

UNIT 4

Probability

(8 Hours) — Probability and Statistics (CACS202)

1. Definition of Probability

1.1 Basic Terms

Term	Meaning
Random experiment	Process whose outcome cannot be predicted with certainty (tossing a coin, rolling a die)
Sample space (S)	Set of all possible outcomes; e.g., for a die $S = \{1, 2, 3, 4, 5, 6\}$
Event	A subset of the sample space (e.g., getting an even number)
Mutually exclusive events	Cannot occur together: $A \cap B = \emptyset$
Independent events	Occurrence of one does not affect the probability of the other
Exhaustive events	Together cover the whole sample space
Complement (A^c)	Event that A does not occur; $P(A^c) = 1 - P(A)$

1.2 Definitions of Probability

- **Classical definition:** if an experiment has n equally likely, mutually exclusive and exhaustive outcomes of which m are favourable to event A , then $P(A) = m / n$.
- **Relative-frequency (empirical) definition:** $P(A) = \lim (\text{number of times } A \text{ occurs} / \text{number of trials})$ as trials $\rightarrow \infty$.
- **Axiomatic definition (Kolmogorov):** (i) $0 \leq P(A) \leq 1$; (ii) $P(S) = 1$; (iii) for mutually exclusive events, $P(A \cup B) = P(A) + P(B)$.

Thus probability is always between 0 (impossible event) and 1 (certain event), and $P(A) + P(A^c) = 1$.

Example 4.1 — Classical probability

Two fair dice are rolled. Find the probability that the sum is 7.

Solution

Total outcomes = $6 \times 6 = 36$. Favourable: (1,6), (2,5), (3,4), (4,3), (5,2), (6,1) = 6.

$P(\text{sum} = 7) = 6/36 = 1/6 \approx 0.1667$.

2. Two Basic Laws of Probability (without proof)

2.1 Addition Law

For any two events A and B: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. If A and B are **mutually exclusive**, $P(A \cap B) = 0$ and so $P(A \cup B) = P(A) + P(B)$. For three events: $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$.

Example 4.2 — Addition law

One card is drawn from a well-shuffled deck of 52. Find the probability that it is a king or a heart.

Solution

$P(\text{King}) = 4/52$, $P(\text{Heart}) = 13/52$, $P(\text{King and Heart}) = 1/52$.

$P(\text{King or Heart}) = 4/52 + 13/52 - 1/52 = 16/52 = 4/13 \approx 0.3077$.

2.2 Multiplication Law

For any two events: $P(A \cap B) = P(A) \cdot P(B | A) = P(B) \cdot P(A | B)$. If A and B are **independent**, $P(B | A) = P(B)$ and so $P(A \cap B) = P(A) \cdot P(B)$.

Example 4.3 — Multiplication law

Two cards are drawn one after the other from a deck of 52, (a) without replacement and (b) with replacement. Find $P(\text{both are aces})$.

Solution

(a) Dependent events: $P = (4/52) \times (3/51) = 12/2652 = 1/221 \approx 0.0045$.

(b) Independent events: $P = (4/52) \times (4/52) = 1/169 \approx 0.0059$.

2.3 Conditional Probability

The **conditional probability** of A given that B has already occurred is $P(A | B) = P(A \cap B) / P(B)$, provided $P(B) > 0$. It restricts the sample space to B. Events A and B are independent if and only if $P(A | B) = P(A)$.

Example 4.4 — Conditional probability

A die is rolled. Given that the number is greater than 3, find the probability that it is even.

Solution

$B = \{4, 5, 6\}$, so $P(B) = 3/6$. $A \cap B = \{4, 6\}$, so $P(A \cap B) = 2/6$.

$P(A | B) = (2/6)/(3/6) = 2/3$.

Bayes' theorem (a useful extension) reverses a conditional probability: $P(A_i | B) = P(A_i) P(B | A_i) / \sum P(A_j) P(B | A_j)$.

Example 4.5 — Bayes' theorem

Machines A, B, C produce 50%, 30% and 20% of the items, with 2%, 3% and 5% defective respectively. An item chosen at random is defective. What is the probability that it came from C?

Solution

$P(D) = 0.5 \times 0.02 + 0.3 \times 0.03 + 0.2 \times 0.05 = 0.010 + 0.009 + 0.010 = 0.029$

$P(C | D) = (0.2 \times 0.05) / 0.029 = 0.010 / 0.029 = 0.3448$.

3. Probability Distributions

3.1 Random Variable and Probability Distribution

A **random variable** X assigns a numerical value to each outcome of an experiment. It is **discrete** (countable values) or **continuous** (any value in an interval). A **probability distribution** lists the possible values with their probabilities. For a discrete variable, the **probability mass function** $p(x) = P(X = x)$ satisfies $p(x) \geq 0$ and $\sum p(x) = 1$; for a continuous variable the **probability density function** $f(x)$ satisfies $f(x) \geq 0$ and $\int f(x) dx = 1$, with probabilities given by areas under the curve. Key measures: **expected value** $E(X) = \sum x p(x)$ and **variance** $V(X) = E(X^2) - [E(X)]^2$.

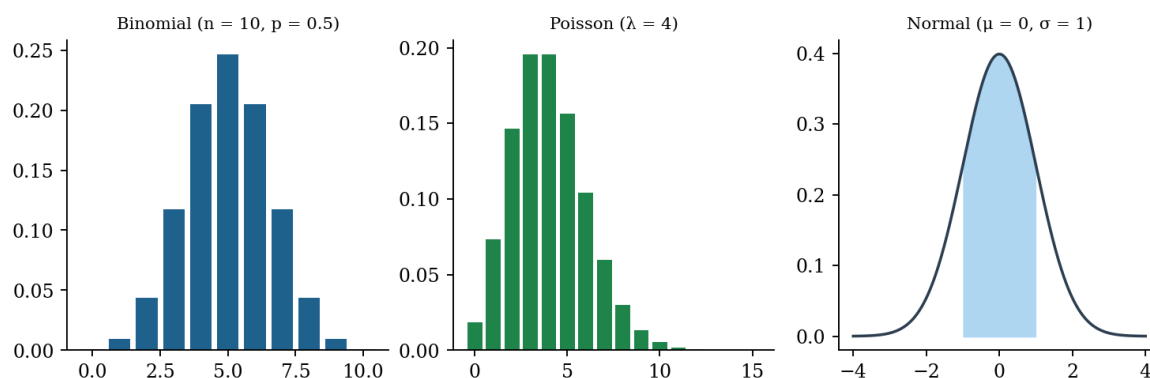


Figure 4.1 — Binomial, Poisson and normal distributions

3.2 Binomial Distribution

A **Bernoulli trial** has two outcomes, success (probability p) and failure ($q = 1 - p$). The binomial distribution applies when: (i) there is a fixed number n of trials; (ii) trials are independent; (iii) each trial has only two outcomes; (iv) p is constant. X = number of successes has

$$P(X = x) = {}^nC_x p^x q^{n-x}, x = 0, 1, 2, \dots, n$$

Property	Value
Mean	$\mu = np$
Variance	$\sigma^2 = npq$
Standard deviation	$\sigma = \sqrt{npq}$
Shape	Symmetric if $p = 0.5$; skewed right if $p < 0.5$, left if $p > 0.5$

Example 4.6 — Binomial

The probability that an item is defective is 0.1. In a random sample of 8 items find (a) exactly 2 defective, (b) at most 1 defective, (c) at least 1 defective, and the mean and variance.

Solution

$$n = 8, p = 0.1, q = 0.9.$$

$$(a) P(X = 2) = {}^8C_2 (0.1)^2 (0.9)^6 = 28 \times 0.01 \times 0.531441 = \mathbf{0.1488}$$

$$(b) P(X \leq 1) = P(0) + P(1) = (0.9)^8 + 8(0.1)(0.9)^7 = 0.4305 + 0.3826 = \mathbf{0.8131}$$

$$(c) P(X \geq 1) = 1 - P(0) = 1 - 0.4305 = \mathbf{0.5695}$$

$$\text{Mean} = np = 0.8; \text{Variance} = npq = 8 \times 0.1 \times 0.9 = \mathbf{0.72}$$

3.3 Poisson Distribution

The **Poisson distribution** models the number of occurrences of a rare event in a fixed interval of time, area or volume (calls per minute, defects per metre, accidents per day) when events occur independently at a constant average rate λ . It is also the limit of the binomial when n is large and p is small with $\lambda = np$ (usually $n \geq 20$, $p \leq 0.05$).

$$P(X = x) = e^{-\lambda} \lambda^x / x!, \quad x = 0, 1, 2, \dots \text{ (with } e \approx 2.71828)$$

Property	Value
Mean	$\mu = \lambda$
Variance	$\sigma^2 = \lambda$ (mean = variance)
Standard deviation	$\sigma = \sqrt{\lambda}$

Example 4.7 — Poisson

A help desk receives on average 3 calls per minute. Find the probability of (a) no call, (b) at least 2 calls, (c) exactly 5 calls in a given minute.

Solution

$$\lambda = 3, \quad e^{-3} = 0.049787.$$

$$(a) P(X = 0) = e^{-3} = \mathbf{0.0498}$$

$$(b) P(X \geq 2) = 1 - [P(0) + P(1)] = 1 - [0.0498 + 0.1494] = \mathbf{0.8009}$$

$$(c) P(X = 5) = e^{-3} 3^5 / 5! = 0.049787 \times 243 / 120 = \mathbf{0.1008}$$

3.4 Normal Distribution

The **normal distribution** is the most important continuous distribution. Its curve is a symmetric, bell-shaped curve determined by mean μ and standard deviation σ , with density

$$f(x) = [1 / (\sigma\sqrt{2\pi})] \exp[-(x - \mu)^2 / (2\sigma^2)], \quad -\infty < x < \infty$$

- Symmetric about $x = \mu$, so mean = median = mode; total area under the curve = 1.
- Points of inflection at $\mu \pm \sigma$; the curve is asymptotic to the x-axis.
- **Empirical rule:** about 68.3% of values lie within $\mu \pm \sigma$, 95.4% within $\mu \pm 2\sigma$, 99.7% within $\mu \pm 3\sigma$.
- The normal distribution is the limit of the binomial for large n , and underlies sampling theory (Central Limit Theorem).

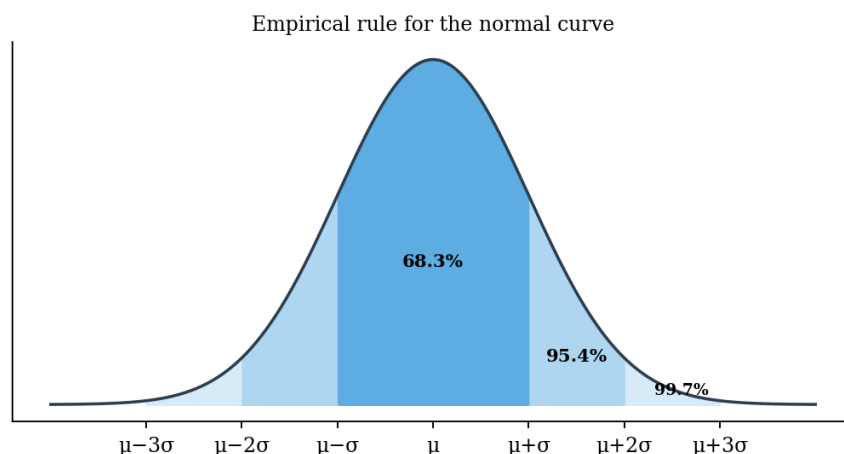


Figure 4.2 — Areas under the normal curve

Standard normal variate: any normal variable X can be converted to $Z = (X - \mu)/\sigma$, which has mean 0 and SD 1. Probabilities are then read from the standard normal table, which gives $P(0 \leq Z \leq z)$ or the cumulative area $P(Z \leq z)$. By symmetry, $P(Z < -z) = P(Z > z)$.

Example 4.8 — Normal distribution

Marks in an examination are normally distributed with mean 60 and standard deviation 10. Find (a) the probability that a student scores more than 75, (b) the probability of scoring between 50 and 70, and (c) the minimum marks needed to be in the top 10%.

Solution

- (a) $Z = (75 - 60)/10 = 1.50$. $P(X > 75) = P(Z > 1.50) = 1 - 0.9332 = \mathbf{0.0668}$
- (b) $Z_1 = (50 - 60)/10 = -1.00$, $Z_2 = (70 - 60)/10 = 1.00$. $P = 0.8413 - 0.1587 = \mathbf{0.6827}$ ($\approx 68.3\%$)
- (c) Top 10% means cumulative area 0.90 $\rightarrow Z = 1.28$. $X = \mu + Z\sigma = 60 + 1.28 \times 10 = \mathbf{72.8}$ marks

Example 4.9 — Choosing the right distribution

Which distribution is suitable for (a) number of heads in 10 tosses of a coin, (b) number of typing errors per page, (c) heights of adult males?

Solution

- (a) **Binomial** ($n = 10$, $p = 0.5$); e.g., $P(\text{exactly 3 heads}) = 0.1172$.
- (b) **Poisson** (rare events per unit of space).
- (c) **Normal** (continuous, symmetric measurement).

Key Takeaways

- ✓ Probability lies between 0 and 1; classical $P(A) = m/n$, empirical and axiomatic definitions also exist.
- ✓ Addition law: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Multiplication law: $P(A \cap B) = P(A)P(B | A)$.
- ✓ Conditional probability $P(A | B) = P(A \cap B)/P(B)$; for independent events $P(A \cap B) = P(A)P(B)$.
- ✓ Binomial: $P(x) = {}^nC_x p^x q^{n-x}$, mean np , variance npq . Poisson: $P(x) = e^{-\lambda} \lambda^x / x!$, mean = variance = λ .
- ✓ Normal: bell-shaped, standardize with $Z = (X - \mu)/\sigma$ and use the standard normal table.

— *End of Unit 4* —

bcalink.com