Bernoulli Distribution

1. Introduction & Intuition

- A **Bernoulli trial** has only two outcomes: *Success* or *Failure*.
- Define a random variable X such that

$$X = \begin{cases} 1 & \text{Success}, \\ 0 & \text{Failure}. \end{cases}$$

  So $X \in \{0, 1\}$.
- Let $p = P(\text{Success})$. Then $q = 1 - p = P(\text{Failure})$.

2. Probability Mass Function (PMF)

2.1 Piecewise form

$$P(X = x) = \begin{cases} p & x = 1, \\ 1 - p & x = 0, \\ 0 & \text{otherwise}. \end{cases}$$

2.2 Compact "exponent-switch" form

$$\boxed{P(X = x) = p^{x}\,(1-p)^{1-x}}, \qquad x \in \{0, 1\}.$$

Why it works (switch intuition):

$$\begin{aligned} &\text{If } x = 1: && p^1 = p, && (1-p)^0 = 1 && \Rightarrow p^x(1-p)^{1-x} = p. \\ &\text{If } x = 0: && p^0 = 1, && (1-p)^1 = 1 - p && \Rightarrow p^x(1-p)^{1-x} = 1 - p. \end{aligned}$$

3. Expectation (Mean)

$$\mathbb{E}[X] = \sum_{x \in \{0,1\}} x\,P(X = x) = 0 \cdot (1 - p) \;+\; 1 \cdot p = p.$$

Interpretation: on average a proportion p of all trials will be successes.

4. Variance

$$\operatorname{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

1. Compute $\mathbb{E}[X^2]$:
$$X^2 = 0^2 \cdot (1-p) + 1^2 \cdot p = p.$$
2. Subtract the square of the mean:
$$\mathrm{Var}(X) = p - p^2 = p(1-p) = pq.$$

Key point: Variance peaks at $p = 0.5$ (maximum uncertainty) and drops to 0 when p is at the extremes 0 or 1.

5. Quick Facts

- **Support:** $\{0, 1\}$.
- **Mode:** 1 if $p > 0.5$ (skew toward success), otherwise 0.
- **Skewness:** distribution is symmetric only when $p = 0.5$; otherwise it is skewed.

6. Worked Examples

1. **Spam detection**: Treat "email flagged correctly as spam" as Success. If the classifier is 95% accurate, then $X \sim \mathrm{Bernoulli}(p = 0.95)$.
2. **Rigged coin**: Coin shows Heads 70% of the time. Define Success = Heads $\Rightarrow X \sim \mathrm{Bernoulli}(0.7)$.

7. Visual Aids

- **Bar chart:** two bars at $x = 0$ and $x = 1$ with heights $(1-p)$ and p.
- **Variance curve:** plot of $p(1-p)$ against p (optional for teaching).
- *Classroom idea:* animate the bars while increasing p so students *see* how $1-p$ decreases by the same amount.

8. Summary

- $\mathrm{Bernoulli}(p)$ is the *building block* of many other discrete distributions (Binomial, Geometric, Negative Binomial, etc.).
- PMF: $P(X = x) = p^x(1-p)^{1-x}$.
- Mean: $\mathbb{E}[X] = p$. Variance: $\mathrm{Var}(X) = p(1-p)$.
- Highest randomness at $p = 0.5$; deterministic at $p = 0$ or $p = 1$.