CLASS NOTES

NORMAL DISTRIBUTION

Applied Mathematics — Probability

1. Continuous Random Variable

- A variable is CONTINUOUS if it can take ANY value in an interval (infinite possible values).
- Examples: height, weight, rainfall, time — anything measured, not counted.
- This is different from a DISCRETE variable (like number of heads in coin tosses), which takes only fixed, countable values.

2. Probability Density Function (PDF)

For a continuous random variable X, the function f(x) is called its Probability Density Function if:

- $f(x) \geq 0$ for all x (never negative)
- Total area under the curve of $f(x) = 1$ (100% probability)
- $P(a \leq X \leq b)$ = Area under f(x) between $x = a$ and $x = b$

In short: PDF turns AREA under a curve into PROBABILITY.

3. Normal Distribution (Gaussian Distribution)

- The most important continuous distribution in statistics.
- Its graph is called the NORMAL CURVE — a symmetric, bell-shaped curve.
- For a normal curve: Mean = Mode = Median (all three lie at the centre).

- The curve is symmetric about the vertical line through the mean (μ).
- It never touches the x-axis; it extends infinitely on both sides.
- Total area under a normal curve = 1 (always).

Normal Density Function

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

(defined for $-\infty < x < \infty$)

- μ (mu) = mean of X
- σ (sigma) = standard deviation of X
- These two parameters completely decide the shape and position of the curve.

4. Standard Normal Distribution

- A special case of normal distribution where $\mu = 0$ and $\sigma = 1$.
- Its random variable is written as Z, called the Z-SCORE.

Standard Normal Density Function

$$f(Z) = [1 / \sqrt{(2\pi)}] \cdot e^{[-Z^2 / 2]}$$

(defined for $-\infty < Z < \infty$)

Converting X to Z (very important formula)

$$Z = (X - \mu) / \sigma$$

- Z tells you HOW MANY standard deviations a value is away from the mean.
- $Z = 0$ → value is exactly at the mean
- $Z > 0$ → value is above the mean
- $Z < 0$ → value is below the mean

5. Cumulative Distribution Function F(Z)

- F(Z) = area under the standard normal curve to the LEFT of the line Z = z.

$F(Z) = P(Z < z)$

- Values of F(Z) are already calculated and given as a READY-MADE TABLE (Z-table) at the back of the book — we don't integrate every time.

Three types of probability we calculate using F(Z):

- Type 1: $P(Z < z) = F(z)$ → area to the LEFT of z
- Type 2: $P(z_1 < Z < z_2) = F(z_2) - F(z_1)$ → area BETWEEN z_1 and z_2
- Type 3: $P(Z > z) = 1 - F(z)$ → area to the RIGHT of z

Quick memory trick:
LEFT of z → use F(z) directly
RIGHT of z → subtract from 1
BETWEEN two values → subtract the two F values (bigger − smaller)

6. Properties of F(Z)

Property 1: $F(\infty) = 1$

- The entire area under the curve is 1, so at $Z = \infty$ the cumulative area becomes 1.

Property 2: $F(0) = 0.5$

- Since the curve is symmetric about Z = 0, exactly half the area (0.5) lies to the left of 0.

Property 3: $F(-z) = 1 - F(z)$

- Area to the left of $-z$ equals area to the right of z (by symmetry).
- Very useful — Z-tables often only give POSITIVE z values, so this property helps find F for negative z.

7. Area Under ANY Normal Curve (not just standard)

- Step 1: Convert X to Z using $Z = (X - \mu) / \sigma$
- Step 2: Look up $F(z)$ in the standard normal table
- Step 3: Apply the correct probability type (left / right / between)

8. Important Remark

- For a continuous random variable, including or excluding the endpoint does NOT change the probability:

$$P(X < a) = P(X \le a) \qquad P(X > b) = P(X \ge b)$$

- This is because a single exact point has ZERO area (zero probability) under a continuous curve.

9. Worked Example — Finding Z-scores

Example: Lengths (in cm) of 7 butterflies: 2, 2, 3, 2, 5, 1, 6

- Step 1 — Find Mean (μ): $\mu = (2+2+3+2+5+1+6)/7 = 21/7 = 3$
- Step 2 — Find Standard Deviation (σ):

$$\sigma = \sqrt{\left[\Sigma(x_i - \mu)^2 / n \right]}$$

- $\Sigma(x_i-\mu)^2 = 1+1+0+1+4+4+9 = 20$ $\qquad \sigma = \sqrt{(20/7)} = \sqrt{2.857} \approx 1.69$

- Step 3 — Find Z for each value using $Z = (x - \mu)/\sigma$:

Length	2	2	3	2	5	1
Z-score	−0.59	−0.59	0	−0.59	1.18	−1.18

(7th value, length 6: $Z = (6-3)/1.69 \approx 1.78$)

10. Quick Revision Summary

- Normal curve → symmetric, bell-shaped, mean = mode = median
- Standard normal → $\mu = 0$, $\sigma = 1$, variable called Z
- $Z = (X - \mu)/\sigma$ → converts any normal variable to standard form
- $F(z) = P(Z < z)$ → area to the left
- $F(\infty) = 1$, $F(0) = 0.5$, $F(-z) = 1 - F(z)$
- $P(Z>z) = 1-F(z)$ | $P(z_1<Z<z_2) = F(z_2)-F(z_1)$