

Chapter 6

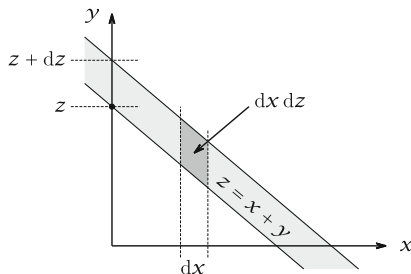
Stable Distributions and Random Walks

Abstract Stable distributions are special types of probability distributions whose origin is a particular limit regime of other types of distributions. They are closely related to the simple convolution process, which is introduced first for continuous and then for discrete random variables. This leads to the central limit theorem as one of the most important results of probability theory, as well as to its generalized version which is useful in the analysis of random walks. Extreme-value distributions are also presented, as they possess a limit theorem of their own (Fisher–Tippett–Gnedenko). The last part is devoted to the discussion of discrete-time and continuous-time random walks, together with their characteristic diffusion properties.

In this chapter we sum independent random variables X_i and discuss what happens to the distribution of their sum, $Y = \sum_i X_i$. We shall see that the distribution of Y is given by the convolution of distributions of individual X_i 's, and that in the case $i \rightarrow \infty$ —under certain conditions—the distributions of Y tend to *stable distributions*, relevant for the processes of *random walks*.

6.1 Convolution of Continuous Distributions

What is the distribution of $Z = X + Y$ if continuous random variables X and Y correspond to densities $f_X(x)$ and $f_Y(y)$? We are interested in the probability that the sum $x + y$ falls within the interval $[z, z + dz]$, where x and y are arbitrary within their own definition domains. All points fulfilling this requirement are represented by the oblique shaded area in the figure.



One must add up all contributions to the probability within this band. The infinitesimal area $dx dz$ (shaded rhomboid) corresponds to the infinitesimal probability $f_X(x)f_Y(y) dx dz$. By integrating over x we obtain the probability $f_Z(z) dz$. Let us write only its density and insert $y = z - x$:

$$f_Z(z) = (f_X * f_Y)(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x) dx. \tag{6.1}$$

This operation is called the *convolution of distributions* and we denote it by the symbol $*$. If you do not trust this geometric argument, one can also reason as follows:

$$\begin{aligned} f_Z(z) dz &= P(z \leq Z \leq z + dz) = P(z \leq X + Y \leq z + dz) \\ &= \int_{-\infty}^{\infty} dx \int_{z-x}^{z-x+dz} f_X(x)f_Y(y) dy = \int_{-\infty}^{\infty} f_X(x)dx \underbrace{\int_{z-x}^{z-x+dz} f_Y(y) dy}_{f_Y(z-x) dz}, \end{aligned}$$

whence (6.1) follows immediately. Convolution is a symmetric operation:

$$(f_X * f_Y)(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x) dx = \int_{-\infty}^{\infty} f_X(z - y)f_Y(y) dy = (f_Y * f_X)(z).$$

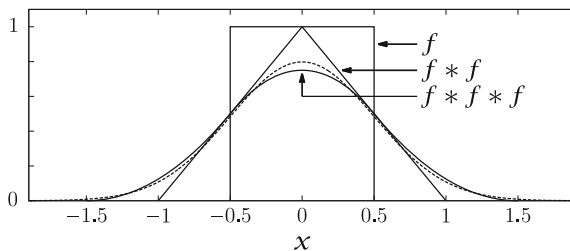
A convolution of three probability distributions is calculated as follows:

$$(f_1 * f_2 * f_3)(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(x_1)f_2(x_2)f_3(z - x_1 - x_2) dx_1 dx_2,$$

and generalizations of higher order are obvious.

Example What do we obtain after two consecutive convolutions of a symmetric uniform distribution $U(-1/2, 1/2)$, corresponding to the “box” probability density f shown in Fig. 6.1? The first convolution yields

Fig. 6.1 Twofold consecutive convolution of $U(-1/2, 1/2)$ with itself



$$g(x) = (f * f)(x) = \int_{-\infty}^{\infty} f(x')f(x - x') dx' = \begin{cases} \int_{-1/2}^{x+1/2} dx' = 1 + x; & -1 \leq x \leq 0, \\ \int_{x-1/2}^{1/2} dx' = 1 - x; & 0 \leq x \leq 1, \end{cases}$$

which is a triangular distribution ($f * f$ in the figure). The second convolution gives

$$(f * f * f)(x) = \int_{-\infty}^{\infty} f(x')g(x - x') dx'$$

$$= \begin{cases} \int_{-1/2}^{x+1} [\bullet] dx' = \frac{9}{8} + \frac{3x}{2} + \frac{x^2}{2} & ; -\frac{3}{2} \leq x \leq -\frac{1}{2}, \\ \int_{-1/2}^x [\bullet] dx' + \int_x^{1/2} [\bullet] dx' = -\frac{x^2}{2} + \frac{3}{4} & ; |x| \leq \frac{1}{2}, \\ \int_{x-1}^{1/2} [\bullet] dx' = \frac{9}{8} - \frac{3x}{2} + \frac{x^2}{2} & ; \frac{1}{2} \leq x \leq \frac{3}{2}, \end{cases}$$

where $\bullet = 1 + (x - x')$. This density is denoted by $f * f * f$ in the figure. Try to proceed yet another step and calculate $f * f * f * f$! (You shall see in an instant where this is leading.) \triangleleft

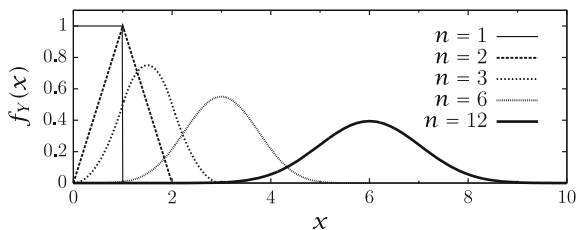
Example What about the convolution of an asymmetric distribution? For instance, what is the distribution of the variable $Y = X_1 + X_2 + \dots + X_n$ if all X_i are uniformly distributed on $[0, 1]$, i.e. $X_i \sim U(0, 1)$?

The density of the variable Y for arbitrary $n \geq 1$ is

$$f_Y(x) = \frac{1}{(n - 1)!} \sum_{h=0}^{\lfloor x \rfloor} \binom{n}{h} (-1)^h (x - h)^{n-1}, \quad n \geq 1, \tag{6.2}$$

and is shown in Fig. 6.2 for $n = 1$ (original distribution), $n = 2$ (single convolution), $n = 3$, $n = 6$ and $n = 12$. As in the previous example, the density after several convolutions reminds one of something “bell-shaped”, one could suspect, the normal distribution. Besides, the distribution of the sum variable creeps away from the origin: this is a cue for the following subsection. \triangleleft

Fig. 6.2 Multiple convolutions of the $U(0, 1)$ distribution with itself



6.1.1 The Effect of Convolution on Distribution Moments

First consider what happens to the average of the sum of two random variables:

$$\begin{aligned} \bar{Z} &= \int_{-\infty}^{\infty} z f_Z(z) dz = \int_{-\infty}^{\infty} z \left[\int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx \right] dz \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} z f_Y(z-x) dz \right] f_X(x) dx = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} (x+y) f_Y(y) dy \right] f_X(x) dx \\ &= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f_Y(y) dy \right] f_X(x) dx + \int_{-\infty}^{\infty} f_X(x) dx \int_{-\infty}^{\infty} y f_Y(y) dy = \bar{X} + \bar{Y}, \end{aligned}$$

thus

$$\overline{X+Y} = \bar{X} + \bar{Y}$$

or $E[X + Y] = E[X] + E[Y]$, which we already know from (4.6). Now let us also calculate the variance of Z ! We must average the expression

$$(Z - \bar{Z})^2 = [(X - \bar{X}) + (Y - \bar{Y})]^2 = (X - \bar{X})^2 + 2(X - \bar{X})(Y - \bar{Y}) + (Y - \bar{Y})^2.$$

Because X and Y are independent, the expected value of the second term is zero, so we are left with only

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

or $\text{var}[X + Y] = \text{var}[X] + \text{var}[Y]$. We know that too, namely, from (4.20), in a slightly different garb also from (4.25) if one sets $Y = X_1 + X_2$. As an exercise, check what happens to the third and fourth moment upon convolution: you will find out that $M_{3,X+Y} = M_{3,X} + M_{3,Y}$, so the third moments of distributions are additive. By taking into account the definition of skewness $\rho = M_3/\sigma^3$ (see (4.18)) this can also be written as

$$\rho_{X+Y} \sigma_{X+Y}^3 = \rho_X \sigma_X^3 + \rho_Y \sigma_Y^3.$$

The fourth moments are not additive, since $M_{4,X+Y} = M_{4,X} + M_{4,Y} + 6M_{2,X}M_{2,Y}$, but by using (4.19) this can be simplified to

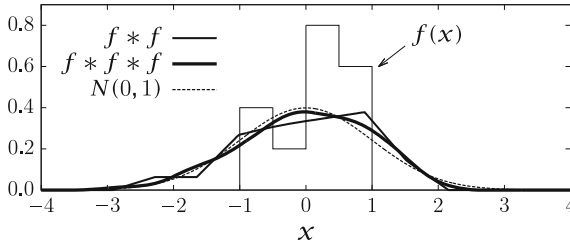


Fig. 6.3 As few as two convolutions may be needed to turn a relatively irregular distribution into a distribution that looks almost like the standardized normal. (We have subtracted the expected value of the distribution obtained at each step and rescaled the variance)

$$\varepsilon_{X+Y}\sigma_{X+Y}^4 = \varepsilon_X\sigma_X^4 + \varepsilon_Y\sigma_Y^4.$$

Example It appears as if even the most “weird” distribution evolves into something “bell-shaped” when it is convoluted with itself a couple of times. Figure 6.3 shows an example in which upon just two convolutions a rather irregular density (fulfilling all requirements for a probability density) turns into a density closely resembling the standardized normal distribution. ◀

Example Still, convolution does not perform miracles. Let us calculate the n -fold convolution of the Cauchy distribution with itself! We obtain

$$\begin{aligned} f^{(1)}(x) &= f(x) = \frac{1}{\pi} \frac{1}{1+x^2}, \\ f^{(2)}(x) &= (f * f)(x) = \frac{1}{\pi} \frac{2}{4+x^2}, \\ f^{(3)}(x) &= (f * f * f)(x) = \frac{1}{\pi} \frac{3}{9+x^2}, \\ &\vdots \\ f^{(n)}(x) &= \underbrace{(f * f * \dots * f)}_n(x) = \frac{1}{\pi} \frac{n}{n^2+x^2}. \end{aligned} \tag{6.3}$$

Certainly $f^{(n)}$ does not approach the density of the normal distribution; rather, it remains faithful to its ancestry. Consecutive convolutions yield just further Cauchy distributions! We say that the Cauchy distribution is *stable with respect to convolution*. The reasons for this behaviour will be discussed below. ◀

6.2 Convolution of Discrete Distributions

The discrete analog of the continuous convolution formula (6.1) for the summation of independent discrete random variables X and Y is at hand: if X takes the value i , then Y must be $n - i$ if their sum is to be n . Since X and Y are independent, the

probabilities for such an “event” should be multiplied, thus

$$P(X + Y = n) = \sum_i P(X = i, Y = n - i) = \sum_i P(X = i)P(Y = n - i) \quad (6.4)$$

or

$$f_{X+Y}(n) = \sum_i f_X(i)f_Y(n - i).$$

Example Let us demonstrate that the convolution of two Poisson distributions is still a Poisson distribution! Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ be mutually independent Poisson variables with parameters λ and μ . For their sum $Z = X + Y$ one then has

$$\begin{aligned} P(Z = n) &= \sum_{i=0}^n P(X = i, Y = n - i) = \sum_{i=0}^n P(X = i)P(Y = n - i) \\ &= \sum_{i=0}^n \frac{\lambda^i e^{-\lambda}}{i!} \frac{\mu^{(n-i)} e^{-\mu}}{(n-i)!} = \frac{e^{-(\lambda+\mu)}}{n!} \sum_{i=0}^n \frac{n!}{i!(n-i)!} \lambda^i \mu^{n-i} = \frac{e^{-(\lambda+\mu)} (\lambda + \mu)^n}{n!}, \end{aligned}$$

thus indeed $Z \sim \text{Poisson}(\lambda + \mu)$. A more elegant solution of this problem will be given by the Example on p. 369 in Appendix B.1. \triangleleft

Example Let us compute the probability distribution of the sum $Z = X + Y$ of independent discrete random variables X and Y , distributed according to

$$f_n = P(X=n) = \begin{cases} 0.15 ; & n = -3, \\ 0.25 ; & n = -1, \\ 0.1 ; & n = 2, \\ 0.3 ; & n = 6, \\ 0.2 ; & n = 8, \\ 0 ; & \text{otherwise,} \end{cases} \quad g_n = P(Y=n) = \begin{cases} 0.2 ; & n = -2, \\ 0.1 ; & n = 1, \\ 0.3 ; & n = 5, \\ 0.4 ; & n = 8, \\ 0 ; & \text{otherwise.} \end{cases}$$

The distributions are shown in Fig. 6.4 (left) [1].

In principle we are supposed to find all values $P(Z = z)$, so we must compute the convolution sum $\{h\} = \{f\} * \{g\}$ for each n separately:

$$h_n = P(Z = n) = \sum_{j=-\infty}^{\infty} f_j g_{n-j}.$$

To make the point, let us just calculate the probability that $X + Y = 4$. We need

$$\begin{aligned} h_4 = P(Z = 4) &= f_{-3}g_7 + f_{-2}g_6 + \underline{f_{-1}g_5} + f_0g_4 + f_1g_3 + f_2g_2 \\ &\quad + f_3g_1 + f_4g_0 + f_5g_{-1} + \underline{f_6g_{-2}} + f_7g_{-3} + f_8g_{-4} \\ &= 0.25 \cdot 0.3 + 0.3 \cdot 0.2 = 0.135. \end{aligned}$$

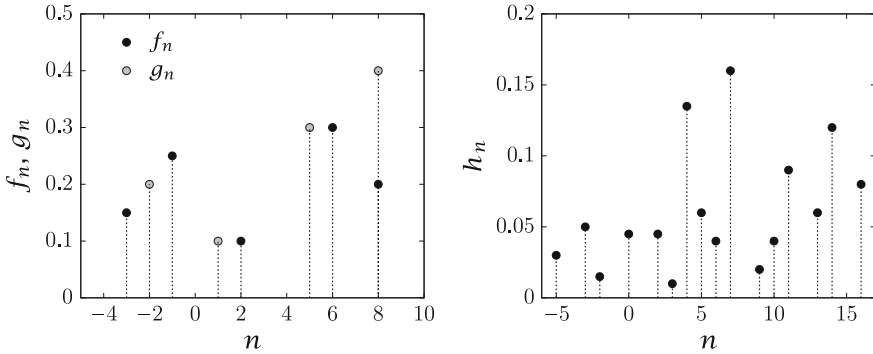


Fig. 6.4 Discrete convolution in the case when the distributions have different supports. [Left] Distributions f and g . [Right] Convolution of f and g

When n and j indices are combed through, many $f_j g_{n-j}$ terms vanish (crossed-out terms above); only the underlined bilinears survive. Such a procedure must be repeated for each n : a direct calculation of convolutions may become somewhat tedious. The problem can also be solved by using generating functions, as demonstrated by the Example on p. 374 in Appendix B.2. \triangleleft

6.3 Central Limit Theorem

Let X_1, X_2, \dots, X_n be real, independent and identically distributed random variables with probability density f_X , whose expected value $\mu_X = E[X_i]$ and variance $\sigma_X^2 = E[(X_i - \mu_X)^2]$ are bounded. Define the sum of random variables $Y_n = \sum_{i=1}^n X_i$. By (4.6) and (4.22), the expected value and variance of Y_n are $E[Y_n] = \mu_{Y_n} = n\mu_X$ and $\sigma_{Y_n}^2 = n\sigma_X^2$, respectively. The probability density f_Y of the sum variable Y_n is given by the n -fold convolution of the densities of the X_i 's,

$$f_{Y_n} = \underbrace{f_X * f_X * \dots * f_X}_n.$$

The example in Fig. 6.2 has revealed that the average of the probability density, calculated by consecutive convolutions of the original density, kept on increasing: in that case, the average in the limit $n \rightarrow \infty$ even diverges! One sees that the variance keeps on growing as well. Both problems can be avoided by defining a rescaled variable

$$Z_n = \frac{Y_n - \mu_{Y_n}}{\sigma_{Y_n}} = \frac{Y_n - n\mu_X}{\sqrt{n}\sigma_X}.$$

This ensures that upon subsequent convolutions, the average of the currently obtained density is subtracted and its variance is rescaled: see Fig. 6.3. In the limit $n \rightarrow \infty$ the

distribution function of the variable Z_n then converges to the distribution function of the standardized normal distribution $N(0, 1)$,

$$\lim_{n \rightarrow \infty} P(Z_n \leq z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt,$$

or, in the language of probability densities,

$$\lim_{n \rightarrow \infty} \sigma_{Y_n} f_{Y_n}(\sigma_{Y_n} z + \mu_{Y_n}) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

In other words, the dimensionless probability density $\sigma_{Y_n} f_{Y_n}$ converges to the standardized normal probability density in the limit $n \rightarrow \infty$, which is known as the *central limit theorem* (CLT).

6.3.1 Proof of the Central Limit Theorem

The central limit theorem can be proven in many ways: one way is to exploit our knowledge on momentum-generating functions from Appendix B.2. Suppose that the momentum-generating function of the variables X_i exists and is finite for all t in some neighborhood of $t = 0$. Then for each standardized variable $U_i = (X_i - \mu_X)/\sigma_X$, for which $E[U_i] = 0$ and $\text{var}[U_i] = 1$ (thus also $E[U_i^2] = 1$), there exists a corresponding momentum-generating function

$$M_{U_i}(t) = E[e^{tU_i}],$$

which is the same for all U_i . Its Taylor expansion in the vicinity of $t = 0$ is

$$M_U(t) = 1 + t \underbrace{E[U]}_0 + \frac{t^2}{2!} \underbrace{E[U^2]}_1 + \frac{t^3}{3!} E[U^3] + \dots = 1 + \frac{t^2}{2} + \mathcal{O}(t^2). \quad (6.5)$$

Let us introduce the standardized variable

$$Z_n = (U_1 + U_2 + \dots + U_n)/\sqrt{n} = (X_1 + X_2 + \dots + X_n - n\mu_X)/(\sigma_X \sqrt{n}).$$

Its momentum-generating function is $M_{Z_n}(t) = E[e^{tZ_n}]$. Since the variables X_i are mutually independent, this also applies to the rescaled variables U_i , therefore, by formula (B.16), we get

$$E[e^{tZ_n}] = E[e^{t(U_1+U_2+\dots+U_n)/\sqrt{n}}] = E[e^{(t/\sqrt{n})U_1}]E[e^{(t/\sqrt{n})U_2}] \dots E[e^{(t/\sqrt{n})U_n}]$$

or

$$M_{Z_n}(t) = [M_U(t/\sqrt{n})]^n, \quad n = 1, 2, \dots$$

By using the expansion of M_U , truncated at second order, we get

$$M_{Z_n}(t) = \left[1 + \frac{t^2}{2n} + \mathcal{O}(t^2/n) \right]^n, \quad n = 1, 2, \dots$$

Hence

$$\lim_{n \rightarrow \infty} M_{Z_n}(t) = \lim_{n \rightarrow \infty} \left(1 + \frac{t^2}{2n} \right)^n = e^{t^2/2}.$$

We know from (B.13) that this is precisely the momentum-generating function corresponding to the normal distribution $N(0, 1)$, so indeed

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2},$$

which we set out to prove. A direct proof (avoiding the use of generating functions) can be found in [2]; it proceeds along the same lines as the proof of the Laplace limit theorem in Appendix B.3.1.

The speed of convergence to the standardized normal distribution $N(0, 1)$ with the distribution function $\Phi(z)$ is quantified by the Berry–Esséen theorem [2]. If the third moment of $|X - \mu_X|$ is bounded ($\rho = E[|X - \mu_X|^3] < \infty$), it holds that

$$|P(Z_n \leq z) - \Phi(z)| \leq \frac{C\rho}{\sqrt{n}\sigma_X^3},$$

where $0.4097 \lesssim C \lesssim 0.4748$ [3]. Now we also realize why consecutive convolutions in (6.3) have not led us to the normal distribution: no moments exist for the Cauchy distribution (see Sect. 4.7.1), so the condition $\rho < \infty$ is not fulfilled. Moreover, one should not truncate the Taylor expansion (6.5).

The central limit theorem and the form of the bound on the speed of convergence remain valid when summing variables X_i distributed according to different (non-identical) probability distributions, if the variables are not too “dispersed” (Lindeberg criterion, see [2]). An excellent (and critical) commentary on “why normal distributions are normal” is also given by [4].

Example Let us revisit the convolution of the uniform distribution in Fig. 6.2. We sum twelve mutually independent variables $X_i \sim U(0, 1)$ and subtract 6,

$$Y = \sum_{i=1}^{12} X_i - 6. \tag{6.6}$$

What are we supposed to get? The averages of all X_i are $1/2$, $E[X_i] = 1/2$, while their variances are $\text{var}[X_i] = 1/12$ (see Table 4.1). Hence, Y should also have an average of zero and a variance of $\text{var}[Y] = \text{var}[X_1] + \dots + \text{var}[X_{12}] = 12/12 = 1$. By the central limit theorem, Y should be almost normally distributed, if we believe that $12 \approx \infty$. How well this holds is shown in Fig. 6.5.

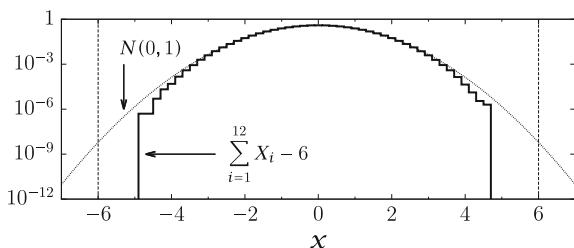


Fig. 6.5 Histogram of 10^7 values Y , randomly generated according to (6.6), compared to the density of the standardized normal distribution (3.10). In effect, the figure also shows the deviation of (6.2) from the normal density. The sharp cut-offs at ≈ -5 and ≈ 4.7 are random: by drawing a larger number of values the histogram would fill the whole interval $[-6, 6]$

We have thus created a primitive “convolution” generator of approximately normally distributed numbers, but with its tails cut off since Y can never exceed 6 and can never drop below -6 . It is a practical generator—which does not mean that it is good. How a “decent” generator of normally distributed random numbers can be devised will be discussed in Sect. C.2.5. \triangleleft

Example (Adapted from [5].) The mass M of granules of a pharmaceutical ingredient is a random variable, distributed according to the probability density

$$f_M(m) = \frac{1}{24m_0^5} m^4 e^{-m/m_0}, \quad m \geq 0, \quad m_0 = 40 \text{ mg}. \quad (6.7)$$

To analyze the granulate, we acquire a sample of 30 granules. What is the probability that the total mass of the granules in the sample exceeds its average value by more than 10%?

The average mass of a single granule and its variance are

$$\bar{M} = \int_0^\infty m f_M(m) dm = 5m_0, \quad \sigma_M^2 = \int_0^\infty (m - \bar{M})^2 f_M(m) dm = 5m_0^2.$$

The probability density f_X of the total sample mass X , which is also a random variable, is a convolution of thirty densities of the form (6.7); this number is large enough to invoke the central limit theorem, so the density f_X is almost normal, with average $\bar{X} = 30\bar{M} = 150m_0$ and variance $\sigma_X^2 = 30\sigma_M^2 = 150m_0^2$:

$$f_X(x) \approx f_{\text{norm}}(x; \bar{X}, \sigma_X^2) = f_{\text{norm}}(x; 150m_0, 150m_0^2).$$

The desired probability is then

$$P(X > 165m_0) \approx \int_{165m_0}^\infty f_{\text{norm}}(x; \bar{X}, \sigma_X^2) dx = \frac{1}{2} \left[1 - \text{erf} \left(\frac{(165 - 150)m_0}{\sqrt{2}\sqrt{150}m_0} \right) \right] \approx 11\%,$$

where we have used Table D.2. \triangleleft

6.4 Stable Distributions ★

The normal distribution as the limit distribution of the sum of independent random variables can be generalized by the concept of *stable distributions* [6, 7].

Suppose we are dealing with independent random variables X_1, X_2 and X_3 with the same distribution over the sample space Ω . We say that such a distribution is *stable*, if for each pair of numbers a and b there exists a pair c and d such that the distribution of the linear combination $aX_1 + bX_2$ is equal to the distribution of $cX_3 + d$, that is,

$$P(aX_1 + bX_2 \in A) = P(cX_3 + d \in A) \quad \forall A \subset \Omega.$$

Such random variables are also called ‘stable’; a superposition of stable random variables is a linear function of a stable random variable with the same distribution.

Stable distributions are most commonly described by their characteristic functions (see Appendix B.3). Among many possible notations we follow [6]. We say that a random variable X has a stable distribution $f_{\text{stab}}(x; \alpha, \beta, \gamma, \delta)$, if the logarithm of its characteristic function (B.17) has the form

$$\log \phi_X(t) = i\delta t - \gamma^\alpha |t|^\alpha [1 - i\beta \Phi_\alpha(t)],$$

where

$$\Phi_\alpha(t) = \begin{cases} \text{sign}(t) \tan(\pi\alpha/2) & ; \alpha \neq 1, \\ -\frac{2}{\pi} \text{sign}(t) \log |t| & ; \alpha = 1. \end{cases}$$

The parameter $\alpha \in (0, 2]$ is the *stability index* or *characteristic exponent*, the parameter $\beta \in [-1, 1]$ describes the skewness of the distribution, and two further parameters $\gamma > 0$ and $\delta \in \mathbb{R}$ correspond to the distribution scale and location, respectively. For $\alpha \in (1, 2]$ the expected value exists and is equal to $E[X] = \delta$. For general $\alpha \in (0, 2]$ there exist moments $E[|X|^p]$, where $p \in [0, \alpha)$.

It is convenient to express X by another random variable Z ,

$$X = \begin{cases} \gamma Z + \delta & ; \alpha \neq 1, \\ \gamma \left(Z + \frac{2}{\pi} \beta \log \gamma \right) + \delta & ; \alpha = 1. \end{cases}$$

Namely, the characteristic function of Z is somewhat simpler,

$$\log \phi_Z(t) = -|t|^\alpha [1 - i\beta \Phi_\alpha(t)],$$

as it depends only on two parameters, α and β . The probability density f_Z of the variable Z is calculated by the inverse Fourier transformation of the characteristic function ϕ_Z :

$$f_Z(z; \alpha, \beta) = \frac{1}{\pi} \int_0^\infty \exp(-t^\alpha) \cos(zt - t^\alpha \beta \Phi_\alpha(t)) dt,$$

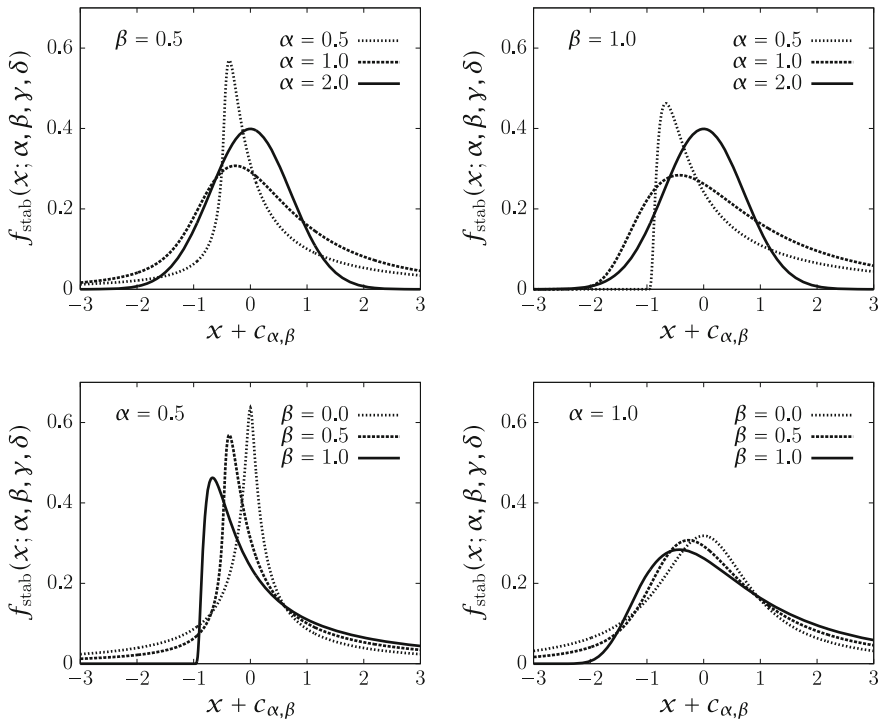


Fig. 6.6 Stable distributions $f_{\text{stab}}(x; \alpha, \beta, \gamma, \delta)$. [Top left and right] Dependence on parameter α at $\beta = 0.5$ and 1.0 . [Bottom left and right] Dependence on parameter β at $\alpha = 0.5$ and 1.0 . At $\alpha \neq 1$ the independent variable is shifted by $c_{\alpha, \beta} = \beta \tan(\pi\alpha/2)$

where $f_Z(-z; \alpha, \beta) = f_Z(z; \alpha, -\beta)$. The values of f_Z and f_X can be computed by using integrators tailored to rapidly oscillating integrands: see [8], p. 660; a modest software support for stable distributions can also be found in [9]. With respect to α and β , the definition domains of f_Z are

$$z \in \begin{cases} (-\infty, 0] & ; \alpha < 1, \beta = -1, \\ [0, \infty) & ; \alpha < 1, \beta = 1, \\ \mathbb{R} & ; \text{otherwise.} \end{cases}$$

The dependence of f_{stab} (f_X or f_Z with appropriate scaling) on the parameter α is shown in Fig. 6.6 (top left and right), while the dependence on β is shown in the same figure at bottom left and right.

By a suitable choice of parameters such a general formulation allows for all possible stable distributions. The most relevant ones are

$$\begin{aligned}
 \text{normal : } \alpha = 2, \beta = 0, f_X(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, & x \in \mathbb{R}; \\
 \text{Cauchy : } \alpha = 1, \beta = 0, f_X(x) &= \frac{1}{\pi} \frac{1}{1+x^2}, & x \in \mathbb{R}; \\
 \text{Lévy : } \alpha = \frac{1}{2}, \beta = 1, f_X(x) &= \frac{1}{\sqrt{2\pi}} e^{-1/(2x)} x^{-3/2}, & x \in \mathbb{R}_+.
 \end{aligned}$$

Stable distributions with $\alpha \in (0, 2)$ have a characteristic behaviour of probability densities known as *power* or *fat tails*. The cumulative probabilities satisfy the asymptotic relations

$$\begin{aligned}
 \beta \in (-1, 1] : \int_x^\infty f_Z(z; \alpha, \beta) dz &\sim \frac{1}{2} c_\alpha (1 + \beta) x^{-\alpha}, & x \rightarrow \infty, \\
 \beta \in [-1, 1) : \int_{-\infty}^x f_Z(z; \alpha, \beta) dz &\sim \frac{1}{2} c_\alpha (1 - \beta) (-x)^{-\alpha}, & x \rightarrow -\infty,
 \end{aligned} \tag{6.8}$$

where $c_\alpha = 2 \sin(\pi\alpha/2)\Gamma(\alpha)/\pi$. For $\beta \in (-1, 1)$ such asymptotic behaviour is valid in both limits, $x \rightarrow \pm\infty$. Note that the probability density has the asymptotics $\mathcal{O}(|x|^{-\alpha-1})$ if the cumulative probability goes as $\mathcal{O}(|x|^{-\alpha})$.

6.5 Generalized Central Limit Theorem ★

Having introduced stable distributions (Sect. 6.4) one can formulate the generalized central (or Lévy’s) limit theorem, elaborated more closely in [2]. Here we just convey its essence.

Suppose we have a sequence of independent, identically distributed random variables $\{X_i\}_{i \in \mathbb{N}}$, from which we form the partial sum

$$Y_n = X_1 + X_2 + \dots + X_n.$$

Assume that their distribution has power tails, so that for $\alpha \in (0, 2]$ the following limits exist:

$$\lim_{x \rightarrow \infty} |x|^\alpha P(X > x) = d_+, \quad \lim_{x \rightarrow -\infty} |x|^\alpha P(X < x) = d_-,$$

and $d = d_+ + d_- > 0$. Then real coefficients $a_n > 0$ and b_n exist such that the rescaled partial sum

$$Z_n = \frac{Y_n - nb_n}{a_n}$$

in the limit $n \rightarrow \infty$ is stable, and its probability density is $f_{\text{stab}}(x; \alpha, \beta, 1, 0)$. Its skewness is given by $\beta = (d_+ - d_-)/(d_+ + d_-)$, while a_n and b_n are

$$a_n = \begin{cases} (dn/c_\alpha)^{1/\alpha} & ; \alpha \in (0, 2), \\ \sqrt{(dn \log n)/2} & ; \alpha = 2, \end{cases}$$

$$b_n = \begin{cases} E[X_i] & ; \alpha \in (1, 2], \\ E[X_i H(|X_i| - a_n)] & ; \text{otherwise,} \end{cases}$$

where H is the Heaviside function. The constant c_α is defined next to (6.8). The coefficient a_n for $\alpha < 2$ diverges with increasing n as $\mathcal{O}(n^{1/\alpha})$.

The generalized central limit theorem is useful in analyzing the process of random walk, which is analogous to extending the partial sum of random numbers Y_n . Such processes are discussed in Sects. 6.7 and 6.8. The convergence to the stable distribution when $n \rightarrow \infty$ is becoming more and more “capricious” when α decreases.

6.6 Extreme-Value Distributions ★

In Sects. 6.3 and 6.4 we have discussed the distributions of values obtained in summing independent, identically distributed random variables $\{X_i\}_{i=1}^n$. Now we are interested in statistical properties of their *maximal* and *minimal* values, i.e. the behaviour of the quantities

$$M_n = \max\{X_1, X_2, \dots, X_n\},$$

$$\tilde{M}_n = \min\{X_1, X_2, \dots, X_n\},$$

when $n \rightarrow \infty$. We thereby learn something about the probability of extreme events, as exceptionally strong earthquakes, unusual extent of floods or inconceivably large amounts of precipitation: “*It rained for four years, eleven months, and two days.*” (See [10], p. 315.) The variables X_i are the values of the process, usually recorded at constant time intervals—for example, $n = 365$ daily temperature averages on Mt. Blanc—while M_n is the corresponding annual maximum. We are interested in, say, the probability that on top of Mt. Blanc, the temperature of $+42^\circ\text{C}$ will be exceeded on any one day in the next ten years.

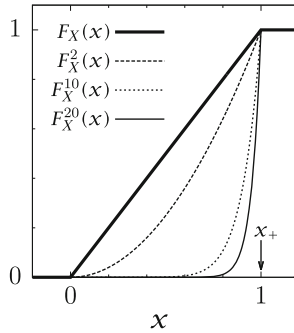
In principle, we have already answered these questions—about both the maximal and minimal value—in Problem 2.11.6: if F_X is the distribution function of the individual X_i 's, the maximal values M_n are distributed according to

$$F_{M_n}(x) = P(M_n \leq x) = [F_X(x)]^n, \quad (6.9)$$

and the minimal as

$$1 - F_{\tilde{M}_n}(x) = 1 - P(\tilde{M}_n \leq x) = P(\tilde{M}_n > x) = [1 - F_X(x)]^n.$$

But this does not help much, as F_X is usually not known! A statistical analysis of the observations may result in an approximate functional form of F_X , but even small errors in F_X (particularly in its tails) may imply large deviations in F_X^n . We therefore accept the fact that F_X is unknown and try to find families of functions F_X^n , by which extreme data can be modeled directly [11, 12].



There is another problem. Define x_+ as the smallest value x , for which $F_X(x) = 1$. Then for any $x < x_+$ we get $F_X^n(x) \rightarrow 0$, when $n \rightarrow \infty$, so that the distribution function of M_n degenerates into a “step” at x_+ . The figure above shows this in the case of uniformly distributed variables $X_i \sim U(0, 1)$ with probability density $f_X(x) = 1$ ($0 \leq x \leq 1$) and distribution function $F_X(x) = x$ ($0 \leq x \leq x_+ = 1$). When $n \rightarrow \infty$, the distribution function F_X^n tends to the step (Heaviside) function at $x = 1$, while its derivative (probability density) resembles the delta “function” at the same point. Our goal is to find a non-degenerate distribution function. We will show that this can be accomplished by a rescaling of the variable M_n ,

$$M_n^* = \frac{M_n - b_n}{a_n}, \tag{6.10}$$

where $a_n > 0$ and b_n are constants. Illustrations of a suitable choice of these constants or of their calculation are given by the following Example and Exercise in Sect. 6.9.5. A general method to determine these constants is discussed in [13, 14].

Example Let X_1, X_2, \dots, X_n be a sequence of independent, exponentially distributed variables, thus $F_X(x) = 1 - e^{-x}$. Let $a_n = 1$ and $b_n = \log n$. Then

$$\begin{aligned} P\left(\frac{M_n - b_n}{a_n} \leq x\right) &= P(M_n \leq a_n x + b_n) = P(M_n \leq x + \log n) \\ &= [F_X(x + \log n)]^n = [1 - e^{-(x + \log n)}]^n = \left[1 - \frac{1}{n} e^{-x}\right]^n \\ &\rightarrow \exp(-\exp(-x)), \quad x \in \mathbb{R}, \end{aligned}$$

when $n \rightarrow \infty$. By a suitable choice of a_n and b_n we have therefore stabilized the location and scale of the distributions of M_n^* in the limit $n \rightarrow \infty$.

Let us repeat this calculation for independent variables with the distribution function $F_X(x) = e^{-1/x}$ and for uniformly distributed variables, $F_X(x) = x$! In the first case we set $a_n = n$ and $b_n = 0$, and get $P(M_n^* \leq x) = e^{-1/x}$ ($x > 0$). In the second case a good choice is $a_n = 1/n$ and $b_n = 1$, yielding $P(M_n^* \leq x) \rightarrow e^x$ ($x < 0$) in the limit $n \rightarrow \infty$. Plot all three functions $F_X(x)$ of this Example and elaborate why one or the other are more or less sensible for the actual physical description of extreme phenomena! \triangleleft

6.6.1 Fisher–Tippett–Gnedenko Theorem

Apparently the choice of constants a_n and b_n is crucial if we wish the distribution of M_n^* in the limit $n \rightarrow \infty$ to be non-trivial (not degenerated into a point); the basic formalism for a correct determination of these constants is discussed e.g. in [14]. In the following we assume that such constants can be found; one can then namely invoke the Fisher–Tippett–Gnedenko theorem [15, 16], which is the extreme-value analog of the central limit theorem of Sect. 6.3: if there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that in the limit $n \rightarrow \infty$ we have

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \rightarrow G(x)$$

for a non-degenerate distribution function G , then G belongs to the family

$$G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \quad (6.11)$$

defined on the set of points $\{x : 1 + \xi(x - \mu)/\sigma > 0\}$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \xi < \infty$. Formula (6.11) defines the family of *generalized extreme-value distributions* (GEV). An individual distribution is described by the location parameter μ (a sort of average of extreme values), the scale parameter σ (their dispersion), and the shape parameter ξ . The value of ξ characterizes three sub-families of the GEV set—Fréchet ($\xi > 0$), Gumbel ($\xi = 0$) and Weibull ($\xi < 0$)—differing by the location of the point x_+ and asymptotics. The Gumbel-type distributions must be understood as the $\xi \rightarrow 0$ limit of (6.11):

$$G(x) = \exp\left\{-\exp\left[-\left(\frac{x - \mu}{\sigma}\right)\right]\right\}, \quad -\infty < x < \infty. \quad (6.12)$$

The corresponding probability density in the case $\xi \neq 0$ is

$$g(x) = G'(x) = \frac{1}{\sigma} [t(x)]^{1+\xi} e^{-t(x)}, \quad t(x) = \left[1 + \xi \frac{x - \mu}{\sigma}\right]^{-1/\xi} \quad (6.13)$$

while for $\xi = 0$ it is

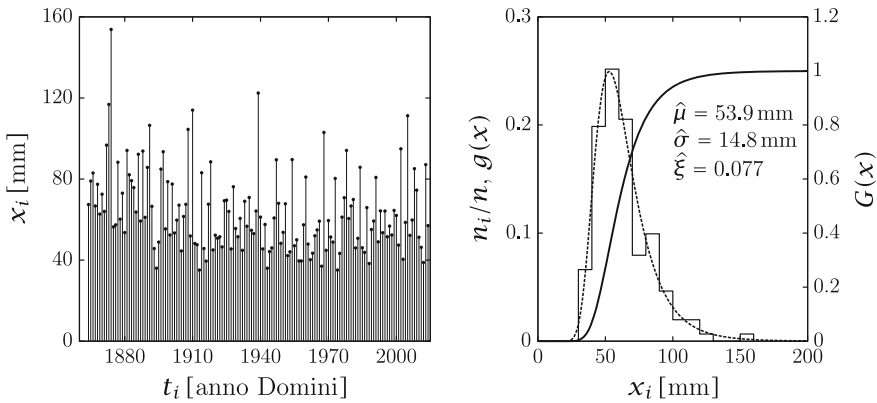


Fig. 6.7 Rainfall in Engelberg (1864–2014). [Left] Time series of annual extreme values. [Right] Histogram of extremes, the corresponding probability density g (dashed curve) and the GEV distribution function (full curve). The optimal parameters $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\xi}$ have been determined by fitting g to the histogram

$$g(x) = \frac{1}{\sigma} \exp \left[-\frac{x - \mu}{\sigma} - \exp \left(-\frac{x - \mu}{\sigma} \right) \right].$$

The predictive power of the Fisher–Tippett–Gnedenko theorem does not lag behind the one of the central limit theorem: if one is able to find suitable $\{a_n\}$ and $\{b_n\}$, the limiting extreme-value distribution is always of the type (6.11), *regardless of the parent distribution F_X that generated these extreme values in the first place!* Different choices of $\{a_n\}$ and $\{b_n\}$ lead to GEV-type distributions with different μ and σ , but with the same shape parameter ξ , which is the essential parameter of the distribution.

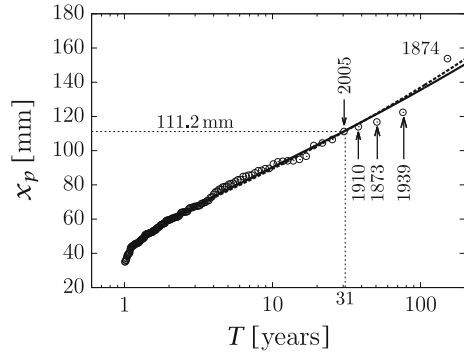
Example Figure 6.7 (left) shows the annual rainfall maxima, measured over 151 years (1864–2014) in the Swiss town of Engelberg [17]. Each data point represents the extreme one-day total (the wettest day in the year): we are therefore already looking at the extreme values and we are interested in *their* distribution, not the distribution of *all* non-zero daily quantities: *that* is most likely normal!

Figure 6.7 (right) shows the histogram of 151 extreme one-day totals, normalized such that the sum over all bins is equal to one. It can therefore be directly fitted by the density (6.13) (dashed curve), resulting in the distribution parameters $\hat{\mu} = 53.9$ mm, $\hat{\sigma} = 14.8$ mm, $\hat{\xi} = 0.077$ (Fréchet family). The corresponding distribution function is shown by the full curve. ◀

6.6.2 Return Values and Return Periods

The extreme-value distribution and its asymptotic behaviour can be nicely illustrated by a *return-level plot*. Suppose that we have measured $n = 365$ daily rainfall amounts

Fig. 6.8 Return values for extreme rainfall in Engelberg (period 1864–2014). The *full curve* is the model prediction with parameters from Fig. 6.7, and the *dashed curve* is the model with parameters obtained by the maximum likelihood method



x_i over a period of N consecutive years, so that their annual maxima are also available:

$$\underbrace{x_1, x_2, \dots, x_n}_{M_{n,1}}, \underbrace{x_{n+1}, x_{n+2}, \dots, x_{2n}}_{M_{n,2}}, \dots, \underbrace{x_{(N-1)n+1}, x_{(N-1)n+2}, \dots, x_{Nn}}_{M_{n,N}}.$$

The quantiles of the annual extremes distribution are obtained by inverting (6.11):

$$x_p = \begin{cases} \mu - \frac{\sigma}{\xi} \left[1 - (-\log(1-p))^{-\xi} \right]; & \xi \neq 0, \\ \mu - \sigma \log(-\log(1-p)) & ; \xi = 0, \end{cases}$$

where $G(x_p) = 1 - p$. We call x_p the *return level* corresponding to the *return period* $T = 1/p$. One may namely expect that the value x_p will be exceeded once every $1/p$ years or that the annual maximum will exceed the value x_p in any year with a probability of $p = 1/T$. From these definitions it follows that

$$T = \frac{1}{p} = \frac{1}{1 - G(x_p)}. \quad (6.14)$$

The model dependence of x_p on T in the case of Engelberg rainfall is shown in Fig. 6.8 by the full curve. On the abscissa one usually uses a logarithmic scale; one thereby shrinks the region of “extreme extreme” values and obtains a clearer picture of the asymptotics in terms of ξ . We must also plot the actually measured extreme observations $M_{n,1}, M_{n,2}, \dots, M_{n,N}$. In general, these are not sorted, so—in the spirit of (6.14)—individual extremes $M_{n,i}$ are mapped to their return periods:

$$T_i = \frac{N}{N + 1 - \text{rank}(M_{n,i})}, \quad i = 1, 2, \dots, N.$$

The points $(T_i, M_{n,i})$ are denoted by circles in the figure. The maximum one-day total of 111.2 mm, recorded in 2005, has an expected return period of 31 years, while the deluge witnessed in 1874 may reoccur every ≈ 150 years on the average.

The fitting of the probability density to the data as in the previous Example depends on the number of bins in the histogram (see Sect. 9.2), so this is not the best way to pin down the optimal parameters. In Problem 8.8.3 the parameters of the GEV distribution and their uncertainties will be determined for the same data set by the method of maximum likelihood. In Fig. 6.7 (right) this distribution is shown by the dashed line.

6.6.3 Asymptotics of Minimal Values

So far we have only discussed the distributions of maximal values, most frequently occurring in practice. On the other hand, the distributions of extremely *small* values, i.e. the asymptotic behaviour of the quantities

$$\tilde{M}_n = \min\{X_1, X_2, \dots, X_n\}$$

when $n \rightarrow \infty$, are also important, in particular in modeling critical errors in systems, where the lifetime of the whole system, \tilde{M}_n , is equal to the minimal lifetime of one of its components $\{X_i\}$.

There is no need to derive new formulas for minimal values; we can simply use the maximal-value results. Define $Y_i = -X_i$ for $i = 1, 2, \dots, n$, so that small values of X_i correspond to large values of Y_i . Thus if $\tilde{M}_n = \min\{X_1, X_2, \dots, X_n\}$ and $M_n = \max\{Y_1, Y_2, \dots, Y_n\}$, we also have

$$\tilde{M}_n = -M_n.$$

In the limit $n \rightarrow \infty$ we therefore obtain

$$\begin{aligned} P(\tilde{M}_n \leq x) &= P(-M_n \leq x) = P(M_n \geq -x) = 1 - P(M_n \leq -x) \\ &\rightarrow 1 - \exp \left\{ - \left[1 + \xi \left(\frac{-x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \\ &= 1 - \exp \left\{ - \left[1 - \xi \left(\frac{x - \tilde{\mu}}{\sigma} \right) \right]^{-1/\xi} \right\} \end{aligned} \quad (6.15)$$

on $\{x : 1 - \xi(x - \tilde{\mu})/\sigma > 0\}$, where $\tilde{\mu} = -\mu$. This means that a minimal-value distribution can be modeled either by directly fitting (6.15) to the observations or by using (6.11) and considering the symmetry exposed above: if, for example, we wish to model the data x_1, x_2, \dots, x_n by a minimal-value distribution (parameters $\tilde{\mu}, \sigma, \xi$), this is equivalent to modeling the data $-x_1, -x_2, \dots, -x_n$ by a maximal-value distribution with the same σ and ξ , but with $\tilde{\mu} = -\mu$.

6.7 Discrete-Time Random Walks ★

Random walks are non-stationary *random processes* used to model a variety of physical processes. A random or *stochastic* process is a generalization of the concept of a random variable: instead of drawing a single value, one “draws” a whole time series (signal), representing one possible *realization* of the random process or its *sample path*. The non-stationarity of the process means that its statistical properties change with time. (A detailed classification of random process can be found in [8].) In this subsection we discuss discrete-time random walks [2, 18, 19], while the next subsection is devoted to their continuous-time counterparts [18–21]. See also Chap. 12.

Imagine a discrete-time random process X , observed as a sequence of random variables $\{X(t)\}_{t \in \mathbb{N}}$. The partial sums of this sequence are

$$Y(t) = Y(0) + \sum_{i=1}^t X(i) = Y(t-1) + X(t) \quad (6.16)$$

and represent a new discrete-time random process Y , i.e. a sequence of random variables $\{Y(t)\}_{t \in \mathbb{N}_0}$. The process Y is called a *random walk*, whose individual step is the process $X(t)$. Let the sample space Ω of X and Y be continuous. We are interested in the time evolution of the probability density $f_{Y(t)}$ of the random variable Y if the initial density $f_{Y(0)}$ is known.

If we assume that Y is a process in which the state of each point depends only on the state of the previous point, the time evolution of $f_{Y(t)}$ is determined by

$$f_{Y(t)}(y) = \int_{\Omega} f(Y(t) = y | Y(t-1) = x) f_{Y(t-1)}(x) dx,$$

where $f(Y(t) = y | Y(t-1) = x)$ is the conditional probability density that Y goes from value x at time $t-1$ to value y at time t . Let us also assume that the process X is independent of the previous states, so that $f(X(t) = x | Y(t-1) = y - x) = f_{X(t)}(x)$. By considering (6.16) and substituting $z = y - x$ we get

$$f_{Y(t)}(y) = \int_{\Omega} f_{X(t)}(z) f_{Y(t-1)}(y - z) dz = (f_{X(t)} * f_{Y(t-1)})(y).$$

By using this formula $f_{Y(t)}$ can be expressed as a convolution of the initial distribution $f_{Y(0)}$ and the distribution of the sum of steps until time t , f_X^{*t} :

$$f_{Y(t)} = f_{Y(0)} * f_X^{*t}, \quad f_X^{*t} = f_{X(1)} * f_{X(2)} * \cdots * f_{X(t)}.$$

The time evolution $f_{Y(t)}(y)$ is most easily realized in Fourier space, where it is given by the product of Fourier transforms \mathcal{F} of the probability densities,

$$\mathcal{F}[f_{Y(t)}] = \mathcal{F}[f_{Y(0)}] \prod_{i=1}^t \mathcal{F}[f_{X(i)}].$$

One often assumes that at time zero the value of the process Y is zero and that $f_{Y(0)}(y) = \delta(y)$. This assumption is useful in particular when one is interested in the qualitative behaviour of $f_{Y(t)}$ at long times.

6.7.1 Asymptotics

To understand the time asymptotics of the distribution $f_{Y(t)}$ it is sufficient to study one-dimensional random walks. Assume that the steps are identically distributed, with the density $f_{X(t)} = f_X$, and $Y(0) = 0$. The distribution corresponding to the process Y is therefore determined by the formula

$$f_{Y(t)} = \mathcal{F}^{-1} \left[\left(\mathcal{F}[f_X] \right)^t \right]$$

for all times t . The behaviour of $f_{Y(t)}$ in the limit $t \rightarrow \infty$ is determined by the central limit theorem (Sect. 6.3) and its generalization (Sect. 6.5). The theorems tell us that at large t , $f_{Y(t)}$ converges to the *limiting* (or *asymptotic*) distribution which can be expressed by one of the stable distributions f_{stab} , such that

$$f_{Y(t)}(y) \sim L(t) f_{\text{stab}}(L(t)y + t\mu(t))$$

with suitably chosen functions L and μ . The function L represents the effective width of the central part of the distribution $f_{Y(t)}$, where the bulk of the probability is concentrated, and is called the *characteristic spatial scale* of the distribution. The function μ has the role of the distribution average.

Furthermore, if the distribution of steps, f_X , has a bounded variance, $\sigma_X^2 < \infty$, the central limit theorem tells us that $f_{Y(t)}$ tends to the normal distribution with a standard deviation of

$$L = \sigma_{Y(t)} \sim t^{1/2}.$$

Such asymptotic dependence of the spatial scale on time defines *normal diffusion*, and this regime of random walks is named accordingly (Fig. 6.9 (left)).

If the probability density f_X asymptotically behaves as

$$f_X(x) \sim \frac{C_{\pm}}{|x|^{\alpha+1}}, \quad x \rightarrow \pm\infty,$$

where C_{\pm} are constants, we say that the distribution has *power* or *fat tails*, a concept familiar from Sect. 6.4. For $\alpha \in (0, 2)$, the second moment of the distribution no longer exists, and $f_{Y(t)}$ at large t tends to a distribution with scale

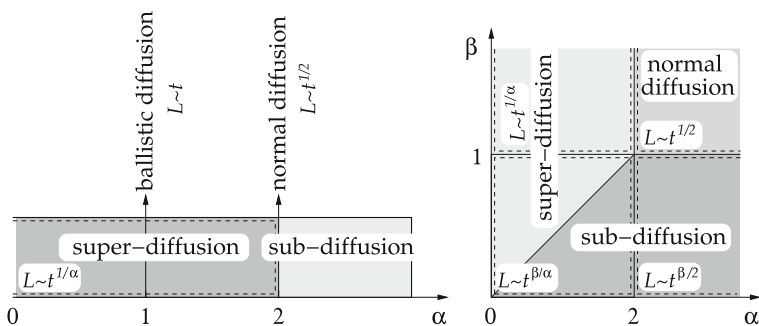


Fig. 6.9 Dependence of the characteristic spatial scale L on time t . [Left] Discrete-time random walks. [Right] Continuous-time random walks

$$L \sim t^{1/\alpha}.$$

Because in this case the characteristic scale changes faster than in the case of normal diffusion, we are referring to *super-diffusion*. The dynamics of the process Y in this regime is known as *Lévy flights*. The diffusion with $\alpha = 1$ is called *ballistic*: particles propagate with no restrictions with given velocities, so that

$$L \sim t.$$

Near $\alpha = 2$ we have $L(t) \sim (n \log n)^{1/2}$, a regime called *log-normal diffusion*.

These properties can be easily generalized to multiple space dimensions. We observe the projection of the walk, $\hat{n}^T Y(t)$, along the direction defined by the unit vector \hat{n} , and its probability density $f_{\hat{n}^T Y(t)}$. For each \hat{n} we apply the central limit theorem or its generalization and determine the scale $L_{\hat{n}}$. A random walk possesses a particular direction \hat{n}^* along which the scale is largest or increases fastest with time. We may take $L_{\hat{n}^*}$ to be the characteristic scale of the distribution $f_{Y(t)}$. An example of a simulation of a two-dimensional random walk where the steps in x and y directions are independent, is shown in Fig. 6.10.

If the densities $f_{X(t)}$ have power tails, $f_{Y(t)}$ also has them. This applies regardless of the central limit theorem or its generalization. Suppose that in the limit $t \rightarrow \infty$ we have $f_{X(t)}(x) \sim C_{\pm,t}|x|^{-\alpha-1}$. When the walk “generates” the density $f_{Y(t)}$, the tails add up, so $f_{Y(t)}(x) \sim (\sum_{i=1}^t C_{\pm,i})|x|^{-\alpha-1}$ when $x \rightarrow \pm\infty$. This means that the probability of extreme events in $Y(t)$ increases with time, since

$$P(|Y(t)| > y) \sim \sum_{i=1}^t P(|X(i)| > y), \quad \text{when } y \rightarrow \infty.$$

To estimate the variance of such processes we therefore apply methods of robust statistics (Sect. 7.4). Instead of calculating the standard deviation $\sigma_{Y(t)}$ in sub-diffusive random walks, for example, one is better off using MAD (7.23).

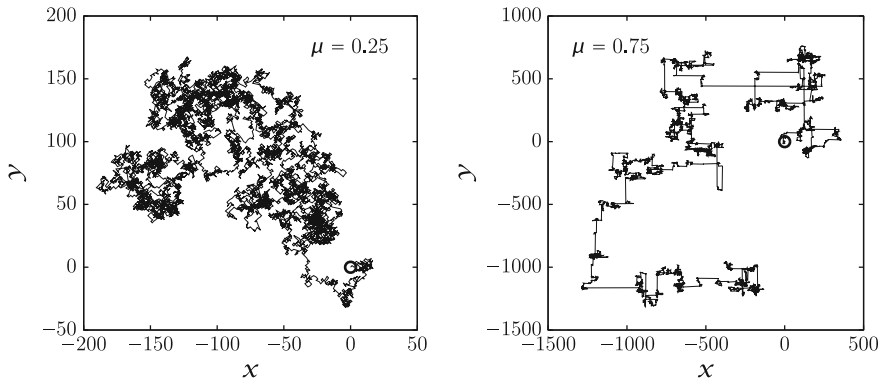


Fig. 6.10 Examples of random walks (x_t, y_t) with 10^4 steps, generated according to $x_{t+1} = x_t + \text{sign}(X)|X|^{-\mu}$ and $y_{t+1} = y_t + \text{sign}(Y)|Y|^{-\mu}$, where X and Y are independent random variables, uniformly distributed on $[-1, 1]$. [Left] $\mu = 0.25$. [Right] $\mu = 0.75$. The circles denote the initial position of the walks, $x = y = 0$

6.8 Continuous-Time Random Walks ★

In continuous-time random walks [18–21] the number of steps $N(t)$ taken until time t becomes a continuous random variable. The definition of a discrete-time random walk (6.16) should therefore be rewritten as

$$Y(t) = Y(0) + \sum_{i=1}^{N(t)} X(i).$$

The expression for $Y(t)$ can not be cast in iterative form $Y(t) = Y(t - 1) + \dots$ as in (6.16). The number of steps $N(t)$ has a probability distribution $F_{N(t)}$. Suppose that $N(t)$ and $X(i)$ are independent processes—which is not always true, as it is not possible to take arbitrary many steps within given time [22, 23]. If $X(i)$ at different times are independent and correspond to probability densities $f_{X(i)}$, the probability density of the random variable $Y(t)$ is

$$f_{Y(t)}(y) = \sum_{n=0}^{\infty} F_{N(t)}(n) (f_{Y(0)} * f_X^{*n})(y),$$

where

$$f_X^{*t} = f_{X(1)} * f_{X(2)} * \dots * f_{X(t)}.$$

In the interpretation of such random walks and the choice of distribution $F_{N(t)}$ we follow [20]. A walk is envisioned as a sequence of steps whose lengths $X(i)$ and waiting time $T(i)$ between the steps are randomly drawn. After N steps the walk makes it to the point $\mathcal{X}(N)$ and the elapsed time is $\mathcal{T}(N)$, so that

$$\mathcal{X}(N) = \sum_{i=1}^N X(i), \quad \mathcal{T}(N) = \sum_{i=1}^N T(i), \quad \mathcal{X}(0) = \mathcal{T}(0) = 0.$$

Within given time, a specific point can be reached in different numbers of steps N . If the step lengths $X(i)$ and waiting times $T(i)$ are independent, the number of steps $N(t)$ taken until time t is determined by the process of drawing the waiting times. Let us introduce the probability that the i th step does not occur before time t ,

$$F_{T(i)}(t) = \int_t^\infty f_{T(i)}(t') dt',$$

where $f_{T(i)}$ is the probability density corresponding to the distribution of waiting times. The probability of making n steps within the time interval $[0, t]$ is then

$$F_{N(t)}(n) = \int_0^t f_T^{*n}(t') F_{T(n+1)}(t - t') dt' = (f_T^{*n} * F_{T(n+1)})(t),$$

where

$$f_T^{*n} = f_{T(1)} * f_{T(2)} * \cdots * f_{T(n)}.$$

The distribution $F_{N(t)}$ can be calculated by using the Laplace transformation in time domain and the Fourier transformation in spatial domain: this allows one to avoid convolutions and operate with products of functions in transform spaces. The procedure, which we can not discuss in detail, leads to the Montroll–Weiss equation [20], helping us to identify four parameter regions corresponding to distributions of step lengths (density f_X) and waiting times (density f_T) with different dependencies of the scale L on time t , which determine the diffusion properties of the random walk. These regions are shown in Fig. 6.9 (right) and quantified below. We assume that the distributions of step lengths and waiting times do not change during the walk, so that $f_{X(i)} = f_X$ and $f_{T(i)} = f_T$.

Normal diffusion with spatial scale $L \sim t^{1/2}$ is obtained when $E[T] < \infty$, $\sigma_X < \infty$.

Sub-diffusion with scale $L \sim t^{\beta/2}$ is obtained with $E[T] = \infty$, $\sigma_X < \infty$ and distribution of waiting times

$$f_T(t) \sim \frac{1}{t^{1+\beta}}, \quad \beta \in (0, 1).$$

Super-diffusion with $L \sim t^{1/\alpha}$ is obtained with $E[T] < \infty$, $\sigma_X = \infty$ and distribution of step lengths

$$f_X(x) \sim \frac{1}{x^{1+\alpha}}, \quad \alpha \in (0, 2).$$

When $E[T] = \infty$ and $\sigma_X = \infty$, and when

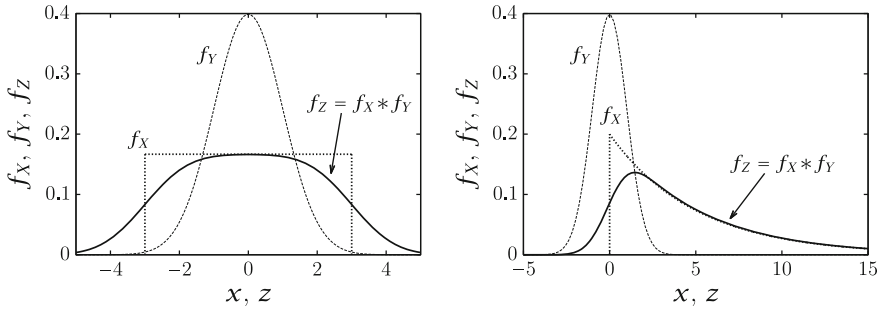


Fig. 6.11 [Left] Convolution of the uniform distribution $U(-3, 3)$ and the standardized normal distribution $N(0, 1)$. [Right] Convolution of the exponential distribution with parameter $\lambda = 0.2$ and the standardized normal distribution

$$f_X(x) \sim \frac{1}{x^{1+\alpha}}, \quad f_T(t) \sim \frac{1}{t^{1+\beta}}, \quad \alpha \in (0, 2), \quad \beta \in (0, 1),$$

the scale is $L \sim t^{\beta/\alpha}$. The walks are super-diffusive if $2\beta > \alpha$ and sub-diffusive otherwise. Processes for which $E[T] = \infty$ are deeply non-Markovian: this means that the values of the process at given time depend on its whole history, not just on the immediately preceding state. Further reading can be found in [19, 21].

6.9 Problems

6.9.1 Convolutions with the Normal Distribution

Calculate the convolution of the normal distribution with the ① uniform, ② normal and ③ exponential distributions!

🔗 ① The convolution of the uniform distribution with the density $f_X(x) = 1/(b - a)$ (see (3.1)) and the normal distribution with the density f_Y (Definition (3.7)) is

$$\begin{aligned} f_Z(z) &= \int_a^b f_X(x)f_Y(z - x) dx = \frac{1}{b - a} \frac{1}{\sqrt{2\pi} \sigma} \int_a^b \exp\left[-\frac{(z - x)^2}{2\sigma^2}\right] dx \\ &= \frac{1}{b - a} \frac{1}{\sqrt{2\pi}} \int_{(a-z)/\sigma}^{(b-z)/\sigma} e^{-u^2/2} du = \frac{1}{2(b - a)} \left[\operatorname{erf}\left(\frac{b - z}{\sqrt{2} \sigma}\right) - \operatorname{erf}\left(\frac{a - z}{\sqrt{2} \sigma}\right) \right]. \end{aligned}$$

This function is shown in Fig. 6.11 (left).

Part ② is easily solved by using characteristic functions (B.20) and property (B.22):

$$\phi_Z(t) = \phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) = e^{i(\mu_X+\mu_Y)t} e^{-(\sigma_X^2+\sigma_Y^2)t^2/2} = e^{i\mu_Z t} e^{-\sigma_Z^2 t^2/2}.$$

From this it is clear that the convolution of two normal distributions with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2 is also a normal distribution, with mean $\mu_Z = \mu_X + \mu_Y$ and variance $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$.

Problem ③ requires us to convolute the distribution with the probability density $f_X(x) = \lambda \exp(-\lambda x)$ (see (3.4)) and the normal distribution, where we set $\mu = 0$:

$$f_Z(z) = \frac{\lambda}{\sqrt{2\pi}\sigma} \int_{-\infty}^z e^{-\lambda(z-y)} e^{-y^2/(2\sigma^2)} dy.$$

Upon rearranging the exponent,

$$\begin{aligned} -\lambda(z-y) - \frac{y^2}{2\sigma^2} &= -\frac{\lambda}{2\sigma^2} \left(2\sigma^2(z-y) + \frac{y^2}{\lambda} + \lambda\sigma^4 - \lambda\sigma^4 \right) \\ &= -\lambda z + \frac{\lambda^2\sigma^2}{2} - \frac{1}{2\sigma^2} (y - \lambda\sigma^2)^2, \end{aligned}$$


it follows that

$$\begin{aligned} f_Z(z) &= \frac{\lambda}{\sqrt{2\pi}\sigma} \exp\left(-\lambda z + \frac{\lambda^2\sigma^2}{2}\right) \int_{-\infty}^{(z-\lambda\sigma^2)/\sigma} e^{-u^2/2} du \\ &= \lambda \exp\left(-\lambda z + \frac{\lambda^2\sigma^2}{2}\right) \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{z - \lambda\sigma^2}{\sqrt{2}\sigma}\right) \right]. \end{aligned}$$

This function is shown in Fig. 6.11 (right).

6.9.2 Spectral Line Width

The lines in emission spectra of atoms and molecules have finite widths [24]. Line broadening has three contributions: the *natural width* (N), *collisional broadening* (C) due to inelastic collisions of radiating particles, and *Doppler broadening* (D). Calculate a realistic spectral line profile by convoluting these three distributions.

 The natural width of the line in spontaneous emission—usually the smallest contribution to broadening—has a Lorentz (Cauchy) profile with a width of $\Delta\nu_N$,

$$\phi_N(\nu) = \frac{1}{\pi} \frac{\Delta\nu_N/2}{(\nu - \nu_0)^2 + (\Delta\nu_N/2)^2}.$$

As noted in the discussion of (3.19), such a distribution embodies a Fourier transformation of the exponential time dependence of the decays into Fourier space.

The broadening due to inelastic collisions depends on pressure and temperature—approximately one has $\Delta\nu_C \propto p/\sqrt{T}$ —and has a Cauchy profile as well:

$$\phi_C(\nu) = \frac{1}{\pi} \frac{\Delta\nu_C/2}{(\nu - \nu_0)^2 + (\Delta\nu_C/2)^2}. \quad (6.17)$$

A convolution of two Cauchy distributions is again a Cauchy distribution,

$$\phi_{N+C}(\nu) = \int_{-\infty}^{\infty} \phi_N(\rho)\phi_C(\nu - \rho) d\rho = \frac{1}{\pi} \frac{(\Delta\nu_N + \Delta\nu_C)/2}{(\nu - \nu_0)^2 + (\Delta\nu_N + \Delta\nu_C)^2/4},$$

where we have shifted the origin of ϕ_C before integrating by setting $\nu_0 = 0$ in (6.17). If we had failed to do that, the peak of the convoluted distribution ϕ_{N+C} would shift from ν_0 to $2\nu_0$ —see Sect. 6.1.1!

The Doppler effect is proportional to the velocity of the radiating objects, which is normally distributed (see (3.14) for a single velocity component), hence the corresponding contribution to the line profile has the form

$$\phi_D(\nu) = \frac{2\sqrt{\log 2}}{\sqrt{\pi}\Delta\nu_D} \exp \left\{ - \left(\frac{2\sqrt{\log 2}}{\Delta\nu_D} (\nu - \nu_0) \right)^2 \right\}.$$

The final spectral line shape is calculated by the convolution of the distributions ϕ_{N+C} and ϕ_D , where again the origin must be shifted. We obtain

$$\phi_V(\nu) = \int_{-\infty}^{\infty} \phi_{N+C}(\rho)\phi_D(\nu - \rho) d\rho = \frac{2\sqrt{\log 2}}{\sqrt{\pi}\Delta\nu_D} \left\{ \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-x^2}}{(w-x)^2 + a^2} dx \right\},$$

where

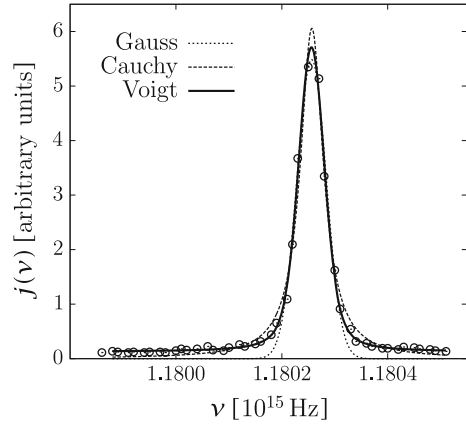
$$a = \sqrt{\log 2} \frac{\Delta\nu_N + \Delta\nu_C}{\Delta\nu_D}, \quad w = 2\sqrt{\log 2} \frac{\nu - \nu_0}{\Delta\nu_D}.$$

This is called the *Voigt* distribution. The natural width is usually neglected because $\Delta\nu_N \ll \Delta\nu_C, \Delta\nu_D$. How well ϕ_V describes an actual line shape (as compared to the Cauchy and Gaussian profiles) is shown in Fig. 6.12.

6.9.3 Random Structure of Polymer Molecules

(Adapted from [5].) A polymer molecule can be envisioned as a chain consisting of a large number of equal, rigid, thin segments of length L . Starting at the origin, a molecule grows by attaching to the current terminal point further and further segments in arbitrary directions in space. ① What is the probability distribution for the position of the terminal point? ② Calculate the expected distance \bar{R} between the initial and terminal point of the chain and \bar{R}^2 !

Fig. 6.12 Description of the Si (III) emission line at the wave-length of 254.182 nm (compare to Fig. 3.6) by a Gaussian (normal), Cauchy (Lorentz) and Voigt distribution with added constant background



① When a new segment is attached to the chain, it “chooses” its orientation at random: the directional distribution is therefore isotropic, $f_{\Theta}(\cos \theta) = dF_{\Theta}/d(\cos \theta) = 1/2$. For a projection of a single segment onto an arbitrary direction (e.g. x) we have

$$\begin{aligned}\overline{X_1} &= L \overline{\cos \Theta} = L \int_0^{\pi} \cos \theta f_{\Theta}(\cos \theta) \sin \theta d\theta = 0, \\ \sigma_{X_1}^2 &= \overline{X_1^2} = L^2 \overline{\cos^2 \Theta} = L^2 \int_0^{\pi} \cos^2 \theta f_{\Theta}(\cos \theta) \sin \theta d\theta = \frac{L^2}{3}.\end{aligned}\quad (6.18)$$

The X -coordinate of the terminal point of an N -segment chain is the sum of independent and identically distributed random variables X_1 so, by the central limit theorem, it is nearly normally distributed at large N , with expected value $\overline{X} = N\overline{X_1} = 0$ and variance $\sigma_X^2 = \overline{X^2} = N\sigma_{X_1}^2 = NL^2/3$. The corresponding probability density is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{x^2}{2\sigma_X^2}\right) = \sqrt{\frac{3}{2\pi NL^2}} \exp\left(-\frac{3x^2}{2NL^2}\right).$$

The x , y and z projections are *not* independent when a single segment is attached, but they are independent on average (after many attachments), so the same reasoning applies to Y and Z coordinates. Since $R^2 = X^2 + Y^2 + Z^2$, the probability density corresponding to the radial distribution of the terminal point of the chain is

$$f_R(r) = f_X(x)f_X(y)f_X(z) = \left(\frac{3}{2\pi NL^2}\right)^{3/2} \exp\left(-\frac{3r^2}{2NL^2}\right).\quad (6.19)$$

② This can be used to calculate the expected values of R and R^2 :

$$\overline{R} = \int_0^{\infty} r f_R(r) 4\pi r^2 dr = L\sqrt{\frac{8N}{3\pi}}, \quad \overline{R^2} = \int_0^{\infty} r^2 f_R(r) 4\pi r^2 dr = NL^2.$$

The latter can also be derived by recalling (6.18), since

$$\overline{R^2} = \overline{X^2 + Y^2 + Z^2} = \overline{X^2} + \overline{Y^2} + \overline{Z^2} = 3 \frac{NL^2}{3} = NL^2.$$

There is yet another path to the same result. Each segment ($n = 1, 2, \dots, N$) is defined by a vector $\mathbf{r}_n = (x_n, y_n, z_n)^T$. We are interested in the average square of the sum vector,

$$R^2 = |\mathbf{R}|^2 = \left(\sum_{m=1}^N \mathbf{r}_m \right)^T \left(\sum_{n=1}^N \mathbf{r}_n \right) = \sum_n r_n^2 + \sum_{m \neq n} \mathbf{r}_m^T \mathbf{r}_n.$$

Averaging the second sum yields zero due to random orientations, $\overline{\mathbf{r}_m^T \mathbf{r}_n} = 0$, hence

$$\overline{R^2} = \sum_{n=1}^N \overline{r_n^2} = N \overline{r_1^2} = NL^2.$$

6.9.4 Scattering of Thermal Neutrons in Lead

(Adapted from [5].) A neutron moves with velocity v in lead and scatters elastically off lead nuclei. The average time between collisions is τ , corresponding to the mean free path $\lambda = v\tau$. The times between consecutive collisions are mutually independent, and each scattering is isotropic. ① What is the (spatial) probability distribution of neutrons at long times? Calculate the average distance \overline{R} of neutrons from the origin and $\overline{R^2}$! ② Demonstrate that $\overline{R^2}$ is proportional to time, so the process has the usual diffusive nature! The diffusion coefficient D is defined by the relation $\overline{R^2} = 6Dt$. How does D depend on λ and v ?

🔍 ① Isotropic scattering implies $f_\Theta(\cos \theta) = dF_\Theta/d(\cos \theta) = 1/2$. But we must also take into account the times between collisions or the distances l traversed by the neutron between collisions, $f_T(t) = dF_T/dt = \tau^{-1} \exp(-t/\tau)$, thus

$$f_L(l) = \frac{dF_L}{dl} = \frac{dF_T}{dt} \frac{dt}{dl} = \frac{1}{\tau} e^{-t/\tau} \frac{1}{v} = \frac{1}{\lambda} e^{-l/\lambda},$$

where $l = vt$. The joint probability density of the linear and angular variable, relevant to each collision, is therefore

$$f_{L,\Theta}(l, \cos \theta) = \frac{1}{2\lambda} e^{-l/\lambda}.$$

The expected value of the projection of the neutron trajectory between two collisions onto the x -axis and the corresponding variance are

$$\bar{X}_1 = \overline{L \cos \Theta} = \int_0^\infty \int_0^\pi l \cos \theta f_{L, \Theta}(l, \cos \theta) dl \sin \theta d\theta = 0,$$

$$\sigma_{X_1}^2 = \overline{X_1^2} = \overline{L^2 \cos^2 \Theta} = \int_0^\infty \int_0^\pi l^2 \cos^2 \theta f_{L, \Theta}(l, \cos \theta) dl \sin \theta d\theta = \frac{2\lambda^2}{3}.$$

Hence, as in Sect. 6.9.3, $\bar{X} = N\bar{X}_1 = 0$ after N scatterings, while $\sigma_X^2 = N\sigma_{X_1}^2 = 2N\lambda^2/3$. Therefore the probability density for the distribution of R (distance from the origin to the current collision point) has the same functional form as in (6.19),

$$f_R(r) = \left(\frac{3}{4\pi N\lambda^2} \right)^{3/2} \exp\left(-\frac{3r^2}{4N\lambda^2} \right),$$

one only needs to insert the variance $2N\lambda^2/3$ instead of $N\lambda^2/3$. It follows that

$$\bar{R} = \int_0^\infty r f_R(r) 4\pi r^2 dr = \lambda \sqrt{\frac{16N}{3\pi}}, \quad \overline{R^2} = \int_0^\infty r^2 f_R(r) 4\pi r^2 dr = 2N\lambda^2.$$

② The elapsed time after N collisions is $t = N\lambda/v$, so that indeed $\overline{R^2}$ is proportional to time, $\overline{R^2} = 2N\lambda^2 = 2(vt/\lambda)\lambda^2 = 2vt\lambda$. From the definition $\overline{R^2} = 6Dt$ it follows that


$$D = \frac{\lambda v}{3}.$$

6.9.5 Distribution of Extreme Values of Normal Variables ★

Let continuous random variables X_i ($i = 1, 2, \dots, n$) be normally distributed, $X_i \sim N(0, 1)$, with the corresponding distribution function $F_X(x) = \Phi(x)$ and probability density $f_X(x) = \phi(x)$ for each variable:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt, \quad \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

What is the distribution function F_{M_n} of the values $M_n = \max\{X_1, X_2, \dots, X_n\}$? This Problem [13] is a continuation of the Example on p. 157 and represents a method to determine the parameters a_n and b_n for the scaling formula (6.10) such that the limiting distribution (6.9) is non-degenerate.

 Let $0 \leq \tau \leq \infty$ and let $\{u_n\}$ be a sequence of real numbers such that

$$1 - F_X(u_n) \rightarrow \frac{\tau}{n} \quad \text{when } n \rightarrow \infty. \quad (6.20)$$

By definition of the exponential function by a series we get

$$F_{M_n}(u_n) = P(M_n \leq u_n) = F_X^n(u_n) = [1 - (1 - F_X(u_n))]^n = \left[1 - \frac{\tau}{n} + \mathcal{O}(1/n)\right]^n \sim e^{-\tau},$$

when $n \rightarrow \infty$. The leading dependence of the distribution function, $F_{M_n} \sim \exp(-\tau)$, follows without explicit reference to the parent distribution $F_X(x)$ being normal! A motivation for a specific form of τ can then be found in the asymptotic property of the normal distribution

$$1 - \Phi(z) \sim \frac{\phi(z)}{z}, \quad n \rightarrow \infty. \quad (6.21)$$

Let $\tau = e^{-x}$. The reason for this choice, fully consistent with (6.20), will become clear in the following: this is the only way to obtain in the final expression a *linear* dependence on x in the rescaled argument of the distribution function. By comparing (6.20) to (6.21) we obtain

$$1 - \Phi(u_n) \sim \frac{e^{-x}}{n} \sim \frac{\phi(u_n)}{u_n} \Rightarrow \frac{1}{n} e^{-x} \frac{u_n}{\phi(u_n)} \rightarrow 1.$$

Taking the logarithm we get $-\log n - x + \log u_n - \log \phi(u_n) \rightarrow 0$ or

$$-\log n - x + \log u_n - \frac{1}{2} \log 2\pi + \frac{1}{2} u_n^2 \rightarrow 0. \quad (6.22)$$

For fixed x in the limit $n \rightarrow \infty$ one therefore has $u_n^2/(2 \log n) \rightarrow 1$, so that taking the logarithm again yields $2 \log u_n - \log 2 - \log \log n \rightarrow 0$ or

$$\log u_n = \frac{1}{2}(\log 2 + \log \log n) + \mathcal{O}(1).$$

Inserting this in (6.22), we get $\frac{1}{2}u_n^2 = x + \log n - \frac{1}{2} \log 4\pi - \frac{1}{2} \log \log n + \mathcal{O}(1)$, hence

$$u_n^2 = 2 \log n \left[1 + \frac{x - \frac{1}{2} \log 4\pi - \frac{1}{2} \log \log n}{\log n} + \mathcal{O}\left(\frac{1}{\log n}\right) \right],$$

and finally, after taking the square root,

$$u_n = \sqrt{2 \log n} \left[1 + \frac{x - \frac{1}{2} \log 4\pi - \frac{1}{2} \log \log n}{2 \log n} + \mathcal{O}\left(\frac{1}{\log n}\right) \right].$$

This expression has the form

$$u_n = a_n x + b_n + \mathcal{O}((\log n)^{-1/2}) = a_n x + b_n + \mathcal{O}(a_n),$$

whence we read off the normalization constants a_n and b_n :

$$a_n = \frac{1}{\sqrt{2 \log n}}, \quad b_n = \sqrt{2 \log n} - \frac{\log \log n + \log 4\pi}{2\sqrt{2 \log n}}. \quad (6.23)$$

These constants imply $P(M_n \leq a_n x + b_n + \mathcal{O}(a_n)) \rightarrow \exp(-e^{-x})$, that is,

$$F_{M_n}(x) = P\left(\frac{M_n - b_n}{a_n} + \mathcal{O}(1) \leq x\right) \rightarrow \exp(-e^{-x}).$$

The distribution of extreme values of normally distributed variables is therefore of the Gumbel type (6.12) with normalization constants (6.23).

References

1. D.L. Evans, L.M. Leemis, Algorithms for computing the distributions of sums of discrete random variables. *Math. Comp. Model.* **40**, 1429 (2004)
2. W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, 2nd edn. (Wiley, New York, 1971)
3. I.S. Tyurin, On the accuracy of the Gaussian approximation. *Dokl. Math.* **80**, 840 (2009)
4. A. Lyon, Why are normal distributions normal? *Brit. J. Phil. Sci.* **65**, 621 (2014)
5. I. Kuščer, A. Kodre, *Mathematik in Physik und Technik* (Springer, Berlin, 1993)
6. J.P. Nolan, *Stable Distributions—Models for Heavy Tailed Data* (Birkhäuser, Boston, 2010)
7. S. Borak, W. Härdle, R. Weron, *Stable Distributions, SFB 649 Discussion Paper 2005–008* (Humboldt University Berlin, Berlin, 2005)
8. S. Širca, M. Horvat, *Computational Methods for Physicists* (Springer, Berlin, 2012)
9. GSL (GNU Scientific Library), <http://www.gnu.org/software/gsl>
10. G.G. Márquez, *One Hundred Years of Solitude* (HarperCollins, New York, 2006)
11. E.J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958)
12. S. Coles, *An Introduction to Statistical Modeling of Extreme Values* (Springer, Berlin, 2001)
13. M.R. Leadbetter, G. Lindgren, H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes* (Springer, New York, 1983)
14. S.I. Resnick, *Extreme Values, Regular Variation, and Point Processes* (Springer, New York, 1987)
15. R.A. Fisher, L.H.C. Tippett, On the estimation of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.* **24**, 180 (1928)
16. B.V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**, 423 (1943)
17. MeteoSwiss IDAWEB, <http://gate.meteoswiss.ch/idaweb/>
18. B.D. Hughes, *Random Walks and Random Environments: Vol. 1: Random Walks* (Oxford University Press, New York, 1995)
19. M. Bazant, *Random walks and diffusion*, MIT OpenCourseWare, Course 18.366, <http://ocw.mit.edu/courses/mathematics/>
20. E.W. Montroll, G.H. Weiss, Random walks on lattices II. *J. Math. Phys.* **6**, 167 (1965)
21. R. Metzler, J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* **339**, 1 (2000)
22. M.F. Shlesinger, B.J. West, J. Klafter, Lévy dynamics of enhanced diffusion: application to turbulence. *Phys. Rev. Lett.* **58**, 1100 (1987)
23. V. Tejedor, R. Metzler, Anomalous diffusion in correlated continuous time random walks. *J. Phys. A: Math. Theor.* **43**, 082002 (2010)
24. J.J. Brehm, W.J. Mullin, *Introduction to the Structure of Matter* (Wiley, New York, 1989)