Random experiments and Sample space

- Random experiment : Experiment involving randomness
  - Coin toss
  - Roll a dice
  - Pick a number at random from $[0, 1]$.
- Sample space Ω: set of all possible outcomes of the random experiment. It could be a finite or infinite set.
  - $\Omega_c = \{H, T\}$
  - $\Omega_d = \{1, 2, \ldots, 6\}$
  - $\Omega_u = [0, 1]$
  - A subset $A \subseteq \Omega$ is called an **event**.
  - Probability of event A is denoted by $\mathbb{P}(A)$.

Probability theory

$\{$Random experiment, Sample space, Events$\}$ are the key ingredients in probability theory.

In probability theory, we are interested in **measuring** the probability of subsets of Ω (events).

Probability measure $\mathbb{P}$ is a **set function**, i.e. it acts on sets and measures the probability of such sets.

sigma-algebra as domain for $\mathbb{P}$

- Event space or *sigma-algebra* $\mathcal{F}$ is a collection of subsets of Ω that satisfy
  $\bullet\ \emptyset \in \mathcal{F} \quad \bullet\ A \in \mathcal{F} \implies A^c \in \mathcal{F}$
  $\bullet A_1, A_2, \ldots A_n, \ldots \in \mathcal{F} \implies \cup_{n=1}^{\infty} A_n \in \mathcal{F}$

- The $\sigma-$algebra is said to be closed under formation of compliments and countable unions.

- It is also closed under the formation of countable intersections

> When Ω is countable and finite, we will consider power-set $\mathcal{P}(\Omega)$ as the domain.

Formal definition of Probability measure $\mathbb{P}$

Definition

A probability measure $\mathbb{P}$ on the *measurable space* $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ s.t.

1. $\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega) = 1$
2. For a disjoint collection of event sets $A_1, A_2, \ldots$ from $\mathcal{F}$ we have
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$
(countable additivity)

- The trio $(\Omega, \mathcal{F}, \mathbb{P})$ is called as a probability space.

Conditional probability

- Given/If dice rolls odd, what is the probability that the outcome is 1?
- The conditional probability of event B given event A is defined as $\mathbb{P}(B/A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$ when $\mathbb{P}(A) > 0$.
- Easy to see that $P(A/B)P(B) = P(B/A)P(A)$.
- This gives the Bayes rule: $P(B/A) = \frac{P(A/B)P(B)}{P(A)}$.

Law of total probability: Let $B_1, B_2, \ldots B_n$ be the partition of the sample space Ω. Then for any event A we have

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A/B_i)P(B_i).$$

Independence & Mutually Exclusive

- Two events A, B are independent iff $P(A/B) = P(A)$ and $P(B/A) = P(B)$.
- Two events A, B are mutually exclusive iff $P(A \cap B) = 0$.

Random variable

- Given a random experiment with associated $(\Omega, \mathcal{F}, \mathbb{P})$, it is sometimes difficult to deal directly with $\omega \in \Omega$. eg. rolling a dice ten times.
- Notice that each sample point $\omega \in \Omega$ is not a number but a sequence of numbers.
- Also, we may be interested in functions of these sample points rather than samples themselves. eg: Number of times 6 appears in the 10 rolls.
- In either case, it is often convenient to work in a new *simpler* probability space rather than the original space.
- Random variable is a device which precisely helps us make this mapping from Ω to a 'simpler' Ω', i.e. $X : \Omega \to \Omega'$

Random variable

- If Ω' is countable, then the random variable is called a discrete random variable.
- If $\Omega' \subseteq \mathbb{R}$ or uncountable, then the random variable is a continuous random variable.
- Notation: Random variables denoted by capital letters like X, Y, Z etc. anf their realizations by small letters x, y, z..

PMF and CDF of a Discrete r.v.

- Let $X : \Omega \to \Omega'$ be a discrete r.v.
- Let $p_X(x)$ for $x \in \Omega'$ denote the probability that X takes the value x.
- $p_X(x)$ is called as a probability mass function.
- The cumulative distribution function (CDF) $F_X(\cdot)$ is defined as $F_X(x_1) := \sum_{x \leq x_1} p_X(x) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq x_1\}$.

Expectation, Moments, Variance

- The mean or expectation of a random variable X is denoted by $E[X]$ and is given by $E[X] = \sum_{x \in \Omega'} x p_X(x)$.
- The n^{th} moment of a random variable X is denoted by $E[X^n]$ and is given by $E[X^n] = \sum_{x \in \Omega'} x^n p_X(x)$.
- Functions of random variables are random variables.
- For a function $g(\cdot)$ of a random variable X, its expectation is given by $E[g(X)] := \sum_{x \in \Omega'} g(x) p_X(x)$
- $Var(X) := E[(X - E[X])^2] = E[X^2] - E[X]^2$
- Linearity of Expectation: For $Y = aX + b$, we have $E[Y] = aE[X] + b$.

Examples of random variable

- Bernoulli random variable $X = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{otherwise.} \end{cases}$
- Basic models of Multi-arm bandit problem assume Bernoulli Bandits. $E[X] = p, E[X^n] = p$.
- X is a Binomial variable with parameters n and p if $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k \in \{0, \dots, n\}$.
- Some other discrete r.v.: Poisson, Geometric

Continuous random variables

- The cdf $F_X(x)$ is defined as $F_X(x) := \mathbb{P}\{\omega \in \Omega : X(\omega) \leq x\}$.
- A random variable X is continuous if there exists a non-negative real valued probability density function (PDF) $f_X(\cdot)$ such that $F_X(x) = \int_{u=-\infty}^{x} f_X(u)du$.
- $P_X(B) = \int_{u \in B} f_X(u)du$. $P_X(\mathbb{R}) = \int_{u=-\infty}^{\infty} f_X(u)du = 1$.
- $P_X(X = a) = 0$. (no mass at any point)

$$\frac{dF_X(x)}{dx} = f_X(x) \text{ or } P_X(x < X \leq x + h) \simeq f_X(x)h.$$

Mean, Variance, Moments

- $E[X] = \int_{-\infty}^{\infty} u f_X(u) du$
- $E[X^n] = \int_{-\infty}^{\infty} u^n f_X(u) du$
- $E[g(X)] = \int_{-\infty}^{\infty} g(u) f_X(u) du$
- $Var[X] = E[g(X)]$ where $g(x) = (x - E[X])^2$.
- For $Y = aX + b$, $E[Y] = aE[X] + b$.
- For $Y = aX + b$, $F_Y(y) = F_X(\frac{y-b}{a})$ when $a \geq 0$.
- For $Y = aX + b$ and $a < 0$, $F_Y(y) = 1 - F_X(\frac{y-b}{a})$.

List of Probability distributions ...

Uniform, Exponential(λ), Gaussian $\mathcal{N}(\mu, \sigma^2)$

Other interesting ones are Beta, Gamma, Erlang, Logistic, Weibull.

`https://en.wikipedia.org/wiki/List_of_probability_distributions`

Summary: Multiple random variables

$p_{XY}(x,y) := \mathbb{P}\{\omega \in \Omega : X(w) = x \text{ and } Y(\omega) = y\}$.
$F_{XY}(x,y) := \mathbb{P}\{\omega \in \Omega : X(w) \leq x \text{ and } Y(\omega) \leq y\}$.

The marginal PMF's p_X and p_Y can be obtained from the joint PMF as follows:
$p_X(x) = \sum_y p_{XY}(x,y)$ and $p_Y(y) = \sum_x p_{XY}(x,y)$.

Two random variables, X and Y are independent if the following is true:
$p_{XY}(x,y) = p_X(x)p_Y(y), F_{XY}(x,y) = F_X(x)F_Y(y)$ and $E[XY] = E[X]E[Y]$.

$E[g(X,Y)] = \sum_{xy} g(xy)p_{XY}(xy)$

The rules for more than 2 discrete random variables are similar.

Summary for Continuous random variable

- $f_{XY}(x, y)$ denotes the joint pdf for X and Y.
- $F_{XY}(x, y) := \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(s, t) ds dt$. $f_{X,Y}(x, y) := \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$.

The marginal pdf's f_X and f_Y can be obtained from the joint PDF as follows:
$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx$

Two random variables, X and Y are independent if the following is true:
$f_{XY}(x, y) = f_X(x) f_Y(y), F_{XY}(x, y) = F_X(x) F_Y(y)$ and $E[XY] = E[X]E[Y]$.

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) dx dy$$

- Rules similar for more than 2 random variables.