Note

• experiment : the process of obtaining an observed result of some phenomenon.
• trial : a performance of an experiment
• outcome : observed result

Definition. *The set of all possible outcomes of an experiment is called the sample space, denoted by S.*

Definition. *If a sample space S is either finite or countably, then it is called a discrete sample space. Otherwise, is called a continuous sample space.*

Definition. *An event is a subset of the sample space S. If A is an event, then "A occurred" if "A contains the out come that occurred"*

Definition. *An event is called and elementary event if it contains exactly one outcome of the experiment.*

Definition.

• *Two events A and B are called mutually exclusive if $A \cap B = \emptyset$*
• *Events $A_1, A_2, A_3, \cdots$ are said to be mutually exclusive if they are pairwise mutually exclusive. That is, if $A_i \cap A_j = \emptyset$ whenever $i \neq j$.*

Definition. *For a given experiment, S denotes the sample space and $A_1, \cdots$ represent possible events. A set function that associates a real value $P(A)$ with each event A is called a probability set function, and $P(A)$ is called the probability of A, if the following properties are satisfied:*

i) $0 \leq p(A)$ *for every* A
ii) $P(S) = 1$
iii) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ *if $A_1, \cdots$ are pairwise mutually exclusive events.*

Theorem.

• $P(A) = 1 - P(A')$
• $P(A) \leq 1$, *for any event* A
• *For any two event A and B,* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- *If* $A \subset B$*, then* $P(A) \leq P(B)$
- $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$ *If* $A_1, \cdots$ *is a sequence of events*
- *If* $A_1, A_2, \cdots, A_k$ *are events, then* $P(\bigcap_{i=1}^{k} A_i) \geq 1 - \sum_{i=1}^{k} P(A_i')$

<u>Definition</u>. *The conditional probability of an event A, given the event B, is defined by*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

if $P(B) \neq 0$

Theorem. *For any events A and B,*

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

Theorem. *If $B_1, \cdots, B_k$ is a collection of mutually exclusive and exhaustive events, then for any event A*

$$P(A) = \sum_{i=1}^{k} P(B_i)P(A \mid B_i)$$

Note. exhaustive events: A collection of event which union is sample space.

Theorem. *If $B_1, \cdots, B_k$ is mutually exclusive and exhaustive events, then for any event A and each $j = 1, \cdots, k$*

$$P(B_j \mid A) = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{k} P(B_i)P(A \mid B_i)} \left(= \frac{P(A \cap B_j)}{P(A)} \right)$$

<u>Definition</u>. *Two events A and B are called independent events if*

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

Otherwise, A and B are called dependent event

Theorem. *If A and B are events such that $P(A) > 0$ and $P(B) > 0$, and A and B are independent, we get*

$$P(A \cap B) = P(A)P(B) \Leftrightarrow P(A \mid B) = P(A) \Leftrightarrow P(B \mid A) = P(B)$$

Theorem.

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ \Leftrightarrow P(A' \cap B) &= P(A')P(B) \\ \Leftrightarrow P(A \cap B') &= P(A)P(B') \\ \Leftrightarrow P(A' \cap B') &= P(A')P(B') \end{aligned}$$

Definition. *The k events $A_1, \cdots, A_k$ are said to b independent or mutually independent if for every $j = 2, 3, \cdots, k$ and every subset of distinct indices $i_1, i_2, \cdots, i_j$*

$$P(A_{i1} \cap A_{i2} \cap \cdots \cap A_{ij}) = P(A_{i1})P(A_{i2}) \cdots P(A_{ij})$$

Definition. *A random variable, say X, is a function defined over a sample space S, that associated a real number with each possible outcome in S*

$$X(e) = x, \text{ where } e \in S$$

Definition. *If the set of all possible values of a random variable, X, is a countable set, $x_1, x_2, \cdots, x_n$, or $x_1, \cdots$, then X is called a discrete random variable.*
The function

$$f(x) = P[X = x] \; x = x_1, x_2, \cdots$$

that assigns the probability to each possible value x will be called the discrete probability density function (discrete pdf).

Property. $f(x_i) \geq 0, \; \sum_{\text{all } x_i} f(x_i) = 1$

Definition. *The cumulative distribution function(CDF) of a random variable X is defined for any real x by*

$$F(x) = P[X \leq x]$$

Theorem. *Let X be a discrete random variable with pdf $f(x)$ and CDF $F(x)$. If the possible values of X are indexed in increasing order, $x_1 < x_2 < x_3 < \cdots$, then:*

(i) $f(x_1) = F(x_i)$
(ii) for any $i > 1,\ f(x_i) = F(x_i) - F(x_{i-1})$
(iii) if $x < x_1$ *then* $F(x) = 0$
(iv) $F(x) = \sum_{x_i \le x} f(x_i)$

Theorem. *A function $F(x)$ is a CDF for some random variable X if and only if it satisfies:*

(i) $\lim_{x\to-\infty} F(x) = 0$
(ii) $\lim_{x\to\infty} F(x) = 1$
(iii) $\lim_{h\to0^+} F(x+h) = F(x)$
(iv) $a < b$ *implies* $F(a) \le F(b)$

Definition. *If X is a discrete random variable with pdf $f(x)$, then the expected value of X is*

$$E(X) = \sum_x x f(x)$$

Definition. *A random variable variable X is called a continuous random variable if there is a function $f(x)$, called the probability density function of X, such that the CDF can be represented as*

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Properties A function $f(x)$ is a pdf for some continuous random variable X if and only if it satisfies:

(i) $f(x) \ge 0\ \forall x$
(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

Definition. *If X is a continuous random variable with pdf $f(x)$, then the expected value of X is defined by*

$$E(X) = \int_{-\infty}^{\infty} x f(x)dx$$

if it's absolutely convergent. Otherwise we say $E(X)$ does not exist.

Properties.

• If X is a random variable with pdf $f(x)$ and $u(x)$ is a real-valued function whose domain includes the possible values of X, then

$$E[u(X)] = \sum_x u(x)f(x) \text{ if } X \text{ is discrete}$$

$$E[u(X)] = \int_{-\infty}^{\infty} u(x)f(x)dx \text{ if } X \text{ is continuous}$$

Note : we can consider $u(X)$ is a new random variable, if $u(x)$ is not one-to-one, $P[u(X) = u(x_1)] \neq P[X - x_1]$, but $P[u(X) = u(x')] = \sum P[X = x_i]$ where $u(x_i) = u(x')$

• If X is a random variable with pdf $f(x)$, a and b are constants, $g(x)$ and $h(x)$ are real-valued functions whose domains include the possible values of X, then

$$E[ag(X) + bh(X)] = aE[g(X)] + bE[h(X)]$$

Note: regard $ag(X) + bh(X)$ as $u(X)$, we can use the above properties to proof it.

Definition. *The variance of a random variable X is given by*

$$Var(X) = E[(X - \mu)]$$

where $\mu = E(X)$

Note.

- kth moment about the origin of a random variable X is

$$\mu'_k = E(X^k)$$

- and kth moment about the mean is

$$\mu_k = E[X - E(X)]^k = E(X - \mu)^k$$

Theorem.

- $Var(X) = E(X^2) - E(X)^2$

Note: Consider $X^2, 1$ as random variable of S and $E(X)$ as constant

- $Var(aX + b) = a^2 \, Var(X)$

Theorem. $P[u(X) \geq c] \leq \dfrac{E[u(X)]}{c}$

Theorem. $P[|X - \mu| \geq k\sigma] \leq \dfrac{1}{k^2}$