Contents

1	Probability Theory			1
	1.1	Elema	antary Probability	1
		1.1.1	Basic principles of Counting	1
		1.1.2	Sample Points, Sample Space and σ -Field	2
		1.1.3	Probability Measures and Spaces	2
	1.2 Conditional Prob		tional Probability	2
		1.2.1	Independence of Events	3
		1.2.2	Bayes' Theorem	
	1.3	Discre	ete Random Variables	
		1.3.1	PDF and CDF	4
		1.3.2	Bernoulli Random Variable	4

1 Probability Theory

1.1 Elemantary Probability

Definition 1.1.1 (Cardano's Principle). A be a random outcome of an experiment that may proceed in various ways. Assume each of these ways is **equally** likely, then, probability P[A] of outcome A is

$$P[A] = \frac{number\ of\ ways\ leading\ to\ outcome\ A}{number\ of\ ways\ the\ experiment\ can\ be\ proceeded} \tag{1}$$

1.1.1 Basic principles of Counting

• Permutation:

$$A_n^k = \frac{n!}{(n-k)!} \tag{2}$$

• Combination:

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \tag{3}$$

• Permutation of k Indisguishable Objects:

$$\frac{n!}{n_1! n_2! n_3! \cdots n_k!} = \binom{n}{n_1} \cdot \binom{n-n_1}{n_2} \cdots \binom{n-(n_1+n_2+\cdots+n_{k-1})}{n_k}$$
(4)

Remark:

In permutation of k indistinguishable objects, the elements has noorder within a group but are different from each other; the groups either has order or are different from each other.

Example:

Consider 10 balls, 5 red, 3 green and 2 blue. How many ways can they be arranged on a line?

$$\frac{10!}{5! \cdot 3! \cdot 2!} = 2520 \tag{5}$$

1.1.2 Sample Points, Sample Space and σ -Field

Definition 1.1.2 (Sample Points). Mathematical objects are called sample points.

Definition 1.1.3 (Sample Space). The sample space S is large enough to accommodate all the sample points.

Definition 1.1.4 (Event). An outcome in the sense of Cardano's principle is interpreted as a subset A of a sample space S abd called an event.

Definition 1.1.5 (Mutually exclusive). Two events A_1 , A_2 are called mutual exclusive if $A_1 \cap A_2 = \emptyset$

Definition 1.1.6 (σ -Field). Suppose that a non-empty set S is given. A σ -field \mathcal{F} on S is a family of subsets of S such that:

- $\bullet \ \emptyset \in \mathcal{F}$
- If $A \in \mathcal{F}$, then $S \setminus A \in \mathcal{F}$
- If $A_1, A_2, \dots \in \mathcal{F}$ is a finit sequence of subsets, then the union $\cup_k A_k \in \mathcal{F}$

1.1.3 Probability Measures and Spaces

Definition 1.1.7 (Probability Measure). Let S be the sample space and \mathcal{F} be a σ -field. Then a function

$$P: \mathcal{F} \to [0, 1], A \to P[A], \tag{6}$$

is called probability measure / probability function on S if

- P[S] = 1
- For any set of events $A_k \subset mathcal F$ such that $A_j \cap A_k = \emptyset$ for $j \neq k$,

$$P[\cup_k A_k] = \sum_k a_n P[A_k] \tag{7}$$

Theorem 1.1.1 (Basic Properties). $P[A_1 \cup A_2] = p[A_1] + P[A_2] - P[A_1 \cap A_2]$

1.2 Conditional Probability

Definition 1.2.1 (Conditional Probability). B occurs given that A has occured

$$P[B \mid A] := \frac{P[A \cap B]}{P[A]} \tag{8}$$

1.2.1 Independence of Events

Definition 1.2.2 (independent). Two events are *independent* if

$$P[A \cap B] = P[A]P[B] \tag{9}$$

equivalent to

$$P[A \mid B] = P[A], P[B] \neq 0$$
 (10)

Definition 1.2.3 (Total Probability).

$$P[B] = P[B \mid A_1] * P[A_1] + \dots + P[B \mid A_n]P[A_n] = \sum_{k=1}^{n} P[B \mid A_k] \cdot P[A_k]$$
 (11)

is called total propability formula for P[B]

1.2.2 Bayes' Theorem

Theorem 1.2.1 (Bayes's Theoremm). Let $A_1, \dots, A_n \subset S$ be a set of pairwise mutually exclusive events whose union is S and who each have non zero probability of occurring. Let $B \subset S$ be any events such that $P[B] \neq 0$. Then for any $A_k, k = 1, \dots, n$

$$P[A_k \mid B] = \frac{P[B \cap A_k]}{P[B]} = \frac{P[B \mid A_k] \cdot P[A_k]}{\sum_{j=1}^n P[B \mid A_j] \cdot P[A_j]}$$
(12)

1.3 Discrete Random Variables

Definition 1.3.1 (Random Variable). Such function X

$$X: S \to \mathbb{R} \tag{13}$$

is a random variable. X has numerical values that are derived from the outcome of a random experiment.

Two types:

 \bullet Discrete $Random\ Variables:$ countable range in $\mathbb R$

Definition 1.3.2 (Discrete Random Variable). Let S be sample space, Ω a countable subset of \mathbb{R} . A discrete random variable is a map

$$X: S \to \Omega$$
 (14)

$$f_X: \Omega \to \mathbb{R}$$
 (15)

A random variable is often given by the pair (X, f_X)

• Continuous Random Variables: having a range equal to \mathbb{R}

Example:

Flip a coin three times, sample space can be given by :

$$S = (t, t, t), (t, t, h), (t, h, t), \cdots, (h, h, h)$$
(16)

Then we can define X as follows:

$$X(t,t,t) = 0, X(t,t,h) = 1, \dots, X(h,h,h) = 3$$
 (17)

X denotes the number of heads

$$P[X = 1] = P[(t, t, h), (t, h, t), (h, t, t)]$$
(18)

We can write

$$P[X = x] = P[A] \tag{19}$$

where $x \in R$ and $A \subset S$ is the event containing all sample points p such that X(p) = x.

1.3.1 PDF and CDF

Random variable comes with a probability density function or probability distribution f_X that allows the calculation of probability directly.

- $f_X(x)>0$ for all x
- $\sum_{X \in \Omega} f_X(x) = 1$

Definition 1.3.3 (Cumulative distribution function). Cumulative distribution function of a random variable is defined as

$$F_X : \mathbb{R} \to \mathbb{R}, F_X(x) := P[X \le x]$$
 (20)

For discrete random variable, $F_X(x) = \sum_{y \le x} f_X(y)$

1.3.2 Bernoulli Random Variable

Definition 1.3.4 (Bernoulli Trial). Consider an experiment can only results in two possible outcomes, such as success or failure. and probability of success 0

Definition 1.3.5 (Bernoulli Random Variable). Let S be a sample space and $X: S \to 0, 1 \in \mathbb{R}$, Let 0 < P < 1, then define the density function

$$f_X:0,1\to R \quad f_X(x) = \begin{cases} 1-p & for \ x=0\\ p & for \ x=1 \end{cases}$$
 (21)

Then X is said to be a *Bernoulli random variable* or follow a *Bernoulli distribution* with parameter p. We indicate this by writing

$$X \sim Bernoulli(p)$$
 (22)