The characteristic function and sums of random variables

Gabriel T. Landi

References:

Ross, chapter 7

Tomé, de Oliveira, chapters 1 and 2.

Van Kampen (Stochastic processes in physics and chemistry), chapter 1.

Definition

Consider a r.v. X with PDF $p(x)$ (continuous or discrete). The characteristic function (CF) of X is defined as

$$\boxed{G(s) = \langle e^{isX} \rangle = \int_{-\infty}^{\infty} e^{isx}\, p(x)\, dx} \qquad (1)$$

That is, $G(s)$ is the Fourier transform of $p(x)$. If we need to be careful we may also write $G_X(s)$. The variable s is an auxiliary variable. It has no important interpretation.

You will be amazed by how useful the CF is. Here are some of its uses

1) If you know $G(s)$ it is easy to find all moments $\langle x^m \rangle$

2) The CF contains the same amount of information as the PDF. But usually it is much easier to work with.

3) The CF completely characterizes the distribution (the Fourier transform is unique). If you ever need to get the PDF back you just need to compute the inverse Fourier transform

$$\boxed{p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx}\, G(s)\, ds} \qquad (2)$$

4) The CF makes it very easy to work with sums of independent r.v.s.

Finding the moments $\langle x^m \rangle$ from the CF

The CF is the average of e^{isx}. Expanding the exponential in a Taylor series we get

$$e^{isx} = 1 + isx + \frac{(is)^2}{2!} x^2 + \frac{(is)^3}{3!} x^3 + \cdots \qquad (3)$$

Now we take the average and use the fact that $\langle\,\rangle$ is a linear operation:

$$\boxed{G(s) = \langle e^{isx} \rangle = 1 + is\langle x \rangle + \frac{(is)^2}{2!}\langle x^2 \rangle + \cdots} \qquad (4)$$

We see from these results that if we know $G(s)$, we may expand it in a Taylor series and the moments will appear as the coefficients of the expansion.

More precisely, the Taylor series expansion of $G(s)$ is

$$G(s) = \sum_{m=0}^{\infty} \frac{s^m}{m!} G^{(m)}(0) \qquad (5)$$

where $G^{(m)} = \frac{d^m G}{ds^m}$, is the m-th derivative of G with respect to s, evaluated at zero. We may write (4) as

$$G(s) = \sum_{m=0}^{\infty} \frac{(is)^m}{m!} \langle x^m \rangle \qquad (6)$$

Thus we conclude that

$$\langle x^m \rangle = \frac{G^{(m)}(0)}{i^m} \qquad (7)$$

Why this is useful: in principle to compute each $\langle x^m \rangle$ you need to do an integral. Using this method you compute just a single integral [Eq (1)] and then find the $\langle x^m \rangle$ by differentiation.

Example: $X \sim \text{Bern}(p)$

The PDF of Bern(p) was

$$\mathbb{P}(X=1) = p \qquad \mathbb{P}(X=0) = q := 1-p \tag{8}$$

Thus

$$G(s) = \langle e^{isX} \rangle = e^{is(0)}\,\mathbb{P}(X=0) + e^{is(1)}\,\mathbb{P}(X=1)$$

$$\therefore \quad \boxed{G(s) = q + pe^{is}} \qquad \text{Bern}(p) \tag{9}$$

To extract the moments we expand $G(s)$ in a power series

$$G(s) = q + p + isp + \frac{(is)^2}{2!}p + \frac{(is)^3}{3!}p + \cdots$$

Comparing with (4) we see that

$$\langle X^m \rangle = p, \quad m = 1, 2, 3, \ldots \tag{10}$$

This, of course, is a consequence of the very simple structure of Bern(p).

<u>Example</u> : $z \sim N(0,1)$

The CF will be

$$G_z(s) = \int_{-\infty}^{\infty} e^{isz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{isz - z^2/2} dz$$

To compute this integral we complete squares

$$isz - \frac{z^2}{2} = -\frac{1}{2}\left(z^2 - 2isz\right) = -\frac{1}{2}\left[(z-is)^2 + s^2\right]$$

$$= -\frac{1}{2}(z-is)^2 - \frac{s^2}{2}$$

Thus we get

$$G_z(s) = \frac{e^{-s^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-is)^2/2} dz \qquad y = z - is$$

$$= e^{-s^2/2} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy}_{1} \quad \text{by the normalization of the normal}$$

Whence

$$\boxed{G_z(s) = e^{-s^2/2}} \qquad z \sim N(0,1) \qquad (11)$$

Now let us find the moments $\langle z^m \rangle$. To do that we use the series expansion of e^x, with $x = -s^2/2$:

$$e^{-s^2/2} = 1 + (-s^2/2) + \frac{1}{2!}(-s^2/2)^2 + \frac{1}{3!}(-s^2/2)^3 + \dots$$

$$= 1 - \frac{s^2}{2} + \frac{s^4}{2!\,2^2} - \frac{s^6}{3!\,2^3} + \frac{s^8}{4!\,2^4} - \dots$$

We compare this with Eq (4)

$$G(s) = 1 + is\langle z \rangle + \frac{(is)^2}{2!}\langle z^2 \rangle + \frac{(is)^3}{3!}\langle z^3 \rangle + \frac{(is)^4}{4!}\langle z^4 \rangle + \dots$$

First, we readily see that there are no terms which are odd in s and, consequently,

$$\langle z \rangle = \langle z^3 \rangle = \langle z^5 \rangle = \dots \qquad (12)$$

As for the even moments, we see that

$$\langle z^2 \rangle = 1$$

$$\langle z^4 \rangle = \frac{4!}{2!\,2^2} = 3$$

$$\langle z^6 \rangle = \frac{6!}{3!\,2^3} = 15$$

and so on. You may convince yourself that, in general

$$\boxed{\langle z^{2m} \rangle = \frac{(2m)!}{2^m\, m!}} \qquad (13)$$

This result, as crazy as it may sound, is also the number of ways of grouping $2m$ people into m pairs.

Example : $X \sim N(\mu, \sigma^2)$

Recall that the general normal was defined from $Z \sim N(0,1)$ as

$$X = \mu + \sigma Z$$

We then have

$$G_X(s) = \langle e^{isX} \rangle = \langle e^{is(\mu + \sigma Z)} \rangle = e^{is\mu} \langle e^{is\sigma Z} \rangle$$

The last term is the CF of Z, but with parameter $s\sigma$ instead of s. That is

$$\langle e^{is\sigma Z} \rangle = G_Z(s\sigma) = e^{-(s\sigma)^2/2}$$

Thus we conclude that

$$\boxed{G_X(s) = \exp\left\{ is\mu - \frac{s^2\sigma^2}{2} \right\}} \qquad X \sim N(\mu, \sigma^2) \qquad (14)$$

If now you expand $G_X(s)$ in a Taylor series the result will be quite messy and will mix a bunch of σ's and μ's.

Sums of independent random variables

Let X_i, $i=1,\dots,m$ be independent (not necessarily identically distributed) random variables, and let

$$Y = X_1 + \dots + X_m \qquad (15)$$

We want to know the properties of Y. Notice how this is precisely the scenario of the random walk problem. The mean and variance of Y are easy to compute because they are additive when the X_i are independent:

$$\langle Y \rangle = \langle X_1 \rangle + \dots + \langle X_m \rangle \qquad (16)$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_m)$$

But finding $P_Y(y)$ turns out to be quite a difficult task. The CF, on the other hand, is trivial

$$G_Y(s) = \langle e^{isY} \rangle = \langle e^{is(X_1 + \dots + X_m)} \rangle$$

$$= \langle e^{isX_1} \dots e^{isX_m} \rangle$$

← Statistical independence

$$= \langle e^{isX_1} \rangle \dots \langle e^{isX_m} \rangle$$

Thus, we see that

$$\boxed{G_Y(s) = G_{X_1}(s) \dots G_{X_m}(s)} \qquad (17)$$

The CF of the sum is just the product of the CFs

In the event that the X_i are also identically distributed, then all $G_{X_i}(s)$ will be equal and we get

$$\boxed{G_Y(s) = [G_X(s)]^M} \qquad (18)$$

where $G_X(s) = \langle e^{isX} \rangle$.

<u>Example</u>: $X_i \sim \mathrm{Bern}(p)$

We have already seen that

$$Y = X_1 + \dots + X_M \sim \mathrm{Bin}(M, p) \qquad (19)$$

the CF of Y will be, using Eq (9)

$$G_Y(s) = [G_X(s)]^M = (q + p e^{is})^M \qquad (20)$$

we may also find this from the definition, but the calculation is a bit more difficult

$$G_Y(s) = \sum_{k=0}^{M} e^{isk} \binom{M}{k} p^k q^{M-k} = \sum_{k=0}^{M} \binom{M}{k} (e^{is} p)^k q^{M-k}$$

$$= (q + p e^{is})^M$$

where, in the last line, I used the Binomial theorem.

<u>The normal sum theorem</u>: the sum of normals continues to be normal

Let $X_i \sim N(\mu_i, \sigma_i^2)$ be a bunch of normals, each with its own mean μ_i and variance σ_i^2. Also, let

$$Y = X_1 + \dots + X_m$$

The CF of Y will be, using Eq (14)

$$\begin{aligned} G_Y(s) &= G_{X_1}(s) \dots G_{X_m}(s) \\ &= e^{is\mu_1 - \frac{s^2\sigma_1^2}{2}} \, e^{is\mu_2 - \frac{s^2\sigma_2^2}{2}} \dots e^{is\mu_m - \frac{s^2\sigma_m^2}{2}} \\ &= \exp\left\{ is(\mu_1 + \dots + \mu_m) - \frac{s^2}{2}(\sigma_1^2 + \dots + \sigma_m^2) \right\} \end{aligned} \tag{21}$$

Thus

$$G_Y(s) = \exp\left(is\mu - \frac{s^2\sigma^2}{2}\right) \tag{22}$$

with parameters

$$\begin{aligned} \mu &= \mu_1 + \dots + \mu_m \\ \sigma^2 &= \sigma_1^2 + \dots + \sigma_m^2 \end{aligned} \tag{23}$$

Since the CF completely characterizes the distribution, it follows that $Y \sim N(\mu, \sigma^2)$.

The journey back

Now let us practice with some examples of how to convert from the CF back to the PDF. For continuous PDFs this task may be somewhat difficult and usually requires the computation of integrals using the residues theorem. For discrete r.v.s, on the other hand, the task is quite simple.

Ex : the δ-distribution : $p(x) = \delta(x - x_0)$.

the CF is

$$G_X(s) = \int_{-\infty}^{\infty} dx\, e^{isx}\, \delta(x - x_0) = e^{isx_0} \tag{24}$$

the inverse Fourier transform, Eq (2), then gives

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} ds\, e^{-isx}\, G_X(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} ds\, e^{-is(x-x_0)}$$

This is precisely the integral representation of the δ function. so $p(x) = \delta(x - x_0)$.

<u>Example</u>: $X \sim Bin(m, p)$

For $X \sim Bin(m,p)$ we found in (20) that

$$G(s) = (q + pe^{is})^m$$

Now we look for the PDF using Eq (2):

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} ds\, e^{-isx}\, G(s)$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} ds\, e^{-isx} (q + pe^{is})^m$$

To compute this integral we expand the term $(q + pe^{is})^m$ using the Binomial theorem:

$$e^{-isx}(q + pe^{is})^m = e^{-isx}\sum_{k=0}^{m} \binom{m}{k} (pe^{is})^k q^{m-k}$$

$$= \sum_{k=0}^{m} \binom{m}{k} p^k q^{m-k} e^{is(k-x)}$$

We then get

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} ds \sum_k \binom{m}{k} p^k q^{m-k} e^{is(k-x)}$$

$$= \sum_k \binom{m}{k} p^k q^{m-k} \underbrace{\frac{1}{2\pi}\int_{-\infty}^{\infty} ds\, e^{is(k-x)}}_{\delta(k-x)}$$

Thus

$$\boxed{p(x) = \sum_{k=0}^{m} \binom{m}{k} p^k q^{m-k} \delta(k-x)} \qquad (25)$$

We see that, for discrete r.v.s, the inverse Fourier transform gives the PDF of the r.v., written in terms of δ-functions.

Proof of the transformation formula

In the previous set of notes I introduced to you the following transformation of variables formula: if $Y = f(X)$ then

$$P_Y(y) = \int_{-\infty}^{\infty} \delta(y - f(x))\, P_X(x)\, dx \tag{26}$$

Now let's prove it. We start with the CF of Y:

$$G_Y(s) = \langle e^{isY} \rangle = \langle e^{isf(X)} \rangle = \int_{-\infty}^{\infty} e^{isf(x)}\, P_X(x)\, dx$$

The PDF of Y is the inverse Fourier transform of $G_Y(s)$:

$$\begin{aligned}
P_Y(y) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} ds\, e^{-isy}\, G_Y(s) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} ds\, e^{-isy} \int_{-\infty}^{\infty} dx\, e^{isf(x)}\, P_X(x) \\
&= \int_{-\infty}^{\infty} dx\, P_X(x) \underbrace{\int_{-\infty}^{\infty} \frac{ds}{2\pi}\, e^{-is(y - f(x))}}_{\delta(y - f(x))} \\
&= \int_{-\infty}^{\infty} dx\, P_X(x)\, \delta(y - f(x))
\end{aligned}$$

which is (26). qed. I love the CF.

Cumulants and the cumulant generating function

From the CF we define the cumulant generating function (CGF) as

$$\boxed{C(s) = \ln G(s) = \ln \langle e^{isx} \rangle} \tag{27}$$

Just like we expanded G(s) in a power series, we may also expand C(s). In fact, the log of a series is also a series. And it turns out to be a quite special one. You may check using Taylor expansions that

$$\ln\left[1 + is\langle x\rangle + \frac{(is)^2}{2!}\langle x^2\rangle + \frac{(is)^3}{3!}\langle x^3\rangle + \dots\right] =$$

$$= is\langle x\rangle + \frac{(is)^2}{2!}\left[\langle x^2\rangle - \langle x\rangle^2\right] \tag{28}$$

$$+ \frac{(is)^3}{3!}\left[\langle x^3\rangle - 3\langle x^2\rangle\langle x\rangle + \langle x\rangle^3\right]$$

$$+ \frac{(is)^4}{4!}\left[\langle x^4\rangle - 4\langle x\rangle\langle x^3\rangle - 3\langle x^2\rangle^2 + 12\langle x\rangle^2\langle x^2\rangle - 6\langle x\rangle^4\right] + \dots$$

We see that when the log of a series is expanded we get another series whose terms are combinations of the originals. These combinations are called <u>cumulants</u>. Notice the appearance of the mean and the variance. They are the first two cumulants

$$\kappa_1 = \langle x\rangle$$

$$\kappa_2 = \mathrm{Var}(x) = \langle x^2\rangle - \langle x\rangle^2 \tag{29}$$

$$\kappa_3 = \langle x^3\rangle - 3\langle x^2\rangle\langle x\rangle + \langle x\rangle^4$$

$$\kappa_4 = \langle x^4\rangle - 4\langle x\rangle\langle x^3\rangle - 3\langle x^2\rangle^2 + 12\langle x\rangle^2\langle x^2\rangle - 6\langle x\rangle^4$$

Now let $Y = X_1 + \dots + X_m$, where the X_i are independent r.v.s the CF of Y is given in (17) so the CGF will be

$$C_Y(s) = \ln G_Y(s) = \ln G_{X_1} + \dots + \ln G_{X_m}$$

or

$$C_Y(s) = C_{X_1}(s) + \dots + C_{X_m}(s) \tag{30}$$

The CGF is an <u>extensive quantity</u>. The CGF of the sum of indep. variables is the sum of the individual CGFs.

Now suppose you expand each of the C_{X_i} in terms of their cumulants

$$C_X(s) = \sum_{j=0}^{\infty} \frac{(is)^j}{j!} \kappa_j(X) \tag{31}$$

where $\kappa_j(X)$ is the j-th cumulant of X. Then, according to (30), we will have

$$\boxed{\kappa_j(Y) = \kappa_j(X_1) + \dots + \kappa_j(X_m)} \tag{32}$$

The cumulants therefore are also extensive. We had already seen this for the mean and the variance. Now we see that it is true for cumulants of any order.

Oops! I forgot to say something: if you look at (29) it becomes quite evident that

$$\boxed{\kappa_j(cX) = c^j \kappa_j(X)} \tag{32'}$$

where c is an arbitrary constant.

Cumulants of the Normal and the δ

The CF of $N(\mu, \sigma^2)$ was computed in (14). The CGF will then be

$$\boxed{C(s) = is\mu - \frac{s^2\sigma^2}{2}} \tag{34}$$

This is already a polynomial in s so we conclude that the Normal has only two non-zero cumulants: the mean μ and the variance σ^2. All higher order cumulants are zero.

This property is absolutely unique of the Normal and, as we will see soon, it is the reason why it is such an important distribution.

If we take the limit $\sigma \to 0$ we obtain the δ-function distribution $\delta(x-\mu)$, with CGF $C(s) = is\mu$. Thus, the δ-function has only one non-zero cumulant.

The δ has 1 cumulant, the normal has 2. Can we come up with a distribution that has 3, 4, 5 or some other finite number of cumulants? The answer is no:

> A distribution either has 1 cumulant (the δ), 2 cumulants (the Normal) or an infinite number of cumulants (all other distributions)

Reason: it can be shown that any CGF which is a polynomial in s of order larger than two will lead to a PDF that cannot be positive everywhere.

The central limit theorem

Now I want to show you the reason why the normal is so important. Consider N independent r.v.s X_i, each with its own distribution. Then let

$$Y = \frac{X_1 + \dots + X_N}{N} \qquad (35)$$

The central limit theorem states that, when N is large Y will tend to a normal distribution

$$\boxed{N \text{ very large}: \quad Y \sim N(\mu, \sigma^2)} \qquad (36)$$

In words: if you sum a bunch of r.v.s the distribution becomes approximately normal. The parameters μ and σ^2 of the distribution are very easy to figure out since they are the mean and variance of Y. From (35),

$$\mu = \langle Y \rangle = \frac{\langle X_1 \rangle + \dots + \langle X_N \rangle}{N} \qquad (37)$$

Moreover, recalling that $\mathrm{Var}(cX) = c^2 \, \mathrm{Var}(X)$ we get

$$\sigma^2 = \mathrm{Var}(Y) = \frac{\mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_N)}{N^2} \qquad (38)$$

The moral of the story is that whenever you have a random quantity which appears as the result of a sum of a large number of events, the distribution will be approximately Gaussian. It is for this reason that we assume, for instance, that the uncertainties in experimental physics are Gaussian.

Proof of the central limit theorem

To prove (36) we actually start with a slightly different random variable:

$$R = \frac{x_1 + \dots + x_N}{\sqrt{N}} \tag{39}$$

comparing with (35) we see that $Y = R/\sqrt{N}$. The reason for this weird choice will become clear in a second.

For simplicity I will also assume that the x_i are identically distributed. The theorem also holds when they are not (they only need to be independent).

Let us look at the cumulants of R: using (32) we get

$$\begin{aligned}
\kappa_j(R) &= \kappa_j(x_1/\sqrt{N}) + \dots + \kappa_j(x_N/\sqrt{N}) \quad \leftarrow \text{Eq (32')} \\
&= \frac{\kappa_j(x_1)}{N^{j/2}} + \dots + \frac{\kappa_j(x_N)}{N^{j/2}} \\
&= N\frac{\kappa_j(x)}{N^{j/2}} \quad \leftarrow \text{they are iid}
\end{aligned}$$

Thus

$$\kappa_j(R) = \frac{\kappa_j(x)}{N^{j/2-1}} \tag{40}$$

For the first few values of j this reads

$$\begin{aligned}
\kappa_1(R) &= \kappa_1(x)\sqrt{N} \\
\kappa_2(R) &= \kappa_2(x) \\
\kappa_3(R) &= \frac{\kappa_3(x)}{\sqrt{N}} \\
\kappa_4(R) &= \frac{\kappa_4(x)}{N}
\end{aligned} \tag{41}$$

etc.

We see from this result that, when N becomes large all cumulants of order $j \geq 3$ will become vanishingly small, so only the first two cumulants will survive. But we have just seen that the distribution with only two cumulants is the Gaussian/Normal. Hence, when N is large, R will become approximately Normal. Consequently, so will $Y = R/\sqrt{N}$, which completes the proof.

If the x_i are iid, Eqs (37) and (38) become

$$\begin{aligned} \langle Y \rangle &= \langle x \rangle \\ \mathrm{Var}(Y) &= \frac{\mathrm{Var}(x)}{N} \end{aligned} \tag{42}$$

Before we were talking about N "large". We now see that if $N \to \infty$ the second cumulant will also vanish and only the mean survives. We may write this mathematically as

$$\boxed{\lim_{N\to\infty} \frac{x_1 + \dots + x_N}{N} = \langle x \rangle} \tag{43}$$

This is known as the <u>law of large numbers</u>: the arithmetic mean of a sufficiently large number of random variables tend to the actual mean. In this case Y becomes a δ-distribution.

These results explain why in thermodynamic experiments we are able to measure deterministic quantities. In principle the response of the system should be a random variable, but the number of particles is so large that, by the law of large numbers, we end up getting a deterministic result.

Example: Bin(n,p) and Pois(λ)

We know that Bin(n,p) is the sum of n Bernoulli trials. that is, if $X_i \sim$ Bern(p) then

$$Y = X_1 + \ldots + X_n \sim \text{Bin}(n,p) \tag{47}$$

By the central limit theorem we expect that when $n \to \infty$ the Binomial will tend to a Gaussian, so

$$Y \sim N(\mu, \sigma^2)$$

with parameters

$$\begin{aligned} \mu &= \langle Y \rangle = np \\ \sigma^2 &= \text{Var}(Y) = np(1-p) \end{aligned} \tag{48}$$

We also showed that Bin(n,p) will tend to Pois(λ) when $n \to \infty$, $p \to 0$ but $\lambda = np$ remains finite. Consequently we conclude that when λ is large, Pois(λ) will also tend to a Normal with $\mu = \lambda$ and $\sigma^2 = \lambda$ (the mean and variance of the Poisson.

Random Walk

Summing random variables is the basic idea of a random walk. If X_i is the step the walker takes at time i, then after a time t the walker will be at position

$$Y_t = X_1 + X_2 + \dots + X_t \tag{49}$$

It is useful to remember that, in the random walk problem, it is very easy to find the mean and variance:

$$\begin{aligned} \langle Y_t \rangle &= \langle X_1 \rangle + \dots + \langle X_t \rangle \\ \mathrm{Var}(Y_t) &= \mathrm{Var}(X_1) + \dots + \mathrm{Var}(Y_t) \end{aligned} \tag{50}$$

As we discussed around Eq (17), it is also easy to find the CF. But the PDF in general is quite complicated. However, if the number of steps is large we may invoke the central limit theorem and approximate the PDF of Y_t to a Gaussian.

Let me illustrate these ideas with a concrete example. Suppose that the walker may take a step to the right with prob. p_+, to the left with prob. p_- or stay still with prob. $1 - p_+ - p_-$. That is

$$\begin{aligned} P(X_i = 1) &= p_+ \\ P(X_i = -1) &= p_- \\ P(X_i = 0) &= 1 - p_+ - p_- = p_0 \end{aligned} \tag{51}$$

we have :

$$\mu := \langle X_i \rangle = (1)\, p_+ + (0)\, p_0 + (-1)\, p_- = p_+ - p_-$$

$$\langle X_i^2 \rangle = (1)^2 p_+ + (0)^2 p_0 + (-1)^2 p_- = p_+ + p_- \qquad (52)$$

$$\Rightarrow \quad \sigma^2 := \mathrm{Var}(X_i) = (p_+ + p_-) - (p_+ - p_-)^2$$

The mean and variance of Y will then be

$$\langle Y_t \rangle = \mu t \qquad (53)$$

$$\mathrm{Var}(Y_t) = \sigma^2 t$$

we may also compute CFs :

$$G_{X_i}(s) = \langle e^{isX_i} \rangle = e^{is} p_+ + p_0 + e^{-is} p_- \qquad (54)$$

thus the CF of Y_t will be

$$\boxed{G_{Y_t}(s) = \left(p_0 + e^{is} p_+ + e^{-is} p_- \right)^t} \qquad (55)$$

we could try to use the inverse Fourier transform to find the PDF of Y_t [Eq (2)]. But that will turn out to be quite complicated. To see why, let us look at $t = 2$ (only two steps).

$$G_{Y_2}(s) = (p_0 + e^{is} p_+ + e^{-is} p_-)^2$$

$$= p_0^2 + p_+ p_- + e^{2is} p_+^2 + e^{-2is} p_-^2 + 2 p_0 p_+ e^{is} + 2 p_0 p_- e^{-is} \qquad (56)$$

We may interpret these terms by recalling Eq (24): a term in the characteristic function of the form e^{isk} is translated into a term in the PDF of the form

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} ds\, e^{-isx}\, e^{isk} = \delta(x-k) \qquad (57)$$

Thus, the PDF associated with (56) will be

$$\begin{aligned}
P_{Y_2}(y) = \left[p_0^2 + p_+ p_- \right] \delta(y) \; + & \quad \text{Prob. of staying in the same place} \\
+\; p_+^2\, \delta(y-2) & \quad \text{Prob. of 2 steps to the right} \\
+\; p_-^2\, \delta(y+2) & \quad \text{Prob. of 2 steps to the left} \\
+\; 2p_0 p_+\, \delta(y-1) & \quad \text{Prob. of 1 step to the right} \\
+\; 2p_0 p_-\, \delta(y+1) & \quad \text{Prob. of 1 step to the left.}
\end{aligned} \qquad (58)$$

You can probably see that for $t=3,4,5,\ldots$ things will start to get quite complicated.

We may consider instead the case of large t. According to the central limit theorem, in this case we may approximate Y_t by a Gaussian. The parameters of the distribution we already know; they are given in (53). Thus, the PDF of Y_t will be

$$\boxed{P_{Y_t}(y) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left\{ - \frac{(y-\nu t)^2}{2\sigma^2 t} \right\}} \qquad (59)$$

The diffusion equation

Eq (59) actually corresponds to the continuum limit, where each step of the random walk is very tiny, but takes place in a small time interval.

You may verify that $p(y,t)$ satisfies the diffusion equation

$$\boxed{\frac{\partial p}{\partial t} = -\mu \frac{\partial p}{\partial y} + \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial y^2}} \tag{60}$$

The first term is called the drift term and the second term is the diffusion term.

We can check that the solution of (60) is Eq (59) directly. Or we can pretend we do not know (59) and try to derive it. That is most easily done using (again) the characteristic function.

Let

$$p(y,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} ds \, e^{-isy} \, G(s,t) \tag{61}$$

Then

$$\frac{\partial p}{\partial t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} ds \, e^{-isy} \frac{\partial G}{\partial t} \tag{62}$$

and

$$\frac{\partial p}{\partial y} = \frac{1}{2\pi} \int_{-\infty}^{\infty} ds \, \frac{\partial}{\partial y} e^{-isy} \, G(s,t)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} ds \, (-is) \, e^{-isy} \, G(s,t) \tag{63}$$

Similarly

$$\frac{\partial^2 p}{\partial y^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} ds\,(-s^2)\,e^{-isy}\,G(s,t) \qquad (64)$$

Combining (62)-(64) into (60) we then get

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} ds\; e^{-isy}\,\frac{\partial G}{\partial t} = \frac{1}{2\pi}\int_{-\infty}^{\infty} ds\; e^{-isy}\left\{ is\mu\,G - \frac{s^2\sigma^2}{2}\,G\right\}$$

The equality must also hold for the integrands due to the uniqueness of the Fourier transform. Thus

$$\frac{\partial G}{\partial t} = \left\{ is\mu - \frac{s^2\sigma^2}{2}\right\} G \qquad (65)$$

The solution of this Equation is

$$G(s,t) = G_0(s)\,e^{\left(is\mu - \frac{s^2\sigma^2}{2}\right)t} \qquad (66)$$

where $G_0(s)$ is related to the initial condition. If we assume that the walker was, for sure, at $y=0$ at $t=0$, then $p(y,0) = \delta(y)$ and, consequently,

$$G_0(s) = \int_{-\infty}^{\infty} dy\; p(y,0)\,e^{isy} = e^{is(0)} = 1$$

Thus

$$\boxed{G(s,t) = \exp\left\{ is\mu t - \frac{s^2(\sigma^2 t)}{2}\right\}} \qquad (67)$$

this is the CF of a Gaussian, Eq (14), with mean μt and variance $\sigma^2 t$. Consequently, $p(y)$ <u>must</u> be given by (59).