1.9 The Binomial Distribution and the Random Walk

In this section, we shall examine another application of statistics to polymer chemistry. This time we shall consider the average spatial dimensions of an isolated polymer molecule, especially with regard to its dependence on molecular weight. In this context, we disregard the fact that polymer samples show polydispersity: The distribution of dimensions we are concerned with would be present even if all polymer chains were identical in length. This is because successive chain units are able to rotate along the bonds of the backbone to acquire a jumbled conformation. No two molecules have identical shapes or dimensions, even if the sample is monodisperse. Likewise, the conformation of any particular molecule continuously changes with time owing to thermal fluctuations. Only a statistical description is adequate to characterize this situation.

As a consequence of these various possible conformations, the polymer chains exist as coils with spherical symmetry. Our eventual goal is to describe these three-dimensional structures, although some preliminary considerations must be taken up first. Accordingly, we begin by discussing a statistical exercise called a one-dimensional random walk.

We start this exercise by considering the placement of \( n \) successive repeat units in a polymer chain along a straight line, say, the \( x \) axis. We assume that the chain is perfectly flexible and that it excludes a negligible volume so that more than one repeat unit can be placed on the same site. We anchor one end of the chain at the origin of the axis and propose to use the toss of a coin to decide on the placement of successive units; that is, if the coin turns up heads (subscript \( H \)), we place the next unit one step ahead an increment of +1 in the \( x \) direction, where 1 is the length of the repeat unit. If the coin shows tails (subscript \( T \)), the next unit is placed back a step for a change in \( x \) value of -1. How far from the origin will the other end of the chain lie after \( n \) units are placed on the basis of this random walk? A possible sequence for 10 tosses might be \( HHTTTTTTTTT \) in a particular exercise; this amounts to 6 steps forward and 4 backward for a net displacement of +2. In this example, the two ends of the chain are separated by only 21, even though the fully extended chain would have a length of 10 1. Since the outcome of 10 tosses could be different in another trial of the same exercise, it is clear that we must turn to statistics to describe the average placement of the chain.
For a one-dimensional random walk, the probability of \( n_H \) heads after \( n \) tosses is supplied by application of the binomial distribution formula:

\[
P(n_H, n) = \frac{n!}{n_H! n_T!} p_H^{n_H} p_T^{n_T}
\]

(1.21)

which the \( p \)'s are the probabilities of either a head or tail in a single toss and the subscripted \( n \)'s are the number of heads and tails in the specific exercise, for a fair coin \( p_H = p_T = \frac{1}{2} \), but for the time being we shall continue using the more general formula.

The binomial distribution function is one of the most fundamental equations in statistics and finds several applications in this volume. To be sure that we appreciate its significance, we make the following observations about the plausibility of Eq. (1.21):

To evaluate the probability that one event and another will occur, we multiply the probabilities of the individual events. Thus the probability of tossing two heads is \( p_H^2 \) and that of tossing \( n_H \) heads is \( p_H^{n_H} \).

Likewise, the probability of tossing \( n_T \) tails is \( p_T^{n_T} \). The probability of tossing \( n_H \) heads and \( n_T \) tails is \( p_H^{n_H} p_T^{n_T} \) by the same principle.

The probability calculated so far is too low because it describes one specific sequence of heads and tails. From the point of view of net displacement, the sequence does not matter. Hence the above results must be multiplied by the number of different ways this outcome can arise. Instead of tossing one coin \( n \) times, we could toss \( n \) coins drawn at random from a piggy bank. For the first, we have a choice of \( n_H \) to draw from; for the second, \( n_H - 1 \); for the third, \( n_H - 2 \), and so on. The total possible ways the toss could be carried out is given by the product of these different choices, that is by \( n! \).

This suggests that \( n! p_H^{n_H} p_T^{n_T} \) gives the desired probability, but now we have gone too far in the opposite direction and overcounted the probability. To appreciate this fact, we recognize that among the \( n! \) ways the coins could be tossed, we have included a number of ways that yield the same net outcome achieved through different sequences of tosses. For example, the \( n! \) count would include "HHHHHTT", "HTHHTHT", and "THHTHHT" as different, although they each consist of the same number of heads and tails.

To correct for this overcounting, we cancel out the number of ways \( n_H \) heads can be permuted and the number of ways \( n_T \) tails can be permuted. Using the same logic as in item (3), these redundant possibilities are given by \( n_H! \) and \( n_T! \), respectively. Dividing the result in item (4) by these factorials gives Eq. (1.21).
If we apply Eq. (1.21) to the problem of determining the number of heads occurring in 10 tosses, we find that the probability of, say, 6 heads is given by \( \frac{10!}{6!4!} \frac{1}{2^{10}} = 0.205 \). That is, heads would occur in 6 of every 10 tosses about 20% of the time. If a large number of 10-mer molecules were being positioned along an axis according to this hypothetical procedure, in about 20% of the cases the end-to-end distance would be 21. For an outcome of 8 heads and 2 tails, corresponding to a net forward displacement of 6 steps or an end-to-end distance of 61, the probability is \( \frac{10!}{8!2!} \frac{1}{2^{10}} = 9.77 \times 10^{-4} \approx 0.1\% \).

It is an easy matter to rewrite Eq. (1.21) in terms of the probability of a displacement \( x \) occurring after \( n \) tosses. We recognize the following:

1. Each toss is either a head or a tail:

   \[ n = n_H + n_T \]  
   \[ \text{(1.22)} \]

2. The displacement \( x \) after \( n \) steps of length 1 is

   \[ x = (n_H - n_T) \]  
   \[ \text{(1.23)} \]

3. Solving Eqs. (1.22) and (1.23) for \( n_H \) and \( n_T \) gives

   \[ n_H = \frac{1}{2} (n + x/l) \]  
   \[ \text{(1.24)} \]

   and

   \[ n_T = \frac{1}{2} (n - x/l) \]  
   \[ \text{(1.25)} \]

4. For a fair coin \( p_H = p_T = \frac{1}{2} \); there Eq. (1.21) becomes

   \[ P(n_H, n) = \frac{n!}{[(n + x/l)/2]! [(n - x/l)/2]!} \left( \frac{1}{2} \right)^n \]  
   \[ \text{(1.26)} \]

This result enables us to calculate the probability of any specified outcome for the one-dimensional random walk. We shall continue to develop this one-dimensional relationship somewhat further, since doing so will produce some useful results.

5. For high molecular weight polymers, \( n \) is large and the logarithm of large factorials is accurately given by Sterling's approximation,

   \[ \ln y! \approx y \ln y - y \]  
   \[ \text{(1.27)} \]

for large \( y \).
Taking logarithms of both sides in Eq. (1.26) and applying Sterling’s approximation gives

\[-\ln P(x, n) = \frac{nl + x}{2l} \ln \left(1 + \frac{x}{nl}\right) + \frac{nl - x}{2l} \ln \left(1 - \frac{x}{nl}\right)\]  

(1.28)

The number of steps is always much larger than the displacement x, since there is a good deal of back-and-forth cancellation. Hence the ratio x/\(nl\) is less than unity and the logarithms may be approximated by the leading terms of a series expansion

\[\ln \left(1 + \frac{x}{nl}\right) \approx \frac{x}{nl} - \frac{1}{2} \left(\frac{x}{nl}\right)^2 + \cdots\]  

(1.29)

in which x can be either positive or negative.

Applying Eq. (1.29) to Eq. (1.28) gives

\[-\ln P(x, n) \approx -\frac{x^2}{2nl^2} + \cdots\]  

(1.30)

or

\[P(x, n) = k \exp\left(-\frac{x^2}{2nl^2}\right)\]  

(1.31)

where the factor k is a constant called a normalization factor.

Well-behaved probability functions total unity when they are summed over all possible outcomes. Since Eq. (1.31) is a continuous function, this has been accomplished by getting rid of the factorials—this sum may be written as an integral over all possible values of x:

\[k \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2nl^2}\right) \, dx = 1\]  

(1.32)

A probability function which satisfies this criterion is said to be normalized. This will be accomplished when a value of k which satisfies Eq. (1.32) is found.

The integral in Eq. (1.32) is known as a gamma function and may be found in tables of integrals. The result of the integration is that

\[k = (2\pi nl^2)^{-\frac{1}{2}}\]  

(1.33)
Therefore an expression which is equivalent to Eq. (1.21) for the case of large \( n \)'s is

\[
P(x, n) = (2\pi nl^2)^{-\frac{1}{2}} \exp \left( \frac{-x^2}{2nl^2} \right)
\]  
(1.34)

Since we have ended up with a continuous distribution function, it is more appropriate to multiply both sides of Eq. (1.34) by \( dx \) and to say that the equation gives the probability of \( x \) values between \( x \) and \( x + dx \) for \( n \) steps of length 1.

In the next section we shall adapt this probability function to the description of a three-dimensional coil. We conclude this section by noting that Eq. (1.21) may be approximated by two other functions which are used elsewhere in this book. For these general relationships we define \( \nu \) to be the number of successes—that is, some specified outcome such as tossing a head—out of \( n \) tries and define \( p \) as the probability of success in a single try. In this amended notation, Eq. (1.21) becomes

\[
P(\nu, n) = \frac{n!}{\nu!(n-\nu)!} p^\nu (1-p)^{n-\nu}
\]  
(1.35)

In terms of this notation we have the following:

1. The average number of successes is given by

\[
\bar{\nu} = np
\]  
(1.36)

and the standard deviation of the distribution of successes is

\[
\sigma = \sqrt{np(1-p)}
\]  
(1.37)

2. If \( n \) is large and \( p \) is very small, Eq. (1.35) is approximated by the Poisson distribution

\[
P_{\text{Poiss}} = \frac{e^{-\bar{\nu}} \bar{\nu}^\nu}{\nu!}
\]  
(1.38)

3. As an example of the Poisson distribution, consider the case of \( n = 1000 \), \( p = 0.01 \), and \( \nu = 5 \). According to the binomial distribution, \( P_{\text{bin}} = (1000!/995!5!)(0.01)^5(0.99)^{995} = 0.0375 \). By Eq. (1.36), \( \bar{\nu} = (0.01)(1000) = 10 \) and \( P_{\text{Poiss}} = e^{-10} \times 10^5/5! = 0.0378 \).
If $n$ is large, but without any particular restriction on $p$, Eq. (1.35) is approximated by the Gaussian or normal distribution

$$P_{\text{norm}} = \frac{1}{(2\pi)^{1/2}} \exp \left[-\frac{1}{2} \left( \frac{\nu - \bar{\nu}}{\sigma} \right)^2 \right]$$  \hspace{1cm} (1.39)

As an example of the normal distribution, consider the case of $n = 60$, $p = 0.20$, and $\nu = 10$. According to the binomial distribution, $P_{\text{bin}} = (60)! / 50!10!(0.2)^{10}(0.8)^{50} = 0.110$. By Eq. (1.36), $\bar{\nu} = (0.2)(60) = 12$, and by (1.37), $\sigma = [(60)(0.2)(0.8)]^{1/2} = 3.10$, therefore

$$P_{\text{norm}} = \frac{1}{3.10(2\pi)^{1/2}} \exp \left[-\frac{1}{2} \left( \frac{10 - 12}{3.10} \right)^2 \right] = 0.105$$

The proof that these expressions are equivalent to Eq. (1.35) under suitable conditions is found in statistics textbooks. We shall have occasion to use the Poisson approximation to the binomial in discussing crystallization of polymers in Chap. 4, and the distribution of molecular weights of certain polymers in Chap. 6. The normal distribution is the familiar bell-shaped distribution that is known in academic circles as “the curve.” We shall use it in discussing diffusion in Chap. 9.

10 Size Parameters for Random Coils

Next let us apply random walk statistics to three-dimensional chains. We begin by assuming isolated polymer molecules which consist of perfectly flexible chains.

To isolate polymer chains from one another, we consider a solution which is efficiently dilute that the domains of the individual polymer molecules are well separated from each other. For the present, we assume the solvent has no influence on the polymer but merely supports the molecule. In fact, this is not always the case, although it can be achieved by proper choice of solvent or temperature.

The chain is imagined to consist of $n$ units connected to one another by perfectly flexible joints. Figure 1.5a shows the $i$th and $(i + 1)$th unit along the backbone of a vinyl polymer according to this model. The three carbon atoms involved in these bonds are numbered for ease of discussing the model. We focus attention at carbon 2, which we picture as lying at the center of a sphere. By hypothesis, the chain is perfectly flexible at $C_2$ so that the angle $\theta$ between the two bonds can have any value between $0^\circ$ and $360^\circ$. Likewise, regardless of the value of $\theta$, there is complete freedom of rotation around bond $i$ connecting $C_1$ and $C_2$; that is, the angle $\phi$ may take on all values between $0^\circ$ and $360^\circ$. 
This kind of perfect flexibility means that $C_3$ may lie anywhere on the surface of the sphere. According to the model, it is not even excluded from $C_1$. This model of a perfectly flexible chain is not a realistic representation of an actual polymer molecule. The latter is subject to fixed bond angles and experiences some degree of hindrance to rotation around bonds. We shall consider the effect of these constraints, as well as the effect of solvent-polymer interactions, after we explore the properties of the perfectly flexible chain. Even in this revised model, we shall not correct for the volume excluded by the polymer chain itself.

If a chain consists of $m$ bonds of this sort, then there are $m + 1$ atoms which can act as centers like $C_2$ in Fig. 1.5. The values of the angles at each successive center are assumed to be independent, so the chain can twist and bend in an enormous number of conformations along its length. This is how the random coil comes about. Figure 1.5b shows a somewhat longer segment of chain. Each carbon atom has the possibility of being anywhere on the surface of the sphere centered at the preceding carbon according to this model.

The one-dimensional random walk of the last section is readily adapted to this problem once we recognize the following connection. As before, we imagine that one end of the chain is anchored at the origin of a three-dimensional coordinate system. Our interest is in knowing, on the average, what will be the distance of the other end of the chain from this origin. A moment's reflection will convince us that the $x$, $y$, and $z$ directions are all equally probable as far as the perfectly flexible chain is concerned. Therefore one-third of the repeat units will be associated with each of the three perpendicular directions.

![Figure 1.5 Placement of successive polymer segments connected by perfectly flexible joints. In (a), the $i$th and $(i + 1)$th bond can be moved through angles $\phi$ and $\theta$ so that carbon $3$ can lie anywhere on the surface of a sphere. In (b), the pattern is illustrated for a longer portion of chain.](image)
of the coordinates system. The probability of a displacement \( x \) in that direction is given by Eq. (1.34), with \( n \) replaced by \( n/3 \). Since the \( x \), \( y \), and \( z \) directions are equivalent, the same expression holds for \( P(y, n/3) \) and \( P(z, n/3) \). Thus the probability that the loose end of the chain will be found in a volume element \( dx \, dy \, dz \) located at a specific set of \( x \), \( y \), \( z \) values is the probability that is has all of the following:

1. A value of \( x \) between \( x \) and \( x + dx \).
2. A value of \( y \) between \( y \) and \( y + dy \).
3. A value of \( z \) between \( z \) and \( z + dz \).

According to the rules for compounding probabilities, this is given by

\[
P(x, y, z, n) \, dx \, dy \, dz = P(x, n/3) \, P(y, n/3) \, P(z, n/3) \, dx \, dy \, dz \quad (1.40)
\]

Substituting Eq. (1.34) with \( n \) replaced by \( n/3 \) for each of the probabilities in (1.40) gives

\[
P(x, y, z, n) \, dx \, dy \, dz = \left( \frac{2\pi^{n/3} l^2}{3} \right)^{-3/2} \exp \left( -\frac{3(x^2 + y^2 + z^2)}{2nl^2} \right) \, dx \, dy \, dz \quad (1.41)
\]

The \( x \), \( y \), and \( z \) coordinates of the loose end of the chain can be related to the radial distance \( r \) from the origin by

\[
x^2 + y^2 + z^2 = r^2 \quad (1.42)
\]

In addition, the volume element of interest is not the box \( dx \, dy \, dz \) shown in Fig. 1.6a but, rather, a spherical shell of radius \( r \) and thickness \( dr \) as shown in Fig. 1.6b. The result of expressing the volume element in spherical coordinates and integrating over all angles is the replacement

\[
dx \, dy \, dz \rightarrow 4\pi r^2 \, dr \quad (1.43)
\]

Making these substitutions gives the probability of finding one end of a perfectly flexible chain of \( n \) units a distance \( r \) from the other end by

\[
P(r, n) \, dr = \left( \frac{2\pi^{n/3} l^2}{3} \right)^{-3/2} 4\pi r^2 \exp \left( -\frac{3r^2}{2nl^2} \right) \, dr \quad (1.44)
\]

Several features of this expression should be noted:

1. The probability is normalized, since Eq. (1.44) was assembled from separately normalized components.
2. The factor containing \( r^2 \) increases with increasing \( r \) and reflects the fact that there are more locations to place the loose chain end within larger spherical shells, but \ldots\.

3. The exponential factor decreases with increasing \( r \) and reflects the fact that large displacements become increasingly probable.

4. The overall probability function for the end-to-end distance is the product of these two considerations. Starting at \( r = 0 \), the probability increases owing to the \( r^2 \) term, passes through a maximum, then decreases as the exponential factor takes over at large \( r \) values.

5. This factor is reminiscent of the radial distribution function for electron probability in an atom and the Maxwell distribution of molecular velocities in a gas, both of which pass through a maximum for similar reasons.

Figure 1.6 A flexible coil attached at the origin at one end and (a) in a volume element \( dx \, dy \, dz \) at the other end and (b) in a spherical shell of volume \( 4\pi r^2 \, dr \). (Reprinted from Ref. 4, p. 116.)