Probability

R4 Cheng

August 13, 2025

Three Axioms of Probability

1. $0 \leq P(A) \leq 1$ for any event A.
2. $P(\Omega) = 1$ (Ω: sample space).
3. $A_1, A_2, \ldots$ are mutually exclusive events $\Rightarrow P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$.
   `> mutually exclusive:` $A_i \cap A_j = \emptyset$ `for` $i \neq j$.

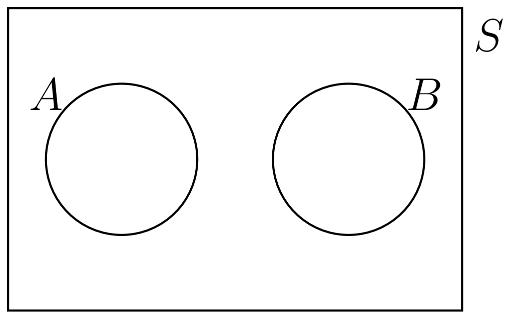
Derived from three axioms

- $P(\emptyset) = 0$ Empty set
- $P(A) + P(\bar{A} = 1)$: Complement
- $P(A) = P(A - B) + P(A \cap B)$: DeMorgan's Law
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$: Union and Intersection
- If $A \subseteq B \Rightarrow P(A) \leq P(B)$: Inclusion-Exclusion Principle
- Bool's Inequality: $P(A \cup B) \leq P(A) + P(B)$
- Bonferroni's Inequality: $P(A \cap B) \geq P(A) + P(B) - 1$ (not sure)

Events Relations and Probability Rules

- Dependent: A event is affected by another event.
- Independent: A event is not affected by another event.
- Mutually Exclusive: Two events cannot happen at the same time.

Mutually Exclusive in Venn diagram:



$$A \cap B = \emptyset$$

Independent Events

We usually use $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$ to check if two events are independent.

$$P(A \cap B) = P(A) \cdot P(B)$$

Counting Principles

Rule of Sum, Rule of Product, Permutations, Combinations

Rule of Product

Permutations

$$P(n, r) = n \cdot (n-1) \cdot \cdots \cdot (n-r+1) = \frac{n!}{(n-r)!}$$

Theorem: Number of permutations of n distinct objects arranged in a **circle** is $(n-1)!$.

Combinations

$$C(n, r) = \frac{P(n, r)}{r!} = \frac{n!}{r!(n-r)!}$$

Finding Probability

Generally, we need to list and find sample space, then count the number of favorable outcomes and find the probability.

If the experiment meets some attributes, we can use the following methods:

Binomial Distribution

If the experiment meets the following conditions:

- The experiment is repeated n times.
- Each trial has two outcomes: success or failure.
- The probability of success is p and the probability of failure is $q = 1 - p$.
- The trials are independent.

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

$$\text{Mean: } E[X] = np$$

$$\text{Variance: } \text{Var}(X) = np(1-p)$$

> Binomial distribution is sum of independent Bernoulli trials.

Bernoulli Distribution

Binomial distribution with $n = 1$.

Hypergeometric Distribution

If the experiment meets the following conditions:

- N items are in the experiment
- Get n items from N without replacement.
- $N = a + (N - a)$ where a is the number of success items and $N - a$ is the number of failure items.
- The probability of success

$$P(X = k) = \frac{\binom{a}{k}\binom{N-a}{n-k}}{\binom{N}{n}}$$

$$\text{Mean: } E[X] = n \cdot \frac{a}{N}$$

$$\text{Variance: } \text{Var}(X) = n \cdot \frac{a}{N} \cdot \frac{N-a}{N} \cdot \frac{N-n}{N-1}$$

Poisson Probability Distribution

The Poisson distribution is used to describe the probability of rare events occurring within a **specific time or area**.

If the experiment meets the following conditions:

- The number of events in a given interval of time or space is counted.
- The rate of occurrence of the event is constant.
- The occurrence of one event does not affect the occurrence of another event.

$$P(X = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad \text{for} \quad x = 0, 1, 2$$

where λ is the average number of events occurring in a unit interval of time or space (e.g., per unit time or per unit area)

> $e = 2.718$

$$\text{Mean: } E(x) = \mu = \lambda t$$

$$\text{Variance: } \text{Var}(x) = \lambda t = \text{mean}$$

> Poisson distribution can be used as an approximation to the binomial distribution when n is large and p is small, where $\mu = np$.