Chapter 1
A First Glimpse of Stochastic Processes

In this introductory chapter we will give a short overview of the history of *probability theory* and *stochastic processes*, and then we will discuss the properties of a simple example of a stochastic process, namely the *random walk* in one dimension. This example will introduce us to many of the typical questions that arise in situations involving *randomness* and to the tools for tackling them, which we will formalize and expand on in subsequent chapters.

1.1 Some History

Let us start this historical introduction with a quote from the superb review article *On the Wonderful World of Random Walks* by E.W. Montroll and M.F. Shlesinger [145], which also contains a more detailed historical account of the development of probability theory:

> *Since traveling was onerous (and expensive), and eating, hunting and wenching generally did not fill the 17th century gentleman's day, two possibilities remained to occupy the empty hours, praying and gambling; many preferred the latter.*

In fact, it is in the area of gambling that the theory of probability and stochastic processes has its origin. People had always engaged in gambling, but it was only through the thinking of the Enlightenment that the outcome of a gambling game was no longer seen as a divine decision, but became amenable to rational thinking and speculation. One of these 17th century gentlemen, a certain Chevalier de Méré, is reported to have posed a question concerning the odds at a gambling game to Pascal (1623–1662). The ensuing exchange of letters between Pascal and Fermat (1601–1665) on this problem is generally seen as the starting point of probability theory.

The first book on probability theory was written by Christiaan Huygens (1629–1695) in 1657 and had the title *De Ratiociniis in Ludo Aleae* (On Reasoning in the Game of Dice). The first mathematical treatise on probability theory in the modern sense was Jakob Bernoulli's (1662–1705) book *Ars Conjectandi* (The Art of Conjecturing), which was published posthumously in 1713. It contained

W. Paul, J. Baschnagel, *Stochastic Processes*, DOI 10.1007/978-3-319-00327-6_1,

- a critical discussion of Huygens' book,
- *combinatorics*, as it is taught today,
- probabilities in the context of gambling games, and
- an application of probability theory to daily problems, especially in *economics*.

We can see that even at the beginning of the 18th century the two key ingredients responsible for the importance of stochastic ideas in economics today are already discernable: combine a probabilistic description of economic processes with the control of risks and odds in gambling games, and you get the risk-management in financial markets that has seen such an upsurge of interest in the last 30 years.

A decisive step in the stochastic treatment of the price of financial assets was made in the Ph.D. thesis of Louis Bachelier (1870–1946), *Théorie de la Spéculation*, for which he obtained his Ph.D. in mathematics on March 19, 1900. His advisor was the famous mathematician Henri Poincaré (1854–1912), who is also well known for his contributions to theoretical physics. The thesis is very remarkable at least for two reasons:

- It already contained many results of the theory of stochastic processes as it stands today which were only later mathematically formalized.
- It was so completely ignored that even Poincaré forgot that it contained the solution to the *Brownian motion* problem when he later started to work on that problem.

Brownian motion is the archetypical problem in the theory of stochastic processes. In 1827 the Scottish botanist Robert Brown had reported the observation of a very irregular motion displayed by a pollen particle immersed in a fluid. It was the *kinetic theory of gases*, dating back to the book *Hydrodynamica, sive de viribus et motibus fluidorum commentarii* (Hydrodynamics, or commentaries on the forces and motions of fluids) by Daniel Bernoulli (1700–1782), published in 1738, which would provide the basis for Einstein's (1879–1955) and Smoluchowski's (1872–1917) successful treatments of the Brownian motion problem in 1905 and 1906, respectively. Through the work of Maxwell (1831–1879) and Boltzmann (1844–1906), *Statistical Mechanics*, as it grew out of the kinetic theory of gases, was the main area of application of probabilistic concepts in theoretical physics in the 19th century.

In the Brownian motion problem and all its variants—whether in physics, chemistry and biology or in finance, sociology and politics—one deals with a phenomenon (motion of the pollen particle, daily change in a stock market index) that is the outcome of many unpredictable and sometimes unobservable events (collisions with the particles of the surrounding liquid, buy/sell decisions of the single investor) which individually contribute a negligible amount to the observed phenomenon, but collectively lead to an observable effect. The individual events cannot sensibly be treated in detail, but their statistical properties may be known, and, in the end, it is these that determine the observed macroscopic behavior.

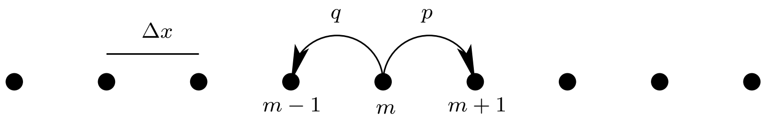

Fig. 1.1 Random walker on a 1-dimensional lattice of sites that are a fixed distance Δx apart. The walker jumps to the right with probability p and to the left with $q = 1 - p$

Another problem, which is closely related to Brownian motion and which we will examine in the next section, is that of a *random walker*. This concept was introduced into science by Karl Pearson (1857–1936) in a letter to Nature in 1905:

> *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 straight line. He repeats this process n times. I require the probability that after these n stretches he is at a distance between r and $r + \delta r$ from his starting point* 0.

The solution was provided in the same volume of Nature by Lord Rayleigh (1842–1919), who pointed out that he had solved this problem 25 years earlier when studying the superposition of sound waves of equal frequency and amplitude but with random phases.

The random walker, however, is still with us today and we will now turn to it.

1.2 Random Walk on a Line

Let us assume that a walker can sit at regularly spaced positions along a line that are a distance Δx apart (see Fig. 1.1); so we can label the positions by the set of whole numbers, $\mathbb{Z}$. Furthermore we require the walker to be at position 0 at time 0. After fixed time intervals Δt the walker either jumps to the right with probability p or to the left with probability $q = 1 - p$; so we can work with discrete time points, labeled by the natural numbers including zero, $\mathbb{N}_0$.

Our aim is to answer the following question: What is the probability $p(m, N)$ that the walker will be at position m after N steps?

For $m < N$ there are many ways to start at 0, go through N jumps to nearest-neighbor sites, and end up at m. But since all these possibilities are independent of each other we have to add up their probabilities. For all these ways we know that the walker must have made $m + l$ jumps to the right and l jumps to the left; and since $m + 2l = N$, the walker must have made

- $(N + m)/2$ jumps to the right and
- $(N - m)/2$ jumps to the left.

So whenever N is even, so is m. Furthermore we know that the probability for the next jump is always p to the right and q to the left, irrespective of what the path of the walker up to that point was. The probability for a sequence of left and right jumps is the product of the probabilities of the individual jumps. Since the probability of the individual jumps does not depend on their position in the sequence, all paths starting at 0 and ending at m have the same overall probability. The probability for making exactly $(N+m)/2$ jumps to the right and exactly $(N-m)/2$ jumps to the left is

$$p^{\frac{1}{2}(N+m)}q^{\frac{1}{2}(N-m)}.$$

To finally get the answer to our question, we have to find out how many such paths there are. This is given by the number of ways to make $(N+m)/2$ out of N jumps to the right (and consequently $N-(N+m)/2=(N-m)/2$ jumps to the left), where the order of the jumps does not matter (and repetitions are not allowed):

$$\frac{N!}{(\frac{N+m}{2})!(\frac{N-m}{2})!}.$$

The probability of being at position m after N jumps is therefore given as

$$p(m,N)=\frac{N!}{(\frac{N+m}{2})!(\frac{N-m}{2})!}p^{\frac{1}{2}(N+m)}(1-p)^{\frac{1}{2}(N-m)}, \tag{1.1}$$

which is the binomial distribution. If we know the probability distribution $p(m,N)$, we can calculate all the moments of m at fixed time N. Let us denote the number of jumps to the right as $r=(N+m)/2$ and write

$$p(m,N)=p_N(r)=\frac{N!}{r!(N-r)!}p^r q^{N-r} \tag{1.2}$$

and calculate the moments of $p_N(r)$. For this purpose we use the property of the binomial distribution that $p_N(r)$ is the coefficient of u^r in $(pu+q)^N$. With this trick it is easy, for instance, to convince ourselves that $p_N(r)$ is properly normalized to one

$$\sum_{r=0}^{N}p_N(r)=\left[\sum_{r=0}^{N}\binom{N}{r}u^r p^r q^{N-r}\right]_{u=1}=\left[(pu+q)^N\right]_{u=1}=1. \tag{1.3}$$

The first moment or *expectation value* of r is:

$$\begin{aligned}
\langle r\rangle &= \sum_{r=0}^{N} r p_N(r)\\
&= \left[\sum_{r=0}^{N} r\binom{N}{r}u^r p^r q^{N-r}\right]_{u=1} = \left[\sum_{r=0}^{N}\binom{N}{r}u\frac{\mathrm{d}}{\mathrm{d}u}\left(u^r p^r q^{N-r}\right)\right]_{u=1}\\
&= \left[u\frac{\mathrm{d}}{\mathrm{d}u}\sum_{r=0}^{N}\binom{N}{r}u^r p^r q^{N-r}\right]_{u=1} = \left[u\frac{\mathrm{d}}{\mathrm{d}u}(pu+q)^N\right]_{u=1}\\
&= \left[Nup(pu+q)^{N-1}\right]_{u=1}
\end{aligned}$$

leading to

$$\mathrm{E}[r] \equiv \langle r \rangle = Np. \tag{1.4}$$

In the same manner, one can derive the following for the second moment:

$$\mathrm{E}\left[r^2\right] \equiv \left\langle r^2 \right\rangle = \left[\left(u\frac{\mathrm{d}}{\mathrm{d}u}\right)^2 (pu+q)^N\right]_{u=1} = Np + N(N-1)p^2. \tag{1.5}$$

From this one can calculate the *variance* or *second central moment*

$$\mathrm{Var}[r] \equiv \sigma^2 := \left\langle \left(r - \langle r \rangle\right)^2 \right\rangle = \left\langle r^2 \right\rangle - \langle r \rangle^2 \tag{1.6}$$

of the distribution, which is a measure of the width of the distribution

$$\sigma^2 = Npq. \tag{1.7}$$

The relative width of the distribution

$$\frac{\sigma}{\langle r \rangle} = \sqrt{\frac{q}{p}} N^{-1/2} \tag{1.8}$$

goes to zero with increasing number of performed steps, N. Distributions with this property are called (strongly) *self-averaging*. This term can be understood in the following way: The outcome for r after N measurements has a statistical error of order σ. For self-averaging systems this error may be neglected relative to $\langle r \rangle$, if the system size (N) becomes large. In the large-N limit the system thus 'averages itself' and r behaves as if it was a non-random variable (with value $\langle r \rangle$). This self-averaging property is important, for instance, in statistical mechanics.

Figure 1.2 shows a plot of the binomial distribution for $N = 100$ and $p = 0.8$. As one can see, the distribution has a bell-shaped form with a maximum occurring around the average value $\langle r \rangle = Np = 80$, and for this choice of parameters it is almost symmetric around its maximum.

When we translate the results for the number of steps to the right back into the position of the random walker we get the following results

$$\langle m \rangle = 2N\left(p - \frac{1}{2}\right) \tag{1.9}$$

and

$$\left\langle m^2 \right\rangle = 4Np(1-p) + 4N^2\left(p - \frac{1}{2}\right)^2, \tag{1.10}$$

$$\sigma^2 = \left\langle m^2 \right\rangle - \langle m \rangle^2 = 4Npq. \tag{1.11}$$

In the case of symmetric jump rates, this reduces to

$$\langle m \rangle = 0 \quad \text{and} \quad \left\langle m^2 \right\rangle = N. \tag{1.12}$$

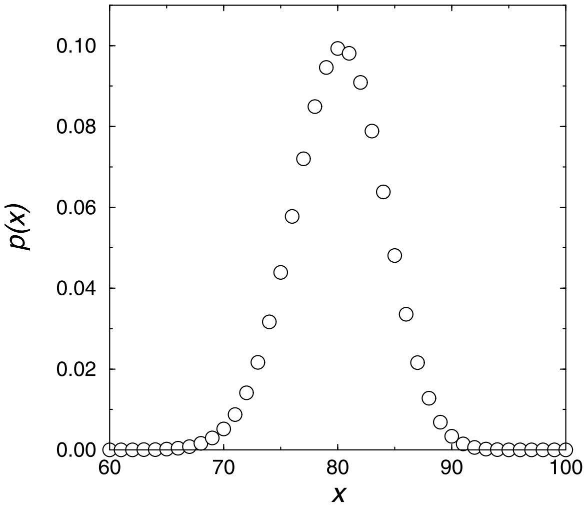

Fig. 1.2 Plot of the binomial distribution for a number of steps $N = 100$ and the probability of a jump to the right $p = 0.8$

This behavior, in which the square of the distance traveled is proportional to time, is called *free diffusion*.

1.2.1 From Binomial to Gaussian

The reader may be familiar with the description of particle diffusion in the context of partial differential equations, i.e., Fick's equation. To examine the relation between our jump process and Fickian diffusion, we will now study approximations of the binomial distribution which will be valid for large number of jumps ($N \to \infty$), i.e., for long times.

Assuming $N \gg 1$ we can use Stirling's formula to approximate the factorials in the binomial distribution

$$\ln N! = \left(N + \frac{1}{2}\right) \ln N - N + \frac{1}{2} \ln 2\pi + \mathrm{O}\big(N^{-1}\big). \tag{1.13}$$

Using Stirling's formula, we get

$$\begin{aligned}\ln p(m, N) = {} & \left(N + \frac{1}{2}\right) \ln N - \left(\frac{N+m}{2} + \frac{1}{2}\right) \ln\left[\frac{N}{2}\left(1 + \frac{m}{N}\right)\right] \\ & - \left(\frac{N-m}{2} + \frac{1}{2}\right) \ln\left[\frac{N}{2}\left(1 - \frac{m}{N}\right)\right] \\ & + \frac{N+m}{2} \ln p + \frac{N-m}{2} \ln q - \frac{1}{2} \ln 2\pi .\end{aligned}$$

Now we want to derive an approximation to the binomial distribution close to its maximum, which is also close to the expectation value $\langle m\rangle$. So let us write

$$m = \langle m\rangle + \delta m = 2Np - N + \delta m,$$

which leads to

$$\frac{N+m}{2} = Np + \frac{\delta m}{2} \quad \text{and} \quad \frac{N-m}{2} = Nq - \frac{\delta m}{2}.$$

Using these relations, we get

$$\begin{aligned}
\ln p(m,N) = {} & \left(N+\frac{1}{2}\right)\ln N - \frac{1}{2}\ln 2\pi \\
& + \left(Np+\frac{\delta m}{2}\right)\ln p + \left(Nq-\frac{\delta m}{2}\right)\ln q \\
& - \left(Np+\frac{\delta m}{2}+\frac{1}{2}\right)\ln\left[Np\left(1+\frac{\delta m}{2Np}\right)\right] \\
& - \left(Nq-\frac{\delta m}{2}+\frac{1}{2}\right)\ln\left[Nq\left(1-\frac{\delta m}{2Nq}\right)\right] \\
= {} & -\frac{1}{2}\ln(2\pi Npq) - \left(Np+\frac{\delta m}{2}+\frac{1}{2}\right)\ln\left(1+\frac{\delta m}{2Np}\right) \\
& - \left(Nq-\frac{\delta m}{2}+\frac{1}{2}\right)\ln\left(1-\frac{\delta m}{2Nq}\right).
\end{aligned}$$

Expanding the logarithm

$$\ln(1\pm x) = \pm x - \frac{1}{2}x^2 + \mathrm{O}\big(x^3\big)$$

yields

$$\ln p(m,N) \simeq -\frac{1}{2}\ln(2\pi Npq) - \frac{1}{2}\frac{(\delta m)^2}{4Npq} - \frac{\delta m(q-p)}{4Npq} + \frac{(\delta m)^2(p^2+q^2)}{16(Npq)^2}.$$

We should remember that the variance (squared width) of the binomial distribution was $\sigma^2 = 4Npq$. When we want to approximate the distribution in its center and up to fluctuations around the mean value of the order $(\delta m)^2 = \mathrm{O}(\sigma^2)$, we find for the last terms in the above equation:

$$\frac{\delta m(q-p)}{4Npq} = \mathrm{O}\big((Np)^{-1/2}\big) \quad \text{and} \quad \frac{(\delta m)^2(p^2+q^2)}{16(Npq)^2} = \mathrm{O}\big((Np)^{-1}\big).$$

These terms can be neglected if $Np \to \infty$. We therefore finally obtain

$$p(m,N) \to \frac{2}{\sqrt{2\pi 4Npq}}\exp\left[-\frac{1}{2}\frac{(\delta m)^2}{4Npq}\right], \tag{1.14}$$

which is the Gaussian (C.F. Gauss (1777–1855)) or normal distribution. The factor 2 in front of the exponential comes from the fact that for fixed N (odd or even) only every other m (odd or even, respectively) has a non-zero probability, so $\Delta m = 2$.

This distribution is called 'normal' distribution because of its ubiquity. Whenever one adds up random variables, x_i, with finite first and second moments, so $\langle x_i \rangle < \infty$ and $\langle x_i^2 \rangle < \infty$ (in our case the jump distance of the random walker is such a variable with first moment $(p-q)\Delta x$ and second moment $(p+q)(\Delta x)^2 = (\Delta x)^2$), then the sum variable

$$S_N := \frac{1}{N}\sum_{i=1}^{N} x_i$$

is distributed according to a normal distribution for $N \to \infty$. This is the gist of the *central limit theorem*, to which we will return in the next chapter.

There are, however, also cases where either $\langle x_i^2 \rangle$ or even $\langle x_i \rangle$ does not exist. In these cases the limiting distribution of the sum variable is not a Gaussian distribution but a so-called *stable* or *Lévy* distribution, named after the mathematician Paul Lévy (1886–1971), who began to study these distributions in the 1930s. The Gaussian distribution is a special case of these stable distributions. We will discuss the properties of these distributions, which have become of increasing importance in all areas of application of the theory of stochastic processes, in Chap. 4.

We now want to leave the discrete description and perform a continuum limit. Let us write

$$\begin{aligned} x &= m\Delta x, \quad \text{i.e., } \langle x \rangle = \langle m \rangle \Delta x, \\ t &= N\Delta t, \\ D &= 2pq\frac{(\Delta x)^2}{\Delta t}, \end{aligned} \tag{1.15}$$

so that we can interpret

$$p(m\Delta x, N\Delta t) = \frac{2\Delta x}{\sqrt{2\pi 2Dt}} \exp\left[-\frac{1}{2}\frac{(x-\langle x \rangle)^2}{2Dt}\right]$$

as the probability of finding our random walker in an interval of width $2\Delta x$ around a certain position x at time t. We now require that

$$\Delta x \to 0, \quad \Delta t \to 0, \quad \text{and} \quad 2pq\frac{(\Delta x)^2}{\Delta t} = D = \text{const.} \tag{1.16}$$

Here, D, with the units length^2/time, is called the *diffusion coefficient* of the walker. For the probability of finding our random walker in an interval of width dx around the position x, we get

$$p(x,t)\mathrm{d}x = \frac{1}{\sqrt{2\pi 2Dt}} \exp\left[-\frac{1}{2}\frac{(x-\langle x \rangle)^2}{2Dt}\right]\mathrm{d}x. \tag{1.17}$$

When we look closer at the definition of $\langle x \rangle$ above, we see that another assumption was buried in our limiting procedure:

$$\langle x \rangle(t) = \Delta x \langle m \rangle = 2\left(p - \frac{1}{2}\right) N \Delta x = 2\left(p - \frac{1}{2}\right) \frac{\Delta x}{\Delta t} t.$$

So our limiting procedure also has to include the requirement

$$\Delta x \to 0, \quad \Delta t \to 0 \quad \text{and} \quad \frac{2(p - \frac{1}{2})\Delta x}{\Delta t} = v = \text{const.} \tag{1.18}$$

As we have already discussed before, when $p = 1/2$ the average position of the walker stays at zero for all times and the velocity of the walker vanishes. Any asymmetry in the transition rates ($p \neq q$) produces a net velocity of the walker. However, when $v = 0$, we have $\langle x \rangle = 0$ and $\langle x^2 \rangle = 2Dt$. Finally, we can write down the probability density for the position of the random walker at time t,

$$p(x, t) = \frac{1}{\sqrt{2\pi 2Dt}} \exp\left[-\frac{1}{2} \frac{(x - vt)^2}{2Dt}\right], \tag{1.19}$$

with starting condition

$$p(x, 0) = \delta(x)$$

and boundary conditions

$$p(x, t) \overset{x \to \pm\infty}{\longrightarrow} 0.$$

By substitution one can convince oneself that (1.19) is the solution of the following partial differential equation:

$$\frac{\partial}{\partial t} p(x, t) = -v \frac{\partial}{\partial x} p(x, t) + D \frac{\partial^2}{\partial x^2} p(x, t), \tag{1.20}$$

which is Fick's equation for diffusion in the presence of a constant drift velocity.

To close the loop, we now want to derive this evolution equation for the probability density starting from the discrete random walker. For this we have to rethink our treatment of the random walker from a slightly different perspective, the perspective of *rate equations*.

How does the probability of the discrete random walker being at position m at time N change in the next time interval Δt? Since our walker is supposed to perform one jump in every time interval Δt, we can write

$$p(m, N + 1) = pp(m - 1, N) + qp(m + 1, N). \tag{1.21}$$

The walker has to jump to position m either from the position to the left or to the right of m. This is an example of a *master equation* for a stochastic process. In the next chapter we will discuss for which types of stochastic processes this evolution equation is applicable.

In order to introduce the drift velocity, v, and the diffusion coefficient, D, into this equation let us rewrite the definition of D:

$$
\begin{aligned}
D &= 2pq\frac{(\Delta x)^2}{\Delta t} \\
&= (2p-1)(1-p)\frac{(\Delta x)^2}{\Delta t} + (1-p)\frac{(\Delta x)^2}{\Delta t} \\
&= v(1-p)\Delta x + (1-p)\frac{(\Delta x)^2}{\Delta t} \\
&= vq\Delta x + q\frac{(\Delta x)^2}{\Delta t}.
\end{aligned}
$$

We therefore can write

$$q = (D - vq\Delta x)\frac{\Delta t}{(\Delta x)^2}, \tag{1.22}$$

$$p = (D + vp\Delta x)\frac{\Delta t}{(\Delta x)^2}. \tag{1.23}$$

Inserting this into (1.21) and subtracting $p(m, N)$ we get

$$
\begin{aligned}
\frac{p(m, N+1) - p(m, N)}{\Delta t} = &\frac{vp}{\Delta x}p(m-1, N) - \frac{vq}{\Delta x}p(m+1, N) \\
&+ D\frac{p(m+1, N) - 2p(m, N) + p(m-1, N)}{(\Delta x)^2} \\
&+ \left(\frac{2D}{(\Delta x)^2} - \frac{1}{\Delta t}\right)p(m, N)
\end{aligned}
$$

and from this

$$
\begin{aligned}
\frac{p(m, N+1) - p(m, N)}{\Delta t} = &-vp\frac{p(m, N) - p(m-1, N)}{\Delta x} \\
&- vq\frac{p(m+1, N) - p(m, N)}{\Delta x} \\
&+ D\frac{p(m+1, N) - 2p(m, N) + p(m-1, N)}{(\Delta x)^2} \\
&+ \left(\frac{2D}{(\Delta x)^2} - \frac{1}{\Delta t} + \frac{vp}{\Delta x} - \frac{vq}{\Delta x}\right)p(m, N).
\end{aligned}
$$

Reinserting the definitions of v and D into the last term, it is easy to show that it identically vanishes. When we now perform the continuum limit of this equation keeping v and D constant, we again arrive at (1.20). The Fickian diffusion equation therefore can be derived via a prescribed limiting procedure from the rate equation (1.21). Most important is the unfamiliar requirement $(\Delta x)^2/\Delta t = \text{const}$ which does not occur in deterministic motion and which captures the diffusion behavior, $x^2 \propto t$, of the random walker.

1.2.2 From Binomial to Poisson

Let us now turn back to analyzing the limiting behavior of the binomial distribution. The Gaussian distribution is not the only limiting distribution we can derive from it. In order to derive the Gaussian distribution, we had to require that $Np \to \infty$ for $N \to \infty$. Let us ask the question of what the limiting distribution is for

$$N \to \infty, \quad p \to 0, \quad Np = \text{const}.$$

Again we are only interested in the behavior of the distribution close to its maximum and expectation value, i.e., for $r \approx Np$; however, now $r \ll N$, and

$$\begin{aligned} p_N(r) &= \frac{N!}{r!(N-r)!} p^r q^{N-r} \\ &= \frac{N(N-1)\cdots(N-r+1)(N-r)!}{r!(N-r)!} p^r (1-p)^{N-r} \\ &\approx \frac{(Np)^r}{r!}(1-p)^N, \end{aligned}$$

where we have approximated all terms $(N-1)$ up to $(N-r)$ by N. So we arrive at

$$\lim_{\substack{N\to\infty, p\to 0 \\ Np=\text{const}}} \frac{(Np)^r}{r!}\left(1 - \frac{Np}{N}\right)^N = \frac{\langle r\rangle^r}{r!} \mathrm{e}^{-\langle r\rangle}. \tag{1.24}$$

This is the Poisson distribution which is completely characterized by its first moment. To compare the two limiting regimes for the binomial distributions which we have derived, take a look at Figs. 1.3 and 1.4.

For the first figure we have chosen $N = 1000$ and $p = 0.8$, so that $\langle r\rangle = Np = 800$. The binomial distribution and the Gaussian distribution of the same mean and width are already indistinguishable. The Poisson distribution with the same mean is much broader and not a valid approximation for the binomial distribution in this parameter range. The situation is reversed in Fig. 1.4, where again $N = 1000$ but now $p = 0.01$, so that $\langle r\rangle = Np = 10 \ll N$. Now the Poisson distribution is the better approximation to the binomial distribution, capturing especially the fact that for these parameters the distribution is not symmetric around the maximum, as a comparison with the Gaussian distribution (which is symmetric by definition)

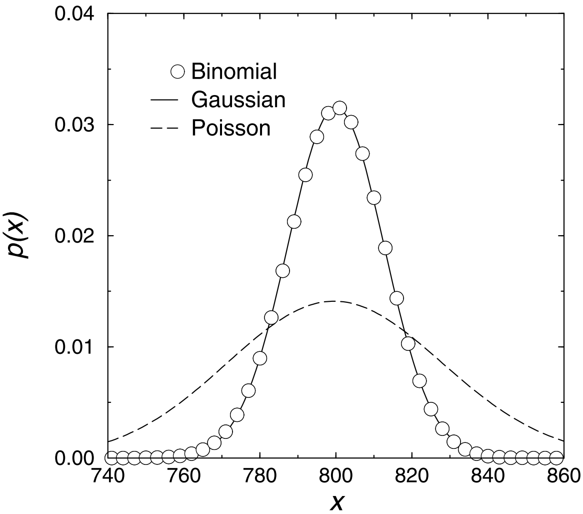

Fig. 1.3 Plot of the binomial distribution for a number of steps $N = 1000$ and the probability of a jump to the right $p = 0.8$ (*open circles*). This is compared with the Gaussian approximation with the same mean and width (*solid curve*) and the Poisson distribution with the same mean (*dashed curve*)

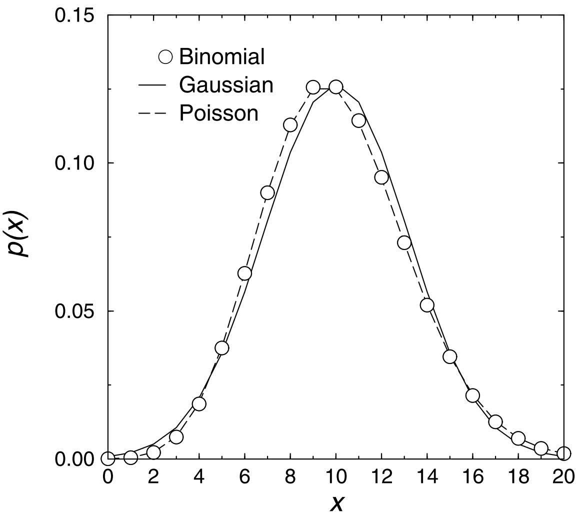

Fig. 1.4 Plot of the binomial distribution for a number of steps $N = 1000$ and the probability of a jump to the right $p = 0.01$ (*open circles*). This is compared with the Gaussian approximation with the same mean and width (*solid curve*) and the Poisson distribution with the same mean (*dashed curve*)

shows. For large r the binomial and Poisson distributions show a larger probability than the Gaussian distribution, whereas the situation is reversed for r close to zero.

To quantify such asymmetry in a distribution, one must look at higher moments. For the Poisson distribution one can easily derive the following recurrence relation between the moments,

$$\langle r^n \rangle = \lambda \left(\frac{\mathrm{d}}{\mathrm{d}\lambda} + 1 \right) \langle r^{n-1} \rangle \quad (n \geq 1), \tag{1.25}$$

where $\lambda = Np = \langle r \rangle$. So we have for the first four moments

$$\begin{aligned}
\left\langle r^0 \right\rangle &= \sum_{r=0}^{\infty} p_N(r) = 1, \\
\left\langle r^1 \right\rangle &= \lambda, \\
\left\langle r^2 \right\rangle &= \lambda(1+\lambda) \Rightarrow \sigma^2 = \lambda, \\
\left\langle r^3 \right\rangle &= \lambda^3 + 3\lambda^2 + \lambda, \\
\left\langle r^4 \right\rangle &= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda.
\end{aligned}$$

We saw that the second central moment, σ^2, is a measure of the width of the distribution and in a similar way the normalized third central moment, or *skewness*, is used to quantify the asymmetry of the distribution,

$$\hat{\kappa}_3 = \frac{\langle (r - \langle r \rangle)^3 \rangle}{\sigma^3} = \lambda^{-1/2}. \tag{1.26}$$

The normalized fourth central moment, or *kurtosis*, measures the 'fatness' of the distribution or the excess probability in the tails,

$$\hat{\kappa}_4 = \frac{\langle (r - \langle r \rangle)^4 \rangle - 3\langle (r - \langle r \rangle)^2 \rangle^2}{\sigma^4} = \lambda^{-1}. \tag{1.27}$$

Excess has to be understood in relation to the Gaussian distribution,

$$p_{\mathrm{G}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \langle x \rangle)^2}{2\sigma^2} \right], \tag{1.28}$$

for which due to symmetry all odd central moments $\langle (x - \langle x \rangle)^n \rangle$ with $n \geq 3$ are zero. Using the substitution $y = (x - \langle x \rangle)^2 / 2\sigma^2$ we find for the even central moments

$$\begin{aligned}
\left\langle \left(x - \langle x \rangle\right)^{2n} \right\rangle_{\mathrm{G}} &= \int_{-\infty}^{+\infty} \left(x - \langle x \rangle\right)^{2n} p_{\mathrm{G}}(x) \mathrm{d}x = \frac{(2\sigma^2)^n}{\sqrt{\pi}} \int_0^{\infty} y^{(n+1/2)-1} \mathrm{e}^{-y} \mathrm{d}y \\
&= \frac{(2\sigma^2)^n}{\sqrt{\pi}} \Gamma\left(n + \frac{1}{2}\right),
\end{aligned} \tag{1.29}$$

where $\Gamma(x)$ is the Gamma function [1]. From this one easily derives that $\hat{\kappa}_4$ is zero for the Gaussian distribution. In fact, for $\lambda = \langle r \rangle \to \infty$, we again recover the Gaussian distribution from the Poisson distribution.

1.2.3 Log–Normal Distribution

As a last example of important distribution functions, we want to look at the so-called log–normal distribution, which is an example of a strongly asymmetric distribution. Starting from the Gaussian distribution $p_G(x)$ and focusing for $\langle x \rangle \gg 1$ on $x > 0$, we can define

$$x = \ln y, \quad \langle x \rangle = \ln y_0, \quad \mathrm{d}x = \frac{\mathrm{d}y}{y} \tag{1.30}$$

and get the log–normal distribution

$$p_G(x)\mathrm{d}x = p_{LN}(y)\mathrm{d}y = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{[\ln(y/y_0)]^2}{2\sigma^2}\right]\frac{\mathrm{d}y}{y}. \tag{1.31}$$

The log–normal distribution is normalized in the following way,

$$\int_0^{\infty} p_{LN}(y)\mathrm{d}y = \int_{-\infty}^{+\infty} p_G(x)\mathrm{d}x = 1. \tag{1.32}$$

The moments are given by

$$\begin{aligned} \langle y^n \rangle &= \int_0^{\infty} y^n p_{LN}(y)\mathrm{d}y \\ &= \int_{-\infty}^{+\infty} \mathrm{e}^{nx} p_G(x)\mathrm{d}x \\ &= y_0^n \mathrm{e}^{n^2\sigma^2/2} \int_{-\infty}^{+\infty} p_G\left(x - n\sigma^2\right)\mathrm{d}x, \end{aligned}$$

leading to

$$\langle y^n \rangle = y_0^n \mathrm{e}^{n^2\sigma^2/2}. \tag{1.33}$$

The maximum of the distribution is at

$$y_{\max} = y_0 \mathrm{e}^{-\sigma^2}. \tag{1.34}$$

A further quantity to characterize the location of a distribution is its median, which is the point where the cumulative probability is equal to one-half,

$$F(y_{\mathrm{med}}) = \int_0^{y_{\mathrm{med}}} p_{LN}(y)\mathrm{d}y = \frac{1}{2}. \tag{1.35}$$

With $x = \ln y$ and $x_{\mathrm{med}} = \ln y_{\mathrm{med}}$, this is equivalent to

$$\int_{-\infty}^{x_{\mathrm{med}}} p_G(x)\mathrm{d}x = \frac{1}{2},$$

or, because of the symmetry of the Gaussian distribution,

$$x_{\mathrm{med}} = \langle x \rangle = \ln y_0 \Rightarrow y_{\mathrm{med}} = y_0. \tag{1.36}$$

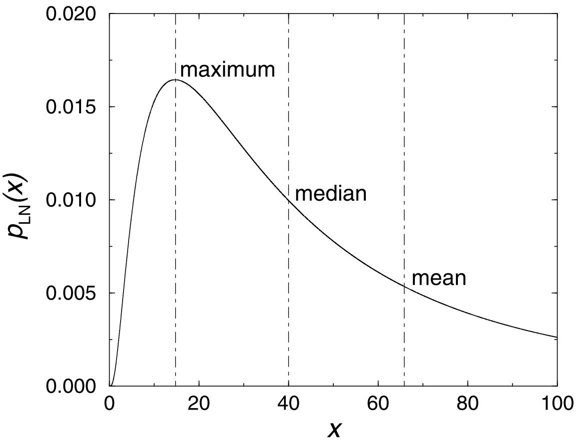

Fig. 1.5 Plot of the log–normal distribution, with *vertical dashed lines* indicating the positions of the maximum, median and mean of the distribution

We therefore find the relative order

$$y_{\max} = y_0 e^{-\sigma^2} < y_{\mathrm{med}} = y_0 < \langle y \rangle = y_0 e^{\sigma^2/2}.$$

In Fig. 1.5 we show the log–normal distribution for $y_0 = 40$ and $\sigma = 1$. This distribution shows a pronounced skewness around the maximum and a fat-tailed behavior for large argument.

1.3 Further Reading

This section was split into two parts. The first one gave a short introduction to the history of stochastic processes and the second one was devoted to the discussion of the one-dimensional random walk.

Section 1.1

- The history of probability theory and the theory of stochastic processes, as we have shortly sketched it, is nicely described in the review article by Montroll and Shlesinger [145]. Those interested in more background information may want to have a look at the books by Pearson [165] or Schneider [183] (in German).

Section 1.2

- The treatment of the random walk on the one-dimensional lattice can be found in many textbooks and review articles on stochastic processes, noticeably that of Chandrasekhar [26]. An elucidating treatment in the context of One-Step or Birth-and-Death processes can be found in the book by van Kampen [101]. Further literature for the random walk can be found at the end of Chap. 3.

- The Fickian diffusion equation is a special case of a Fokker-Planck equation, which we will discuss in the second part of Chap. 2. Its applications in the natural sciences are widespread. All transport phenomena obeying the diffusion approximation $\Delta x^2/\Delta t = \text{const.}$, can be modeled by this equation. Examples include energy transport (following a temperature gradient), mass transport (following a density gradient) and interdiffusion (following a concentration gradient).
- The Gaussian limit of probability distributions is of central importance in statistical physics and is therefore discussed in many textbooks, e.g. in Reichl [176].
- The log–normal distribution is of importance in finance [18, 217], as will be seen in Chap. 5. A detailed account of its properties as well as further applications may be found in [99].

Springer

http://www.springer.com/978-3-319-00326-9

Stochastic Processes
From Physics to Finance
Paul, W.; Baschnagel, J.
2013, XIII, 280 p. 43 illus., Hardcover
ISBN: 978-3-319-00326-9