1 Basic probability theory

1.1 Random variables and distributions

Suppose X is an observable which can be measured in an experiment. Because of lack of information let us assume that we cannot definitely predict the outcome of an experiment measuring X. Instead, let us say that we can only give the probability that the variable X will take any specific value or some value within a range of possible values. Then we say that X is a **random variable**. The outcome of the experiment can be described by a **probability distribution** $P(X)$.

Examples of random variables and probability distributions:

1. A coin tossing experiment – we toss an un-biased coin whose two faces have the values $+1$ (for head) and -1 (for tail) written on them . The outcome X can take two possible values $X = 1$ or $X = -1$. In this case $P(X)$ is given by:

$$P(X = +1) = 1/2$$
$$P(X = -1) = 1/2 \ .$$

2. A dice rolling experiment – we throw a dice with six faces with the values $1, 2, 3, 4, 5, 6$ written respectively on each of them. In this case the allowed values of X are the six numbers $1 - 6$ and $P(X)$ is given by:

$$P(X = i) = 1/6 \quad \text{for } i = 1, 2, \ldots, 6 \ .$$

3. Tossing N number of coins – we toss N coins and measure the number of heads obtained. The outcome X can take values $0, 1, \ldots, N$ and $P(X)$ is given by the **Binomial** distribution

$$P(X = n) = \frac{N!}{n!(N-n)!}\frac{1}{2^N} \ . \tag{1}$$

4. Number of air molecules in a small volume. Let the mean density of air in a room be ρ and let us consider a volume v which is a small fraction of the total volume V of the room. The mean number of molecules in v will then be $\lambda = \rho v = Nv/V$, where N is the total number of molecules in the room. However if we make a measurement we will not get exactly the value λ. The measured number of air molecules will be a random variable and in this case X can take values $n = 0, 1, 2, \ldots, N$. If we assume that the air density is low and therefore the molecules hardly interact then the distribution is again given by the **Binomial distribution**

$$P(X = n) = \frac{N!}{n!(N-n)!}\left(\frac{v}{V}\right)^n\left(1 - \frac{v}{V}\right)^{N-n} \ . \tag{2}$$

For large N and large V with $\rho = N/V$ kept finite, this reduces to the **Poisson distribution**

$$P(X = n) = \frac{\lambda^n}{n!}e^{-\lambda} \ .$$

5. The number of photons hitting a telescope lens in a time interval τ is also given by the **Poisson distribution**, now with $\lambda = r\tau$, where r is the rate of arrival of the photons.

6. Time taken for a radioactive atom to decay. The time $X = t$ can take any value between 0 and ∞. If the rate of decay is r then $P(t)$ is given by the **exponential** distribution

$$P(t)\ dt = \text{Prob}(t < X < t + dt) = re^{-rt}dt\ .$$

Note that $P(t)$ is now a probability density and has the dimension of 1/time.

7. $x-$component of the velocity $\mathbf{v} = (v_x, v_y, v_z)$ of an air molecule. The velocity component v_x can take any values in the range $(-\infty, \infty)$ and $P(v_x)$ is given by the Maxwell distribution

$$P(v_x) = \left(\frac{m}{2\pi k_B T}\right)^{1/2} e^{-mv_x^2/2k_BT}\ .$$

The Maxwell distribution is an example of a **Gaussian** or a **Normal** distribution whose general form, for a variable X is:

$$P(X) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} e^{-(X-\mu)^2/2\sigma^2}\ , \tag{3}$$

where μ is the mean value of the random variable and σ is the root-mean-square deviation around the mean value. Note that here X is a continuous variable and $P(X)dX$ is the probability (dimensionless number) of a measurement yielding a value between X and $X+dX$.

Experimental determination of probability distributions: We have seen in the above examples that a random variable X can either take a discrete set of values or a continuous range of values. Let us see how we can determine the distribution function $P(X)$ for these two cases in an experiment (or simulation).

Case I – Discrete X (only integer values allowed)

- Perform the experiment to measure X a large number (say R) of times.
- Every measurement will yield a value. Count the number of times we get any particular value. Suppose $X = n$ occurs R_n times.
- $$P(X = n) = \lim_{R\to\infty} \frac{R_n}{R}\ .$$
- In an experiment the larger we take R the better our answer will get.

Case II – Continuous X (all real values allowed)

- Perform the experiment to measure X a large number (say R) of times.
- Every measurement will yield a value. Count the number of times we get a value of X in the range x to $x + \Delta x$, where Δx is chosen small. Suppose this occurs R_x number of times.

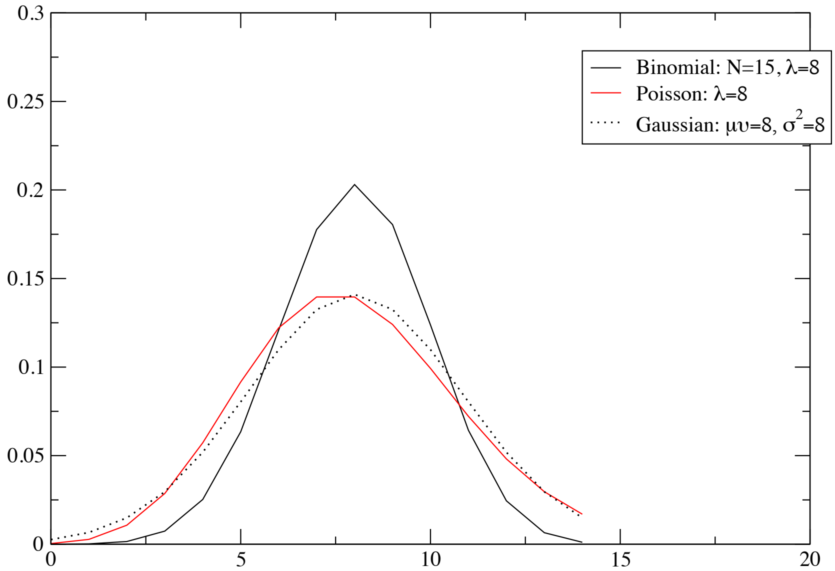

Figure 1: Here we show plots of the Binomial distribution, the Poisson distribution and the Gaussian distribution.

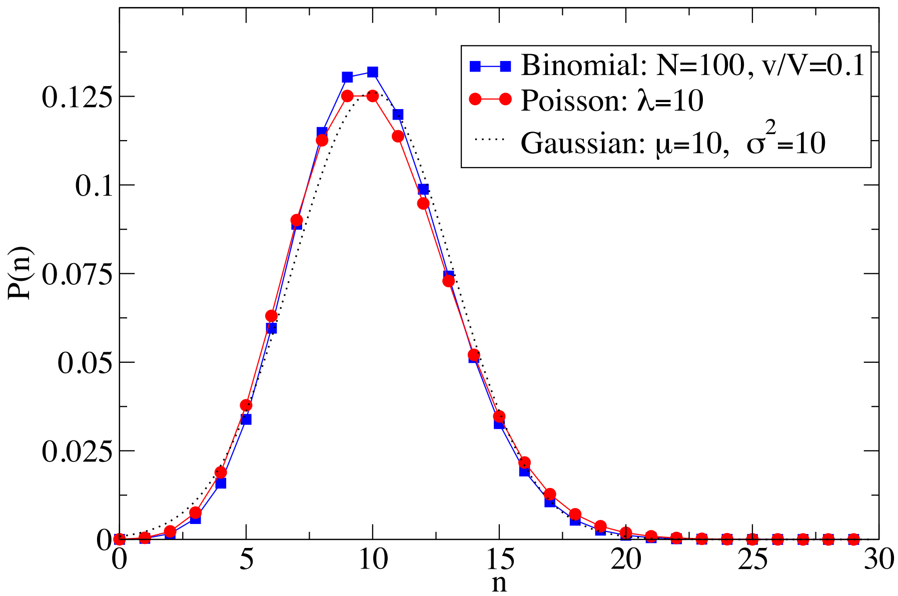

Figure 2: Here we show plots of the Binomial distribution, the Poisson distribution and the Gaussian distribution.

- Then

$$Prob(x < X < x + \Delta x) = \lim_{R\to\infty} \frac{R_x}{R} \ .$$

  and

$$P(x) = \lim_{\Delta x \to 0} \frac{Prob(x < X < x + \Delta x)}{\Delta x} \ .$$

- In an experiment we need to take R as large as possible and Δx as small as possible.

1.2 Properties of probability distributions

Case I – Discrete $X = n$

1. Positivity and normalization.

$$\begin{aligned} P(n) &\geq 0 \\ \sum_{n=-\infty}^{\infty} P(n) &= 1 \ . \end{aligned}$$

2. Two important properties of a random variable are its mean value and the mean square deviation from the mean value. These are respectively denoted by μ and σ^2 and defined as:

$$\begin{aligned} \mu &= \langle n \rangle = \sum_{n=-\infty}^{\infty} n \ P(n) \\ \sigma^2 &= \langle (n-\mu)^2 \rangle = \sum_{=-\infty}^{\infty} (n-\mu)^2 \ P(n) \ . \end{aligned}$$

3. We can define the **moments** of the distribution:

$$M_s = \langle n^s \rangle = \sum_{n=-\infty}^{\infty} n^s \ P(n) \quad \text{for } s = 1, 2, ...$$

Prob: Verify the above properties for the binomial distribution and find μ and σ.
Prob: Show that $\sigma^2 = M_2 - M_1^2$.

Case II – continuous $X = x$

1. Positivity and normalization.

$$\begin{aligned} P(x) &\geq 0 \\ \int_{-\infty}^{\infty} dx \ P(x) &= 1 \ . \end{aligned}$$

2. The mean and mean square deviation are given by

$$\begin{aligned}\mu &= \langle x \rangle = \int_{-\infty}^{\infty} dx\ x\ P(x) \\ \sigma^2 &= \langle (x-\mu)^2 \rangle = \int_{-\infty}^{\infty} dx\ (x-\mu)^2 P(x)\ .\end{aligned}$$

3. The **moments** of the distribution are defined as:

$$M_s = \langle x^s \rangle = \int_{-\infty}^{\infty} dx\ x^s\ P(x) \quad \text{for } s = 1, 2, ... \tag{4}$$

Prob: Verify the above properties for the Maxwell distribution and find μ and σ.

1.3 Multivariate distributions, joint and conditional distributions

In general we can consider more than one random variable. Our observables can be a set of N random variables $(X_1, X_2, \ldots, X_N)$ and we can ask for their **joint probability distribution** $P(X_1, X_2, \ldots, X_N)$.

Examples:

1. Tossing two **independent** coins simultaneously – the first can be in two possible states $X_1 = +1$ or $X_1 = -1$ and the second coin can also be in two possible states $X_2 = +1$ or $X_2 = -1$. The joint distribution $P(X_1, X_2)$ is given by:

$$\begin{aligned}P(+1,+1) &= 1/4 \\ P(-1,+1) &= 1/4 \\ P(+1,-1) &= 1/4 \\ P(-1,-1) &= 1/4\ .\end{aligned} \tag{5}$$

2. All the three components of the velocity $\mathbf{v} = (v_x, v_y, v_z)$ of an air molecule. In this case we have three random variables $X_1 = v_x, X_2 = v_y, X_3 = v_z$ and their joint distribution is given by:

$$P(v_x, v_y, v_z) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-mv^2/2k_BT}\ ,$$

where $v^2 = v_x^2 + v_y^2 + v_z^2$. In this case $P(\mathbf{v})dv_x dv_y dv_z$ gives the probability that the components of the measured velocity lie within the ranges $(v_x, v_x + dv_x), (v_y, v_y + dv_y), (v_z, v_z + dv_z)$ respectively.

Independent random variables: The set of random variables $(X_1, X_2, \ldots, X_N)$ are said to be independent if their joint distribution can be written in the following product form:

$$P(X_1, X_2, \ldots, X_N) = P_1(X_1)P_2(X_2)\ldots P_N(X_N) \ ,$$

where P_i denotes the probability distribution of the i^{th} variable. If the individual distributions are all identical, *i.e* $P_1 = P_2 \ldots = P_N$ then the random variables are said to be *iid* (independent-identically-distributed).

Prob: Verify that in e.g (1) above X_1 and X_2 are *iid* variables since $P(X_1, X_2) = P(X_1)P(X_2)$. In e.g (2), verify that v_x, v_y and v_z are *iid* variables since $P(v_x, v_y, v_z) = P(v_x)P(v_y)P(v_z)$.

For simplicity let us now stick to the case when we have two random variables (X_1, X_2) with a joint distribution $P(X_1, X_2)$. In general the variables can be **correlated** instead of being independent and in that case $P(X_1, X_2) \neq P_1(X_1)P_2(X_2)$. We define the conditional probability $P^c(X_2|X_1)$ of X_2 given X_1 through the relation:

$$P^c(X_2|X_1) = \frac{P(X_1, X_2)}{P_1(X_1)} \ . \tag{6}$$

Similarly we can define

$$P^c(X_1|X_2) = \frac{P(X_1, X_2)}{P_2(X_2)} \ . \tag{7}$$

From the definition of the conditional probability we get **Bayes' theorem**

$$P^c(X_2|X_1)P_1(X_1) = P^c(X_1|X_2)P_2(X_2) \ .$$

Understanding the definition of conditional probability: Consider an experiment with two coins which are connected to each other by a rod. We put the coins in a box, shake it and look at the values X_1 and X_2 take. Let us do ths experiment R number of times. There are four possible outcomes and let their frequencies be:

1. $(X_1 = +1, X_2 = +1)$ occurs R_{++} times ,
2. $(X_1 = -1, X_2 = +1)$ occurs R_{-+} times ,
3. $(X_1 = +1, X_2 = -1)$ occurs R_{+-} times ,
4. $(X_1 = -1, X_2 = -1)$ occurs R_{--} times ,

where $R = R_1 + R_2 + R_3 + R_4$. As we have seen earlier we can determine various probabilities accurately if we make R large. Suppose we want to find the probabilty that $X_1 = +1$ given we know that $X_2 = -1$. Then clearly (for $R \to \infty$):

$$\begin{aligned} P^c(X_1 = +1|X_2 = -1) &= \frac{R_{+-}}{R_{+-} + R_{--}} \\ &= \frac{R_{+-}/R}{(R_{+-} + R_{--})/R} = \frac{P(X_1 = +1, X_2 = -1)}{P(X_2 = -1)} \ , \end{aligned}$$

which agrees with our previous definition.

If the rod connecting the coins is very rigid (with same faces pointing in the same direction), then X_1 and X_2 are strongly correlated and we would then have

$$P(X_1 = +1, X_2 = +1) = P(X_1 = -1, X_2 = -1) = 1/2 \ ,$$
$$P(X_1 = +1, X_2 = -1) = P(X_1 = -1, X_2 = +1) = 0 \ ,$$

Clearly in this case $P(X_1, X_2) \neq P_1(X_1)P_2(X_2)$. On the other hand if the rod is very flexible then $P(X_1, X_2)$ is given by Eq. (5) and X_1, X_2 are un-correlated.

1.4 Characteristic functions

Consider a continuous random variable with a distribution $P(x)$. It's characteristic function is defined by

$$\widetilde{P}(k) = \int_{-\infty}^{\infty} dx e^{ikx} P(x) \ , \tag{8}$$

which is thus just the Fourier transform of $P(x)$. The characteristic function $\widetilde{P}$ has the same information as $P(x)$, which can be obtained from it by an inverse Fourier transform. Thus

$$P(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} \widetilde{P}(k) \ . \tag{9}$$

If we use the expansion $e^{ikx} = \sum_{s=0}^{\infty} (ikx)^s / s!$ in Eq. (8) then we get:

$$\begin{aligned} \widetilde{P}(k) &= \sum_{s=0}^{\infty} \frac{(ik)^s}{s!} \int_{-\infty}^{\infty} dx x^s P(x) \ , \\ &= \sum_{s=0}^{\infty} \frac{(ik)^s}{s!} M_s \ , \end{aligned} \tag{10}$$

where M_s are the moments defined in Eq. (4) . Comparing this series with the Taylor series expansion of $\widetilde{P}(k)$ about $k = 0$:

$$\widetilde{P}(k) = \sum_{s=0}^{\infty} \frac{k^s}{s!} \frac{d^s \widetilde{P}(k)}{dk^s} \Big|_{k=0} \ ,$$

we immediately see that

$$M_s = (-i)^s \frac{d^s \widetilde{P}(k)}{dk^s} \Big|_{k=0} \ .$$

Example: Consider the Gaussian distribution in Eq. (3) for the case $\mu = 0$. The characteristic function is given by:

$$\widetilde{P}(k) = \int_{-\infty}^{\infty} dx e^{ikx} \frac{1}{(2\pi)^{1/2} \sigma} e^{-x^2/2\sigma^2} = e^{-\sigma^2 k^2/2} \ .$$

Using the expansion of $e^{-\sigma^2 k^2/2}$ and comparing with Eq. (10) we get:

$$\widetilde{P}(k) = \sum_{s=0}^{\infty} \frac{(-k)^{2s}\sigma^{2s}}{2^s s!} = \sum_{s=0}^{\infty} \frac{(ik)^s}{s!} M_s \ . \tag{11}$$

Comparing the terms of the two series we then get

$$M_{2s} = \frac{2s!}{2^s s!}\sigma^{2s} \quad \text{for} \ \ s = 0, 1, \dots \ ,$$

and all odd moments vanish.

1.5 Central limit theorem

Suppose we have a set of N continuous random variables $(x_1, x_2, \dots, x_N)$. Each of the random variable are chosen from the distribution $p(x)$ which has a mean μ and a mean square deviation σ^2. Let us take the sum of these variables and denote it by y. Thus

$$y = \sum_{l=1} x_l \ . \tag{12}$$

Clearly y is also a random variable and it will have a distribution function $P(y)$. How do we find this distribution and how does it depend on the distribution $p(x)$? The central limit theorem states that if we take N to be very large then the distribution of y has a simple form given by the Gaussian distribution

$$P(y) \ = \ \frac{1}{(2\pi N\sigma^2)^{1/2}} e^{-(y-N\mu)^2/2N\sigma^2} \ .$$

Proof:

$$P(y) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} dx_1 \dots dx_N \ \delta[y - (x_1 + x_2 + \dots + x_N)] \ P(x_1)P(x_2)\dots P(x_N) \ . \tag{13}$$

We now use the follwoing representation of the Dirac-*delta* function:

$$\delta(y-a) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-ik(y-a)}$$

in Eq. (13) to get

$$\begin{aligned} P(y) \ &= \ \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-iky} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} dx_1 \dots dx_N \ e^{ik(x_1+x_2+\dots+x_N)} \ p(x_1)p(x_2)\dots p(x_N) \\ &= \ \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-iky} [\tilde{p}(k)]^N \ , \\ \text{where} \ \ \tilde{p}(k) \ &= \ \int_{-\infty}^{\infty} dx e^{ikx} p(x) \end{aligned}$$

is the characteristic function of the distribution $p(x)$. Now we write

$$\begin{aligned} [\tilde{p}(k)]^N \ &= \ \exp[N \log \tilde{p}(k)] \\ &= \ \exp[N \log(1 + ik\mu - k^2 M_2/2 + \mathrm{O}(k^3))] \\ &= \ \exp[N(ik\mu - k^2\sigma^2/2 + \mathrm{O}(k^3))] \ . \end{aligned}$$

Plugging this into Eq. (14) we note that for large N, the integral gets its contribution mainly from small values of k. Hence it is alright to neglect $O(k^3)$ terms and we get:

$$\begin{aligned} P(y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-ik(y-N\mu)} \, e^{-k^2 N\sigma^2/2} \\ &= \frac{1}{(2\pi N\sigma^2)^{1/2}} e^{-(y-N\mu)^2/2N\sigma^2} \; . \end{aligned}$$

Some exact results that are valid for any N, not necessarily large.

- $\langle y\rangle = \langle (x_1 + x_2 + x_3 + \cdots + x_N)\rangle$
  $= \langle x_1\rangle + \langle x_2\rangle + \langle x_3\rangle + \cdots + \langle x_N\rangle = N\mu$.
- $\langle (y - N\mu)^2\rangle = \langle [(x_1 - \mu) + (x_2 - \mu) + (x_3 - \mu) + \cdots + (x_N - \mu)]^2\rangle$
  $= \langle (x_1 - \mu)^2\rangle + \langle (x_2 - \mu)^2\rangle + \langle (x_3 - \mu)^2\rangle + \cdots + \langle (x_N - \mu)^2\rangle$
  where we have used the fact that $\langle (x_l - \mu)(x_n - \mu)\rangle = \langle (x_l - \mu)\rangle \, \langle (x_n - \mu)\rangle = 0$ for $l \neq n$.

2 Random walks

History of the random walk problem: Karl Pearson introduced the term "Random Walk". He was interested in describing the spatial/temporal evolutions of mosquito populations invading cleared jungle regions. He found it too complex to model deterministically, so he conceptualized a simple random model. Pearson posed his problem in Nature (27 July 1905) as follows:

A man starts from a point 0 and walks ℓ yards in a straight line; he then turns through any angle whatever and walks another ℓ yards in a second straight line. He repeats this process n times. I require the probability that after n of these stretches he is at a distance between r and $r + dr$ from his starting point.

The question was answered the following week by Lord Rayleigh, who pointed out the connection between this problem and an earlier paper of his published in 1880* concerned with sound vibrations. Rayleigh pointed out that, for large values of n, the answer is given by:

$$P(r, n)dr = \frac{2}{n\ell^2} \; e^{-r^2/n\ell^2} \; r \; dr \; .$$

2.1 Random walk in one-dimension

Consider a walker on a one dimensional lattice. At every time step the walker tosses an unbiased coin and moves to the left if it is a head and to the right if it is a tail. Thus for every step there are 2 possibilities and the walker chooses either of them with equal probability. After taking n steps the walker can be anywhere between $-n$ to n and we would like to know: what is the probability $P(X, n)$ that the walker is at some point X ? We note that there are 2^n distinct possible walks, each of which occurs with the same probability. Out of these let us say $T(X,n)$ walks end up at the point X. Then clearly

$$P(X,n) = \frac{T(X,n)}{2^n}. \tag{14}$$

We can find $T(X,n)$ as follows. For any given realization of the walk let n_R be the number of right steps and let n_L be the number of left steps. Then $n_R + n_L = n$ and $n_R - n_L = X$.

Example: One possible realization of a 10−step walk is $LRRLLRLLRL$. In this case $n = 10,\ n_R = 4,\ n_L = 6$ and $X = -2$. A different realization of the walk which leads to the same values of n, n_R, n_L, X is $LLLLRRRRLL$.

Clearly there are many possible ways of arranging the n_R Rs' and n_L L's and the number of ways would give us $T(X,n)$. This is a combinatorial problem whose answer is:

$$T(X,n) = \frac{n!}{n_R! n_L!}. \tag{15}$$

Prob: Check this formula for $n = 4$.

Now since $n_R = (n+X)/2$ and $n_L = (n-X)/2$ we therefore get, using Eq. (14) and Eq. (15),

$$P(X,n) = \frac{n!}{\frac{n+X}{2}!\frac{n-X}{2}!}\frac{1}{2^n}. \tag{16}$$

Prob: Check normalization: $\sum_{X=-n}^{n} P(X,n) = 1$. Use the fact that $(1/2+1/2)^n = 1$.

Now what we would eventually like to get is a "continuum description", that is we want to look at length scales much larger than the lattice spacing (say a) and time scales much larger than time taken for each step (say τ). Let us try to get this.

For this we use Stirling's approximation. This states that for large n we get $n! \approx n^n e^{-n}\sqrt{2\pi n}$. This formula is infact very good even for n as small as 5; in that case $5! = 120$ while the Stirling approximation gives ≈ 118.

Proof of Stirling's approximation: One derivation of the Stirling formula is to use the following result:

$$n! = \int_0^\infty dx\ e^{-x} x^n\ . \tag{17}$$

Let us try to evaluate this integral by finding out where it takes its maximum value and then approximating the integrand by a Gaussian around the maximum. We can write the integrand in the form $e^{-x}\ x^n = e^{-x+n\log x} = e^{f(x)}$, where $f(x) = -x + n\log x$. The point x^* where $f(x)$ and therefore $e^{f(x)}$ is maximum is given by the condition $f'(x^*) = 0$ which gives $x^* = n$. Expanding around this point we get

$$\begin{aligned} f(x) &= f(x^*) + \frac{1}{2}f''(x^*)(x-x^*)^2 + \dots \\ \text{where } f(x^*) &= -n + n\log n\ , \quad f''(x^*) = -1/n\ . \end{aligned} \tag{18}$$

Using this in Eq. (17) we get

$$\begin{aligned} n! &= \int_0^\infty dx e^{f(x)} \\ &\approx e^{f(x^*)} \int_0^\infty dx \; e^{f''(x^*)(x-x^*)/2} \\ &\approx e^{f(x^*)} \int_{-\infty}^\infty dx \; e^{f''(x^*)(x-x^*)/2} \\ &= e^{f(x^*)} \left(\frac{2\pi}{-f''(x^*)} \right)^{1/2} = e^{-n+n\log n}(2\pi n)^{1/2} \; , \end{aligned} \tag{19}$$

Which is the Stirling formula.

QED

Using Stirling's formula in Eq. (16) we get:

$$P(X,n) = \frac{n^n e^{-n}(2\pi n)^{1/2}}{(\frac{n+X}{2})^{\frac{n+X}{2}} e^{-\frac{n+X}{2}} [2\pi(\frac{n+X}{2})]^{1/2} (\frac{n-X}{2})^{\frac{n-X}{2}} e^{-\frac{n-X}{2}} [2\pi(\frac{n-X}{2})]^{1/2}} \frac{1}{2^n} \; .$$

After simplification this reduces to

$$P(X,n) = \frac{(2\pi n)^{1/2}}{(1+\frac{X}{n})^{\frac{n+X}{2}}(1-\frac{X}{n})^{\frac{n-X}{2}} \pi n (1-\frac{X^2}{n^2})^{1/2}} \; .$$

We now consider $X << n$ only or more precisely $X \lesssim O(\sqrt{n})$. In this limit, we get

$$P(X,n) = (\frac{2}{\pi n})^{1/2} e^{-\frac{X^2}{2n}} \; . \tag{20}$$

This is easiest to obtain by taking $\ln P(X,n)$ and expanding. Now let $x = Xa$ and $t = n\tau$. Then the *probability density* for the walker to be between x and $x + dx$ is

$$\begin{aligned} p(x,t) &= P(X,n)/(2a) \\ &= \frac{1}{(4\pi(a^2/2\tau)t)^{1/2}} e^{-\frac{x^2}{4(a^2/2\tau)t}} . \end{aligned} \tag{21}$$

The reason we divide by $2a$ and not a is because after n steps the walker can be located either on even sites (if n is even) or on odd sites (n odd). Now defining the diffusion constant $D = a^2/(2\tau)$ we finally get

$$p(x,t) = \frac{1}{(4\pi Dt)^{1/2}} e^{-\frac{x^2}{4Dt}} . \tag{22}$$

Prob: Check that $\int_{-\infty}^{\infty} dx \; p(x,t) = 1$. Also verify that $\langle x \rangle = 0$ and $\langle x^2 \rangle = 2Dt$.

The moments $\langle x \rangle$ and $\langle x^2 \rangle$ can be obtained more directly. The position of the walker $x(t)$ after $n = t/\tau$ time steps is

$$x(t) = a \sum_{i=1}^{n} \xi_i \tag{23}$$

where ξ_i is $+1$ or -1 with equal probability (thus $\langle \xi_i \rangle = 0$) and ξ_i and ξ_j are uncorrelated or independent, which means that $\langle \xi_i \xi_j \rangle = 0$. Therefore

$$\begin{aligned}
\langle x(t) \rangle &= a \sum_{i=1}^{n} \langle \xi_i \rangle = 0 \\
\langle x^2(t) \rangle &= a^2 \sum_{i,j=1}^{n} \langle \xi_i \xi_j \rangle \\
&= a^2 (\sum_{i=1}^{n} \langle \xi_i^2 \rangle + \sum_{i \neq j} \langle \xi_i \xi_j \rangle) \\
&= a^2 n = 2[a^2/(2\tau)](n\tau) = 2Dt
\end{aligned} \tag{24}$$

Prob: Write a Monte-carlo program to generate $1-D$ random walks and verify the law $\langle x^2 \rangle = 2Dt$.

Prob: Let $\xi_i = 1$ with probability p and -1 with probability $q = 1 - p$. Find $\langle x(t) \rangle$ and $\langle x^2(t) \rangle - \langle x(t) \rangle^2$.

Prob: Let $x_i = 2$ with probability $1/2$ and $x_i = -1$ or 0 with probabilities $1/4$ each. Find $\langle x(t) \rangle$ and $\langle x^2(t) \rangle - \langle x(t) \rangle^2$.

2.2 Random walk and the Diffusion equation

Another method to get Eq. (22) : Since a random walk is like diffusion of particles we expect Eq. (22) to be the solution of the diffusion equation. Let us see how this comes about. As before, $P(X, n)$ is the probability that a particle is at the site X after n steps. It satisfies the following equation

$$P(X, n+1) = \frac{1}{2} \ [\ P(X+1, n) + P(X-1, n)\ ] \tag{25}$$

Subtract $P(X, n)$ from both sides. We then get

$$\begin{aligned}
&P(X, n+1) - P(X, n) = \frac{1}{2} \ [\ P(X+1, n) - 2P(X, n) + P(X-1, n)\ ] \\
&\Rightarrow \frac{p(x, t+\tau) - p(x, t)}{\tau} = \frac{a^2}{2\tau} \frac{[p(x+a, t) - 2p(x, t) + p(x-a, t)]}{a^2} \\
&\Rightarrow \frac{\partial p(x, t)}{\partial t} = D \frac{\partial^2 p(x, t)}{\partial x^2}
\end{aligned} \tag{26}$$

which is the *diffusion equation*. Normally in the diffusion equation we have density of particles $\rho(x, t)$ instead of the probability density $p(x, t)$. But they are simply related by $\rho(x, t) = N\ p(x, t)$ where N is the total number of diffusing particles.

Now Eq. (26) is a linear equation which can be easily solved by Fourier transforming. Solving means: given an initial probability distribution $p(x, t = 0)$ find $p(x, t)$ at some later time t. Let us solve for the initial condition $p(x, t = 0) = \delta(x)$, which corresponds to the case when the particle is initially located at the origin. Taking the Fourier transform

$$p(x, t) = \int_{-\infty}^{\infty} \widetilde{p}(k, t) e^{ikx} dk; \quad \widetilde{p}(k, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} p(x, t) e^{-ikx} dx$$

gives

$$\int_{-\infty}^{\infty} \frac{\partial \widetilde{p}(k,t)}{\partial t} e^{ikx} dk = \int_{-\infty}^{\infty} -Dk^2 \widetilde{p}(k,t) e^{ikx} dk$$
$$\Rightarrow \frac{\partial \widetilde{p}(k,t)}{\partial t} = -Dk^2 \widetilde{p}(k,t)$$
$$\Rightarrow \widetilde{p}(k,t) = e^{-Dk^2 t} \widetilde{p}(k,0) = \frac{1}{2\pi} e^{-Dk^2 t} \qquad (27)$$

Taking the inverse Fourier transformation we then get

$$p(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-Dk^2 t} e^{ikx} dk = \frac{1}{(4\pi D t)^{1/2}} e^{-\frac{x^2}{4Dt}} \ . \qquad (28)$$

as before. Note that the diffusion equation can also be written in the following form:

$$\frac{\partial p(x,t)}{\partial t} + \frac{\partial J(x,t)}{\partial x} = 0 \quad \text{where} \qquad (29)$$
$$J(x,t) = -D \frac{\partial p(x,t)}{\partial x}.$$

Higher dimensions: We can consider a random walk on a $2-$dimensional, $3-$dimensional or in general a $d-$dimensional lattice and ask the same questions. The combinatorial approach becomes difficult but we get the same diffusion equation which can again be solved quite easily.

To see that we do get the same equation, consider the $2-$dimensional case, where a random walker can move up, down, left or right with equal probabilities. Thus if at some time $t = n\tau$ the walker is at the point $\mathbf{x} = (x,y)$ then, at time $t+\tau$, it can be at either of the 4 points $(x+a,y)$, $(x-a,y)$, $(x,y+a)$, $(x,y-a)$. The probability of it being at any of these 4 points is clearly $1/4$. Let $P(\mathbf{x},t)$ be the probability for the walker to be at $\mathbf{x}$ at time t. Then Eq. (25) gets modified to

$$P(\mathbf{x}, t+\tau) = \frac{1}{4} \ [\ P(x+a,y,t) + P(x-a,y,t) + P(x,y+a,t) + P(x,y-a,t)\ ] \ . \qquad (30)$$

Subtracting $P(\mathbf{x},t)$ from both sides we get:

$$\begin{aligned} \frac{P(\mathbf{x},t+\tau) - P(\mathbf{x},t)}{\tau} &= \frac{a^2}{4\tau} \frac{[P(x+a,y,t) - 2P(x,y,t) + P(x-a,y,t)]}{a^2} \\ &+ \frac{a^2}{4\tau} \frac{[P(x,y+a,t) - 2P(x,y,t) + P(x,y-a,t)]}{a^2} \\ \Rightarrow \frac{\partial p(\mathbf{x},t)}{\partial t} &= D \left[\frac{\partial^2 p(\mathbf{x},t)}{\partial x^2} + \frac{\partial^2 p(\mathbf{x},t)}{\partial y^2} \right] \ , \end{aligned} \qquad (31)$$

which is the $2-$dimensional diffusion equation and we have defined $D = a^2/(4\tau)$. Similarly in $3-$dimensions we get [with $D = a^2/(6\tau)$]

$$\frac{\partial p(\mathbf{x},t)}{\partial t} = D \nabla^2 P(\mathbf{x},t). \qquad (32)$$

To solve this we again Fourier transform

$$p(\mathbf{x},t) = \int_{-\infty}^{\infty} \tilde{p}(\bar{k},t) e^{i\bar{k}.\mathbf{x}} d\bar{k}; \quad \tilde{p}(\bar{k},t) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} p(\mathbf{x},t) e^{-i\bar{k}.\mathbf{x}} d\mathbf{x}.$$

Proceeding exactly as in the $1-D$ case we get

$$p(\mathbf{x},t) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\bar{k} e^{-Dk^2t} e^{i\bar{k}.\mathbf{x}} = p(x,t)p(y,t)p(z,t) = \frac{1}{(4\pi Dt)^{3/2}} e^{-\frac{r^2}{4Dt}} \tag{33}$$

The mean square distance traveled by the walker is $\langle r^2 \rangle = \langle x^2 + y^2 + z^2 \rangle = 6Dt$. We can verify this directly. Since $\mathbf{x}(t) = a\sum_{i=1}^{n} \bar{\xi}_i$, therefore $\langle \mathbf{x}^2 \rangle = a^2 \sum_{i=1}^{n} \langle \bar{\xi}_i^2 \rangle = na^2 = 6Dt$. Note that the number of walks of length n, from the origin to $\mathbf{x}$ is $6^n P(\mathbf{x},t)$.

3 Langevin equations and Brownian motion

The random walk is basically described by the equation

$$X_{n+1} = X_n + \xi_n$$

where ξ_n is uncorrelated noise with zero mean *i*.e

$$\langle \xi_n \rangle = 0; \quad \langle \xi_n \xi_m \rangle = \delta_{nm}.$$

For continuous space and time we write, as usual, $t = n\tau$, $x = X_n a$ and the above equation gives

$$\begin{aligned} \frac{dx(t)}{dt} &= \xi(t) \quad \text{with} \\ \langle \xi(t) \rangle = 0; &\quad \langle \xi(t)\xi(t') \rangle = 2D\delta(t-t'). \end{aligned} \tag{34}$$

This is the *Langevin equation* describing diffusion. It is the equation of motion for a free Brownian particle (a particle whose velocity is a random function of time). If we have a large number of such particles then we have seen that their density obeys the diffusion equation which can be written in the form

$$\begin{aligned} \frac{\partial \rho(x,t)}{\partial t} + \frac{\partial J_{diff}(x,t)}{\partial x} &= 0 \quad \text{where} \\ J_{diff}(x,t) &= -D\frac{\partial \rho(x,t)}{\partial x}. \end{aligned} \tag{35}$$

Now consider a different situation where a large number of particles are moving with a deterministic (that is non-random) velocity. The equation of motion of each particle is then

$$\frac{dx(t)}{dt} = f(x). \tag{36}$$

In this case also we can define the density distribution of particles and ask the question: how does the density change with time ? This is basically determined from the conservation of particles which is given by the following continuity equation:

$$\begin{aligned} \frac{\partial \rho(x,t)}{\partial t} + \frac{\partial J_{drift}(x,t)}{\partial x} &= 0 \quad \text{where} \\ J_{drift}(x,t) = \dot{x}\rho(x,t) &= f(x,t)\rho(x,t). \end{aligned} \tag{37}$$

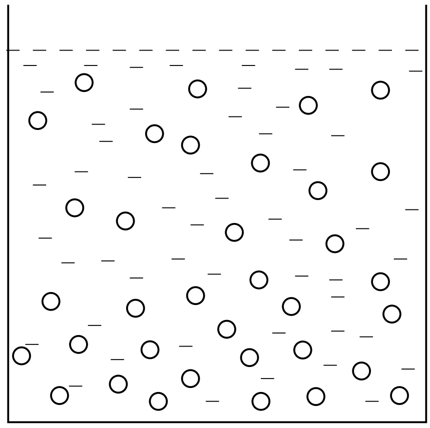

Figure 3: Large number of colloidal particles inside a fluid.

Prob: Derive the continuity equation from the condition for conservation of particles which is $\rho(x,t)dx = \rho(x',t')dx'$ where $t' = t + dt$ and $x' = x + dx$. This is infact just the Liouville's equation in classical mechanics.

The Einstein fluctuation-dissipation relation: we will now use the previous results to derive a formula which will enable us to make estimates of the diffusion constant D in real physical systems.

Consider a large number of colloidal particles inside a fluid as shown in Fig. (3). Each colloidal particle is much larger than the fluid particles which constantly bombard it. The net effect of all the forces imparted on the colloid by the fluid particles can effectively be described very accurately by just two forces:

(i) A *dissipative* part: This is the viscous drag on the particle and is a deterministic force given by

$$F_d = -\gamma \dot{x} \tag{38}$$

where $\gamma = -6\pi\eta a$ is the Stokes formula with a the particle radius and η the fluid viscosity.

(ii) A *fluctuating* part: this is a random force with zero average and which is totally uncorrelated in time, that is we take

$$F_f = \alpha(t); \quad \langle \alpha(t) \rangle = 0 \quad \langle \alpha(t)\alpha(t') \rangle = 2D'\delta(t-t'). \tag{39}$$

What the fluctuation-dissipation theorem (or Einstein relation) tells us is that the two parts of the forces mentioned above are related to each other. To see how this comes about, consider the state when the colloidal particles have reached a steady state and are in thermal equilibrium at some temperature T. We know that because of gravity there will be a concentration gradient of colloidal particles in the system and their density will vary as:

$$\rho(x) = \rho(x=0)e^{-\frac{mgx}{k_B T}}, \tag{40}$$

where m is mass of each particle. Also there will be two currents set up in the system:

(i) a drift current: each colloidal particle is acted upon by two deterministic forces, one is the drag force given by $\gamma \dot{x}$ and the other is gravity given by mg. In the steady state they are equal, which means that the colloid attains a steady speed $\dot{x} = mg/\gamma$. This, from Eq. (37), implies a drift current given by $J_{drift} = mg\rho(x)/\gamma$,

(ii) a diffusive current arising from the random forces: $J_{diff} = -D\frac{\partial \rho(x)}{\partial x}$.

In the equilibrium state there is no net current of particles which means $J_{diff} = J_{drift}$. Hence

$$\frac{\partial \rho(x)}{\partial x} = -\frac{mg}{\gamma D}\rho(x).$$

But from Eq. (40) we get

$$\frac{\partial \rho(x)}{\partial x} = -\frac{mg}{k_B T}\rho(x).$$

Comparing the two equations above we get the Einstein relation

$$D = \frac{k_B T}{\gamma}. \tag{41}$$

Note that it relates the fluctuating (D) and dissipative (η) parts of the fluid forces acting on the colloid. Later we will also determine the strength of the fluctuating force given by D'. Since $\gamma = 6\pi\eta a$, everything on the right hand side of Eq. (41) is known, and we can then use it to estimate the value of D.

Prob: Estimate the value of D for a colloid of size $a = 1\mu m$ in water and in air. How far does it travel in a minute? and in an hour?

3.1 Fokker-Planck equation for a Brownian particle in a potential

The full equation of motion for a Brownian particle moving inside a potential $U(x)$ is given by:

$$\begin{aligned}
&m\frac{dv}{dt} = -\gamma v + F(x) + \alpha(t) \quad \text{where} \\
&F(x) = -\frac{dU(x)}{dx} \\
&\langle \alpha(t) \rangle = 0; \quad \langle \alpha(t)\alpha(t') \rangle = 2D'\delta(t-t') \ .
\end{aligned}$$

If we are in the high viscosity limit (overdamped, low Reynolds number limit) it can be shown that it is alright (for times $t >> m/\gamma$) to neglect inertial terms, namely the term on the left hand side of the above equation. In that case we get

$$\begin{aligned}
&\gamma\frac{dx}{dt} = F(x) + \alpha(t) \\
&\Rightarrow \frac{dx}{dt} = \frac{F(x)}{\gamma} + \xi(t) \quad \text{where} \\
&\xi(t) = \frac{\alpha(t)}{\gamma} \ \Rightarrow \ \langle \xi(t)\xi(t') \rangle = 2D\delta(t-t'); \ D = D'/\gamma^2 \ .
\end{aligned} \tag{42}$$

This is the Langevin equation of a Brownian particle in the overdamped limit. Using the Einstein relation $D = k_B T/\gamma$ we see that $D' = \gamma k_B T$.

If $p(x,t)$ is the probabilty distribution of the particle then, as we saw in the previous section, the probability current corresponding to this Langevin equation is:

$$\begin{aligned} J &= J_{diff} + J_{drift} \\ &= \frac{F(x)}{\gamma} p(x,t) - D \frac{\partial p(x,t)}{\partial x}. \end{aligned}$$

Using the continuity equation $\partial p/\partial t + \partial J/\partial x = 0$, this then leads to the following *Fokker-Planck* equation for $p(x,t)$:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[\frac{F(x)}{\gamma} p(x,t)\right] + D \frac{\partial^2 p(x,t)}{\partial x^2} \ . \tag{43}$$

A systematic derivation of the Fokker-Planck equation: We now give a derivation of the Fokker-Planck equation using the fact that the Langevin equation with δ-correlated noise describes a Markov process and hence we can write a evolution equation for the probability density — this will be a generalization of Eq. (25) for the random walk. Let us assume that starting from point x' at time t', the particle can make a transition of size Δx that is chosen from a distribution $\phi(\Delta x; x', t')$. Then we have:

$$p(x,t) = \int_{-\infty}^{\infty} d\Delta x \ p(x-\Delta x, t-\Delta t) \ \phi(\Delta x; x-\Delta x, t-\Delta t) \ .$$

We now do a Taylor-expansion of the function $f(x-\Delta x) = p(x-\Delta x, t-\Delta t) \ \phi(\Delta x; x-\Delta, t-\Delta t)$ around the point x. We then get:

$$\begin{aligned} p(x,t) &= \sum_{n=0}^{\infty} \frac{\partial^n}{\partial x^n}\left[p(x,t-\Delta t) \int_{-\infty}^{\infty} d\Delta x \ \frac{(-\Delta x)^n}{n!} \ \phi(\Delta x; x, t-\Delta t) \right] \\ &= p(x,t-\Delta t) + \sum_{n=1}^{\infty} \frac{\partial^n}{\partial x^n} \left[p(x,t)(-1)^n \frac{\langle (\Delta x)^n \rangle}{n!}\right] \ , \end{aligned} \tag{44}$$

where we have used the fact that $\int \ d\Delta x \ \phi(\Delta x; x, t-\Delta t) = 1$ and defined the moments of the jump distribution $\langle (\Delta x)^n \rangle = \int_{-\infty}^{\infty} \ d\Delta x \ (\Delta x)^n \ \phi(\Delta x; x, t-\Delta t)$. As we will see, for our process described by the Langevin equations, the coefficients $D_n(x,t) = \lim_{\Delta t \to 0} \langle (\Delta x)^n \rangle / (n! \Delta t)$ vanish for $n > 2$. Hence we get

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}[D_1 p(x,t)] + \frac{\partial^2}{\partial x^2}[D_2 p(x,t)] \ . \tag{45}$$

For the overdamped Langevin equation described by Eq. (42) we have:

$$\Delta x = x(t+\Delta t) - x(t) = \frac{F(x)}{\gamma}\Delta t + \int_t^{t+\Delta t} dt' \xi(t') = \frac{F(x)}{\gamma}\Delta t + \nu(t) \ ,$$

where $\nu(t)$ is a Gaussian distributed number with zero mean and variance $2D\Delta t$. Hence we get:

$$\begin{aligned} D_1 &= \lim_{\Delta t \to 0} \frac{\langle \Delta x \rangle}{\Delta t} = \frac{F(x)}{\gamma} \ , \\ D_2 &= \lim_{\Delta t \to 0} \frac{\langle (\Delta x)^2 \rangle}{2\Delta t} = D \ . \end{aligned}$$

Hence we get the Fokker-Planck equation in Eq. (43).