 minutes:

$$P(T \leq 3) = 1 - e^{-0.2 \times 3} \approx 0.451$$

9. Summary Table of Distributions

Distribution	Type	Parameters	Mean	Variance	Example in Medicine
Binomial	Discrete	n, p	$n \cdot p$	$n \cdot p \cdot q$	Patients recovering in a trial
Poisson	Discrete	λ	λ	λ	ER visits per hour
Normal	Continuous	μ, σ	μ	σ^2	Blood pressure, cholesterol
Exponential	Continuous	λ	$1/\lambda$	$1/\lambda^2$	Waiting time between events