Chapter 1

Probability

1.1 Introduction

Consider an *experiment*, result of which is random, and is one of the finite number of *outcomes*.

Example 1. Examples of experiments and possible outcomes:

Experiment	Outcomes
Toss a coin	$\{\mathrm{H}, \mathrm{T}\}$
Roll a dice	$\{1, 2, 3, 4, 5, 6\}$
Toss two coins	$\{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$

Let $\Omega = \{w_1, w_2, \ldots, w_n\}$ be the set of outcomes of an experiment, called as *sample space*.

To each outcome, a probability P is assigned such that $P(w_i) \in [0, 1]$ and $\sum_1^n P(w_i) = 1$.

Example 2. A *fair* coin is tossed. By fair coin, one means that the probability of getting H is same as that of getting T. Thus $P(\mathrm{H}) = P(\mathrm{T}) = 1/2$.

Here the meaning of the probability is in *frequency* sense. There are two ways of interpreting this.

- ▷ The experiment is performed N times. Let n_i is the number of times outcome w_i appears. Then

$$\lim_{N \to \infty} \frac{n_i}{N} = P(w_i)$$

- ▷ Alternatively, N identical experiments are performed simultaneously. Such collection of experiments is called *ensemble*.

Example 3. A jar contains 6 yellow, 4 white and 7 blue balls. The balls, except color, are identical. Experiment is to find the color of a ball picked randomly. Sample space is $\{\text{yellow}, \text{white}, \text{blue}\}$. The probabilities are

$$\begin{aligned} P(\text{yellow}) &= \frac{6}{17} \\ P(\text{white}) &= \frac{4}{17} \\ P(\text{blue}) &= \frac{7}{17} \end{aligned}$$

An event is a subset of the sample space. Let A be an event. Then the probability of A is defined as

$$P(A) = \sum_{w \in A} P(w).$$

Example 4. A fair dice is rolled. What is the probability that outcome is less than or equal to 3? Then this event is $A = \{1, 2, 3\}$. Then

$$P(A) = P(1) + P(2) + P(3) = \frac{1}{2}.$$

Theorem 5. *Probabilities assigned to events in sample space Ω satsify following properties*

1. $P(\phi) = 0$ *and* $P(\Omega) = 1$.
2. *If* $A \subset B \subset \Omega$ *then* $P(A) \leq P(B)$.
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
4. $P(\bar{A}) = 1 - P(A)$

A physicist deals with measurable quantities which are represented by real numbers. Thus, in the rest of this chapter, sample spaces will be assumed to be numerical.

1.2 Discrete Probability Distributions

Definition 6. Consider an experiment whose outcome depends on chance. We represent the outcome of the experiment by a capital Roman letter, such as X, called a *random variable*. The *sample space* of the experiment is the set of all possible outcomes. If the sample space is either finite or countably infinite, the random variable is said to be discrete.

The assignment of probabilities to outcomes is described next.

Definition 7. Let X be a random variable which denotes the value of the outcome of a certain experiment, and assume that this experiment has only finitely many possible outcomes. Let Ω be the sample space of the experiment. A distribution function for X is a real-valued function P whose domain is Ω and which satisfies:

1. $P(\omega) \geq 0$ for all $\omega \in \Omega$;
2. $\sum_{\omega \in \Omega} P(\omega) = 1$.

Example 8. Consider four magnetic dipoles which can be either in *up* or *down* state. Magentic moment of a dipole is $+\mu$ when it is in up state and is $-\mu$ when it is in down state. The experiment is to find total magnetism given by the sum of all dipole moments, that is

$$M = \sum_{i=1}^{4} \mu_i.$$

Now, let up state be represented by letter u and down state by d. Then, the configuration $uudd$ means that the first dipole is up, the second is up, third and fourth are down. And magnetization of this configuration is 0. There are total 16 possible configurations. The sample space is set of all possible values for total magnetism. Let

$$\Omega_M = \{-4\mu, -2\mu, 0, 2\mu, 4\mu\}.$$

Now if all configuration are equally likely to occur, then

configurations	# of Arrangements	Magnetization	Probability
$uuuu$	1	4μ	1/16
$uuud$	4	2μ	1/4
$uudd$	6	0	3/8
$uddd$	4	-2μ	1/4
$dddd$	1	-4μ	1/16

Normally, it is not interesting to know the detailed probability distribution. Usually the *moments* of the distribution are quoted. Two interesting moments are average and variance.

Definition 9. Let X be a discrete random variable with sample space Ω and distribution funtion P, then the expectation value of X, denoted by $E(X)$ or $\langle X \rangle$, is defined as

$$E(X) = \sum_{\omega \in \Omega} \omega P(\omega)$$

provided the sum converges.

The expectation value is interpreted again in frequecy sense. If the experiment is performed large number of times, then the *sample mean* is close to the expectation value.

Example 10. In magnetization example above, the exectation value of the magnetization is given by

$$\begin{aligned} E(M) &= (-4\mu)\frac{1}{16} + (-2\mu)\frac{1}{4} + (0)\frac{3}{8} + (2\mu)\frac{1}{4} + (4\mu)\frac{1}{16} \\ &= 0 \end{aligned}$$

Theorem 11. *Let X be a discrete random variable with sample space Ω and distribution funtion P. Let $\phi : \omega \to \mathbb{R}$. Then the expectation value of $\phi(X)$ is*

$$E(\phi(X)) = \sum_{\omega \in \Omega} \phi(\omega) P(\omega)$$

provided the sum converges.

Example 12. Again in magnetization example,

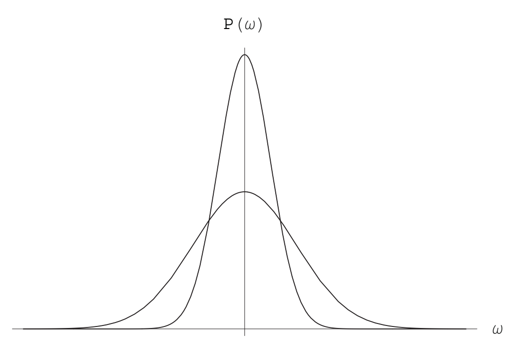
$$\begin{aligned} E\left(M^2\right) &= (-4\mu)^2 \frac{1}{16} + (-2\mu)^2 \frac{1}{4} + (0)^2 \frac{3}{8} + (2\mu)^2 \frac{1}{4} + (4\mu)^2 \frac{1}{16} \\ &= 4\mu^2 \end{aligned}$$

Definition 13. Let X be a discrete random variable with sample space Ω and distribution funtion P. Let $E(X) = \mu$. The variance of X, denoted by $V(X)$, is defined as

$$\begin{aligned} V(X) &= E\left((X-\mu)^2\right) \\ &= \sum_{\omega \in \Omega} (\omega - \mu)^2 P(\omega) \end{aligned}$$

The standard deviation of X, denoted by σ_X, is defined as $\sqrt{V(X)}$.

The variance is a measure of the spread of the the probability distributions about the mean value. In the figure below, the red probability distribution has larger variance.



Theorem 14. *If X is a discrete random variable, then*

$$V(X) = E\left(X^2\right) - (E(X))^2$$

Example 15. In magnetization example, $E(M) = 0$. Then, $V(M) = E(M^2) = 4\mu^2$.

1.3 Continuous Probability Distribution

A random variable is said to be continuous if the sample space is continuous, we will assume that the sample is either $\mathbb{R}$ or intervals in $\mathbb{R}$.

Example 16. Suppose a particle is trapped in a 1D box (say in $[0, L]$), and is going back and forth with speed u. It was set in motion in past from some random location in the box. The experiment is to find the position of particle at some time. Let $P(X < x)$ be the probability of finding particle to the left of x. Then

$$\begin{aligned} P(X < x) &= 0 \quad x < 0 \\ &= 1 \quad x > 1 \\ &= x/L \quad \text{otherwise.} \end{aligned}$$

This is intutive, since all points in box are equally likely, probability that particle is in some interval $[a, b] \subset [0, L]$ is proportional to length $b - a$. Then, probability of *finding the particle at a point* becomes 0 because length of a point is 0! This is a contradiction, because then the net probability of finding the particle in the box will be 0. The way out is to define probability density function.

Definition 17. Let X be a continuous real valued random variable. A probability density function f is a real function such that

$$P(a < X < b) = \int_a^b f(x)dx.$$

Thus, probability of finding particle at x is 0. However, probability that particle is near x, that is the particle is in $[x, x + dx]$ is given by $f(x)dx$.

The function $F(x) = P(X < x)$ is called a cumulative distribution function. Clearly,

$$f(x) = \frac{dF}{dx}(x).$$

Example 18. For particle-in-a-box example, the probability density function is

$$f(x) = \begin{cases} \frac{1}{L} & x \in [0, L] \\ 0 & \text{otherwise.} \end{cases}$$

The expectation value and variance are defined exactly in the same way as in case of discrete variables. Thus

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

and

$$V(x) = E\left(X^2\right) - (E(x))^2.$$

Example 19. For particle-in-a-box example,

$$E(X) = \int_0^L x\left(\frac{1}{L}\right)dx = \frac{L}{2}$$

and

$$V(X) = \int_0^L x^2\left(\frac{1}{L}\right)dx = \frac{L^2}{3}.$$

Another Example.

Example 20. A ball is released from height H. What is pdf for finding particle at height y measured downwards from the point of release?

Consider a small interval dy at y. Probability that particle is found in $[y, y+dy]$ is proportional to the time the ball spends in that section. That is,

$$\begin{aligned} f(y)dy &= P(X \in [y, y+dy]) \\ &= K\,dt \\ &= K\frac{dy}{v(y)} \end{aligned}$$

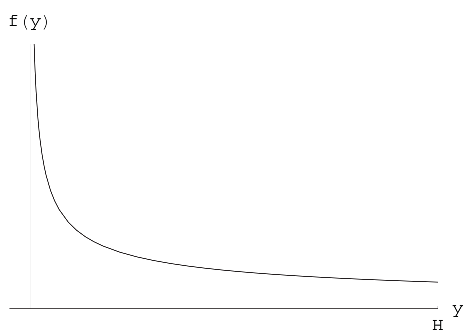
where, K is a constant and $v(y) = \sqrt{2gy}$ is the speed of the ball at y. Thus, $f(y) = K/\sqrt{2gy}$. The constant of proportionality can be found by *normalizing* f. The normalization condition,

$$\int_0^H f(y)dy = 1$$

gives $f(y) = 1/\left(2\sqrt{Hy}\right)$. The cdf $F(x) = \sqrt{y/H}$. The average height is given by

$$E(y) = \int_0^H y\left(\frac{1}{2\sqrt{Hy}}\right)dy = \frac{H}{3}$$

The graph below shows the pdf



Exercise 21. A particle is performing simple harmonic motion with amplitude A and frequency ω. Show that the pdf for finding particle at a position x is given by

$$f(x) = \frac{1}{\pi}\frac{1}{\sqrt{A^2 - x^2}}.$$

Plot this function and find the variance of the distribution.