Appendix A
Probability as Measure ⋆

Abstract Probability is defined in a mathematically strict manner as a measure. The Dirac delta is defined.

Here we give a mathematically more strict definition of probability and the Dirac delta "function" (phenomenologically introduced in (2.1)) as *measures*.

σ-Algebra

Let X be a non-empty set. A family of subsets $\mathcal{A}$ of X is called a *σ-algebra on X* if it has the following properties:

1. $X \in \mathcal{A}$;
2. for each subset $S \in \mathcal{A}$ also $X \backslash S \in \mathcal{A}$;
3. for each countable family $\{A_i : i \in \mathbb{N}\}$ of elements from $\mathcal{A}$, the union $\bigcup_{i\in\mathbb{N}} A_i$ also belongs to $\mathcal{A}$.

The elements of the family $\mathcal{A}$ are called *measurable sets*, and the set X, furnished with $\mathcal{A}$, is called a *measurable space*. A measurable space is the pair $(X, \mathcal{A})$.

Positive Measure

Examples of positive measures are: length of subset (interval) in $\mathbb{R}$, area of planar geometric shapes, volume of bodies in space. To generalize these special cases to arbitrary measurable spaces, one defines a *positive measure* (or simply *measure*) on a measurable space $(X, \mathcal{A})$ as the function

$$\mu : \mathcal{A} \to [0, \infty],$$

satisfying the conditions

1. $\mu(\{\,\}) = 0$ and
2. $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$

for each countable family of disjoint subsets A_n and $\mathcal{A}$. A positive measure μ on a measurable space $(X, \mathcal{A})$ is called a *finite measure* if $\mu(X) < \infty$.

S. Širca, *Probability for Physicists*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-31611-6

Probability as a Positive Measure

For random experiments with sample space S we define the event space $\mathcal{E}$, which is the power set of S, i.e. the set of all subsets of S, including the empty set and S itself. The mapping $P : \mathcal{E} \to \mathbb{R}$ is called a *probability measure* on measurable space $(S, \mathcal{E})$ if the following holds true:

1. $P(A) \geq 0$ for each $A \in \mathcal{E}$;
2. $P(S) = 1$;
3. if $A_1, A_2, \ldots$ are mutually exclusive events in $\mathcal{E}$, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

In simple cases, e.g. in throwing a die, $\mathcal{E}$ may indeed be identified with the power set of S, but often we restrict ourselves to a much smaller set: for example, it turns out [1] that there does not exist a probability measure P that would be defined on *all* subsets of the interval [0, 1] and would satisfy the requirement $P(\{x\}) = 0$ for each $x \in [0, 1]$.

Dirac Measure

Let X be an arbitrary non-empty set, $\mathcal{A}$ its power set, and $x \in X$ its arbitrary element. Then the prescription

$$\mu(A; x) = \begin{cases} 1 \ ; \ x \in A, \\ 0 \ ; \ x \in X \backslash A, \end{cases}$$

defines a positive finite measure on measurable space $(X, \mathcal{A})$ called the *Dirac measure at* x and denoted by δ_x. For each function $f : X \to \overline{\mathbb{R}}$ the integration with respect to the Dirac delta represents the evaluation of the function at x,

$$\int_X f \, \mathrm{d}\delta_x = f(x).$$

In the special case $X = \mathbb{R}$ we have

$$\int_A f \, \mathrm{d}\delta_x = f(x)\delta_x(A) = \int_A f(t)\delta(t - x) \, \mathrm{d}t, \qquad A \subseteq X,$$

for each measurable function $f : \mathbb{R} \to \mathbb{R}$, where $\mathrm{d}\delta_x(t) = \delta(t - x) \, \mathrm{d}t$ and δ is the Dirac delta “function”.

Reference

1. G. Grimmett, D. Welsh, *Probability. An introduction*, 2nd edn. (Oxford University Press, Oxford, 2014)

Appendix B
Generating and Characteristic Functions ⋆

Abstract Generating and characteristic functions are introduced as transformations of random variables that facilitate the calculation of certain distribution properties, in particular its moments and their convolutions. The problems of inverting probability-generating and characteristic functions, as well as of the existence of generating functions are presented.

Generating and characteristic functions are transformations of probability functions. As such they are not as easy to interpret as the distributions themselves, but in certain cases they offer immense benefits in terms of elegant calculation of distribution properties—for example, their moments—or quantities relating the distribution to each other, in particular their convolutions.

B.1 Probability-Generating Functions

Generating functions are applicable to random variables whose possible values are non-negative integers or their subsets. Such variables are called *non-negative integer random variables*. Let X be such a variable with the probability function

$$f_n = P(X = n). \tag{B.1}$$

Then the function of a real variable

$$G_X(z) = \sum_{n=0}^{\infty} P(X = n) z^n = \sum_{n=0}^{\infty} f_n z^n, \qquad |z| \le 1,$$

is the *[probability]-generating function* of the random variable X, distributed according to (B.1). The coefficients in this power expansion are probabilities with values

S. Širca, *Probability for Physicists*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-31611-6

between 0 and 1. Since they are bounded, the series is absolutely convergent for any $|z| < 1$, and due to

$$G(1) = \sum_{n=0}^{\infty} f_n = 1$$

the series converges at least on $[-1, 1]$. By comparison to (4.7) we also see that the generating function is equal to the expected value of the random variable z^X,

$$G_X(z) = E\left[z^X\right]. \tag{B.2}$$

The generating function G_X uniquely determines the probability function of X. This can be seen if we take the derivative of the series with respect to z:

$$\frac{\mathrm{d}^r}{\mathrm{d}z^r} G_X(z) = \sum_{n=r}^{\infty} n(n-1)\cdots(n-r+1) z^{n-r} f_n, \qquad r = 1, 2, \ldots$$

Namely, by setting $z = 0$ we obtain

$$f_r = P(X = r) = \frac{1}{r!}\left[\frac{\mathrm{d}^r}{\mathrm{d}z^r} G_X(z)\right]\bigg|_{z=0}, \qquad r = 0, 1, 2, \ldots, \tag{B.3}$$

so indeed by taking consecutive derivatives the complete distribution is determined, as all its components f_r are combed through. Why does this matter? Frequently only the generating function of X is available, while its probability function is not explicitly known. In such cases its components can be calculated by using (B.3): see Sect. B.1.2. Besides, taking the derivatives of the generating function is an easy way to produce the moments of X. For instance, by taking the first and second derivative we get

$$G_X'(z) = \sum_{n=1}^{\infty} n z^{n-1} f_n, \qquad G_X''(z) = \sum_{n=2}^{\infty} n(n-1) z^{n-2} f_n.$$

On the other hand,

$$E[X] = \sum_{n=1}^{\infty} n f_n = \lim_{z \nearrow 1} G_X'(z) = G_X'(1), \tag{B.4}$$

$$E[X(X-1)] = \sum_{n=2}^{\infty} n(n-1) f_n = \lim_{z \nearrow 1} G_X''(z) = G_X''(1),$$

therefore

$$E[X^2] = E[X(X-1)] + E[X] = G_X''(1) + G_X'(1),$$

$$\mathrm{var}[X] = E[X^2] - (E[X])^2 = G_X''(1) + G_X'(1)\left[1 - G_X'(1)\right]. \tag{B.5}$$

Individual moments can be calculated without such detours by using the formula

$$E[X^r] = \left[\left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)^r G_X(z)\right]\bigg|_{z=1}.$$

Example The generating function of the binomial distribution (Definition (5.1)) with the probability function $f_n = P(X = n; N, p)$ is

$$G_X(z) = \sum_{n=0}^{N} f_n z^n = \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} z^n = \sum_{n=0}^{N} \binom{N}{n} (pz)^n q^{N-n} = (pz+q)^N.$$

Its first derivative is $G'_X(z) = Np(pz+q)^{N-1}$, and the second derivative is $G''_X(z) = N(N-1)p^2(pz+q)^{N-2}$, thus $G'_X(1) = Np$ and $G''_X(1) = N(N-1)p^2$. From (B.4) and (B.5) it follows that

$$E[X] = Np, \qquad \mathrm{var}[X] = N(N-1)p^2 + Np(1-Np) = Npq,$$

which are familiar expressions (5.5). ◁

Example The generating function of the Poisson distribution (5.11) is

$$G_X(z) = \sum_{n=0}^{\infty} f_n z^n = \sum_{n=0}^{\infty} \frac{\lambda^n \mathrm{e}^{-\lambda}}{n!} z^n = \mathrm{e}^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda z)^n}{n!} = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda z} = \mathrm{e}^{\lambda(z-1)}.$$

Differentiation gives $G'_X(z) = \lambda \mathrm{e}^{\lambda(z-1)}$ and $G''_X(z) = \lambda^2 \mathrm{e}^{\lambda(z-1)}$, hence $G'_X(1) = \lambda$ and $G''_X(1) = \lambda^2$. From (B.4) and (B.5) it follows that

$$E[X] = \lambda, \qquad \mathrm{var}[X] = \lambda^2 + \lambda(1-\lambda) = \lambda,$$

which, again, we know from (5.12). ◁

B.1.1 Generating Functions and Convolution

Let us discuss mutually independent integer random variables X and Y with the probability functions

$$f_n = P(X = n), \qquad g_n = P(Y = n), \qquad n = 0, 1, 2, \ldots$$

Their sum $Z = X + Y$ is also an integer random variable with the corresponding probability function

$$h_n = P(Z = n), \qquad n = 0, 1, 2, \ldots$$

The sum Z has value n when the variables X and Y have values $(X, Y) = (0, n)$ or $(1, n-1)$ or $(2, n-2)$, and so on. Since X and Y are independent, the probabilities of these simultaneous events are $P(X = 0)P(Y = n)$ or $P(X = 1)P(Y = n - 1)$ or $P(X = 2)P(Y = n - 2)$, and so on. In other words,

$$h_n = P(Z = n) = \sum_{j=0}^{n} P(X = j)P(Y = n - j) = \sum_{j=0}^{n} f_j g_{n-j}, \qquad n = 0, 1, 2, \ldots$$

We are looking at a discrete convolution of the sequences $\{f_n\}$ and $\{g_n\}$, which we denote as

$$\{h_n\} = \{f_n\} * \{g_n\}.$$

This a discrete analogue of the Definition (6.1) or

$$h_Z(z) = \int_{-\infty}^{\infty} f_X(x) g_Y(z - x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f_X(z - y) g_Y(y)\,\mathrm{d}y.$$

Where do generating functions come into play? Let

$$G_X(z) = \sum_{n=0}^{\infty} f_n z^n, \qquad G_Y(z) = \sum_{n=0}^{\infty} g_n z^n$$

be generating functions of X and Y. The generating function of their sum is then

$$G_Z(z) = \sum_{n=0}^{\infty} h_n z^n = \sum_{n=0}^{\infty} \left(\sum_{j=0}^{n} f_j g_{n-j} \right) z^n = \sum_{n=0}^{\infty} \sum_{j=0}^{n} f_j z^j g_{n-j} z^{n-j}.$$

The series on the right is just the product of the series $G_X(z)$ and $G_Y(z)$, so

$$G_Z(z) = G_X(z) G_Y(z). \tag{B.6}$$

The generating function of the sum of independent integer variables is therefore equal to the product of the generating functions of the two terms. An even faster way to this result would be to consider (B.2): if X and Y are independent, the variables $U = z^X$ and $V = z^Y$ are independent, too; since for independent variables U and V one has $E[UV] = E[U]E[V]$, this also means

$$E\big[z^{X+Y}\big] = E\big[z^X z^Y\big] = E\big[z^X\big] E\big[z^Y\big], \tag{B.7}$$

whence (B.6) follows. This should not be read in the opposite direction: having $G_Z(z) = G_X(z)G_Y(z)$ does not necessarily mean that X and Y are independent. But the relation *can* be generalized to several independent variables: if $X_1, X_2, \ldots, X_n$ are mutually independent random variables with generating functions $G_{X_1}(z), G_{X_2}(z), \ldots, G_{X_n}(z)$ and Z is their sum with the generating function $G_Z(z)$, then

$$G_Z(z) = G_{X_1}(z) G_{X_2}(z) \cdots G_{X_n}(z). \tag{B.8}$$

Multiplying generating functions is a much simpler operation than computing convolution sums, so convolution of independent integer random variables is most easily performed by using (B.6) and (B.8).

Example We demonstrate that the convolution of two Poisson distributions is a Poisson distribution. In the Example on p. 148 we have derived this result by a direct calculation of the convolution sum. But if one calls generating functions $G_X(z) = \mathrm{e}^{\lambda(z-1)}$ and $G_Y(z) = \mathrm{e}^{\mu(z-1)}$ to the rescue, the effort is minimal:

$$G_Z(z) = G_X(z) G_Y(z) = \mathrm{e}^{\lambda(z-1)} \mathrm{e}^{\mu(z-1)} = \mathrm{e}^{(\lambda+\mu)(z-1)}.$$

Clearly the variable Z has the generating function of the Poisson distribution with parameter $\lambda + \mu$, so indeed $Z \sim \mathrm{Poisson}(\lambda + \mu)$. ◁

B.1.2 Inverting the Probability-Generating Function

The functional form $G_Z(z) = \mathrm{e}^{(\lambda+\mu)(z-1)}$ in the preceding Example immediately allowed us to conclude that Z is Poissonian, as we already knew the relation between the generating function and its inverse beforehand. The same procedure can be used for more complicated generating functions, as long as they can be split into sums of terms whose inverses are known.

But how do we compute the inverse of an arbitrary generating function? Formula (B.3) can be used for simple explicit functions, but analytic differentiation may be strenuous and is numerically unstable. The solution—in particular when the generating function is only known at discrete points—is offered by the Cauchy integral formula

$$G_X(a) = \frac{1}{2\pi \mathrm{i}} \oint_{\partial D} \frac{G_X(z)}{z - a} \, \mathrm{d}z,$$

where $D = \{z : |z - z_0| \le R\}$ is a subset completely contained in the definition domain of G_X (neighborhood of z_0), ∂D is its boundary and a is any point in the interior of D. For the nth derivative of G_X it holds that

$$G_X^{(n)}(a) = \frac{n!}{2\pi \mathrm{i}} \oint_{\partial D} \frac{G_X(z)}{(z - a)^{n+1}} \, \mathrm{d}z,$$

so the components f_n of the probability distribution of X—use of (B.3) requires derivatives of G_X at $a = 0$—are given by the integral

$$f_n = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{G_X(z)}{z^{n+1}}\,\mathrm{d}z.$$

The closed curve C is a circle in the complex plane. By the substitution $z = R\,\mathrm{e}^{\mathrm{i}\,u}$, where R must be such that G_X is analytic on D, we get

$$f_n = \frac{1}{2\pi R^n} \int_0^{2\pi} G_X\big(R\,\mathrm{e}^{\mathrm{i}\,u}\big)\,\mathrm{e}^{-\mathrm{i}\,nu}\,\mathrm{d}u.$$

The integral can be evaluated by using the trapezoidal rule, resulting in the following approximation for the true distribution f_n ($n = 0, 1, \ldots, N-1$):

$$\widetilde{f}_n \approx \frac{1}{N R^n} \sum_{m=0}^{N-1} G_X\left(R\,\mathrm{e}^{\mathrm{i}\,2\pi m/N}\right)\,\mathrm{e}^{-\mathrm{i}\,2\pi mn/N}, \qquad \widetilde{f}_{n+N} = \widetilde{f}_n, \tag{B.9}$$

which is the inverse discrete Fourier transformation scaled by R. Due to the discrete nature of the approximation, discretization and aliasing errors are thereby introduced (see, for example, [1], page 166), which can be controlled by the parameter R. For details see [2–4].

Example Let us pretend that we do not know the probability function of the Poisson distribution $f_n(\lambda) = \lambda^n \mathrm{e}^{-\lambda}/n!$, but only its generating function $G_X(z) = \mathrm{e}^{\lambda(z-1)}$. Take $\lambda = 2$, for instance: the exact values f_n up to $n = N = 20$ are shown in Fig. B.1 (left). We compute the approximations for f_n by inverting the generating function via (B.9) with different R, say, $R = 1.0$, $R = 2.0$ and $R = 0.5$. The absolute errors of these reconstructed probability functions are shown in Fig. B.1 (right). Note the absence of the value at $n = N$: due to the periodicity of the Fourier transform one has $f_N = f_0$. ◁

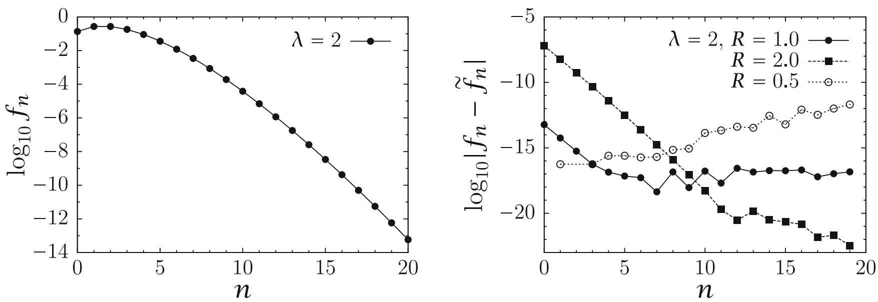


Fig. B.1 [Left] Poisson distribution with parameter $\lambda = 2$ in logarithmic scale. [Right] Difference between the exact values f_n and their approximations, calculated by inverting the probability-generating function by the discrete Fourier transformation (B.9), at several values of the parameter R

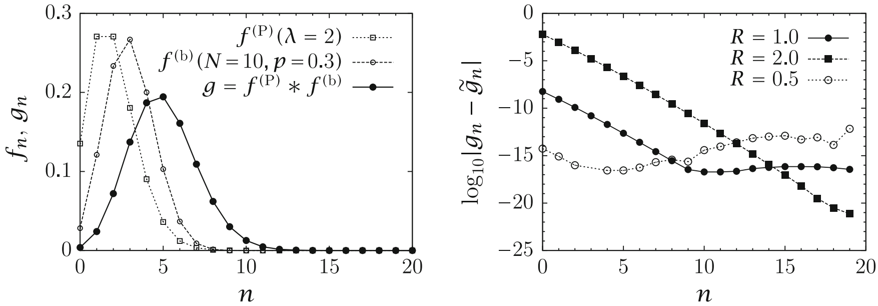


Fig. B.2 [Left] Poisson distribution with parameter $\lambda = 2$, binomial distribution with $N = 10$, $p = 0.3$, and their convolution. [Right] The difference between the exact probabilities g_n and their approximations $\widetilde{g}_n$, calculated by inverting the generating function and by using the discrete Fourier transformation with various values of R

Example It is instructive to compare the convolution of discrete distributions, calculated by the basic formula (6.4), and by multiplying generating functions according to (B.6). Take, for instance, the Poisson distribution

$$P(X = n) = f_n^{(P)} = \frac{\lambda^n e^{-\lambda}}{n!}, \qquad n = 0, 1, 2, \ldots,$$

where $\lambda = 2$, and the binomial distribution

$$P(Y = n) = f_n^{(b)} = \binom{N}{n} p^n q^{N-n}, \qquad n = 0, 1, 2, \ldots, N,$$

where $N = 10$ and $p = 0.3$. These distributions are shown in Fig. B.2 (left) at its left edge. By the definition of convolution we obtain the distribution

$$P(X + Y = n) = g_n = \left(f^{(P)} * f^{(b)}\right)_n = \sum_{i=0}^{n} f_i^{(P)} f_{n-i}^{(b)}, \tag{B.10}$$

indicated by full circles in the figure. We should expect the same result by multiplying the generating functions of both distributions and computing the inverse Fourier transformation of the product. Thus we compute

$$G_Z(z) = G_{X+Y}(z) = G_X(z)G_Y(z) = e^{\lambda(z-1)}(pz + q)^N,$$

and then use this function in formula (B.9):

$$\widetilde{g}_n \approx \frac{1}{N_{\mathrm{DFT}} R^n} \sum_{m=0}^{N_{\mathrm{DFT}}-1} G_Z\left(R\, e^{i\, 2\pi m/N_{\mathrm{DFT}}}\right) e^{-i\, 2\pi mn/N_{\mathrm{DFT}}}.$$

(Think about it: what N_{DFT} should one take in the above equation and what is the range of n in (B.10), considering that the definition domains of the distributions differ?) We thereby obtain the probabilities $\widetilde{g}_n$ that should be equal to g_n. How well this holds is shown in Fig. B.2 (right). ◁

B.2 Moment-Generating Functions

Probability-generating functions have been defined for random variables with non-negative integer values. The concept can be extended to random variables with arbitrary real values, if $E[z^X]$ (see (B.2)) is replaced by $E\big[\mathrm{e}^{tX}\big]$. If this expected value is finite for t on the interval $[t-T, t+T]$ for some $T > 0$, we may define the *moment-generating function*

$$M_X(t) = E\big[\mathrm{e}^{tX}\big] = \int_{-\infty}^{\infty} \mathrm{e}^{tx} f_X(x)\,\mathrm{d}x, \tag{B.11}$$

which is nothing but the continuous Laplace transform. In the case of a discrete probability distribution of X, which we shall not discuss separately from now on, the corresponding definition is

$$M_X(t) = E\big[\mathrm{e}^{tX}\big] = \sum_i \mathrm{e}^{tx_i} P(X = x_i).$$

The 'moment-generating' attribute is easy to explain if one expands e^{tX} in a power series and exchanges the order of summation and taking the expected values:

$$E\big[\mathrm{e}^{tX}\big] = E\left[\sum_{k=0}^{\infty} t^k \frac{X^k}{k!}\right] = 1 + tE[X] + \frac{t^2}{2!}E\big[X^2\big] + \frac{t^3}{3!}E\big[X^3\big] + \cdots.$$

Namely, individual distribution moments can be obtained by taking consecutive derivatives

$$E\big[X^r\big] = \left[\frac{\mathrm{d}^r M_X(t)}{\mathrm{d}t^r}\right]\Bigg|_{t=0}, \qquad r = 1, 2, \ldots, \tag{B.12}$$

thus $E[X] = M_X'(0)$, $E[X^2] = M_X''(0)$, and so on. Compare (B.3) and (B.12)!

Example Let us calculate the moment-generating function of a random variable distributed according to the Cauchy distribution (3.18):

$$M_X(t) = E\big[\mathrm{e}^{tX}\big] = \int_{-\infty}^{\infty} \mathrm{e}^{tx}\frac{1}{\pi}\frac{1}{1+x^2}\mathrm{d}x = \begin{cases} 1 & ;\ t = 0, \\ \infty & ;\ \text{otherwise}. \end{cases}$$

Now we see why Definition (B.11) had to be formulated so carefully: the expected value $E\left[e^{tX}\right]$ for arbitrary t may not even exist! This obstacle will be circumvented in Sect. B.3. ◁

Example What about the moment-generating function of a random variable distributed according to the standardized normal distribution (3.10)? By elementary integration[1] we immediately obtain

$$M_X(t) = E\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = e^{t^2/2}. \tag{B.13}$$

The main dish follows: we expand the exponential in a power series

$$M_X(t) = e^{t^2/2} = \sum_{k=0}^{\infty} \frac{(t^2/2)^k}{k!} = \sum_{k=0}^{\infty} \frac{(2k)!}{k!2^k} \frac{t^{2k}}{(2k)!} = \sum_{k=0}^{\infty} E\left[X^k\right] \frac{t^k}{k!}$$

and compare the terms with equal powers of t on both sides of the last equality. This gives us the odd moments

$$E\left[X^k\right] = 0, \qquad k \text{ odd},$$

while the even ones are

$$E\left[X^{2k}\right] = \frac{(2k)!}{k!2^k} = 1 \cdot 3 \cdot 5 \cdot (2k-1),$$

thus $E\left[X^2\right] = 1$, $E\left[X^4\right] = 3$, $E\left[X^6\right] = 15$, $E\left[X^8\right] = 105$, and so on. The first two values should already be familiar from Sect. 4.7, while the others were derived elegantly, with minimal effort. ◁

Let X and Y be random variables with moment-generating functions $M_X(t)$ and $M_Y(t)$. If X and Y are mutually independent, the same reasoning that brought us to (B.6) implies also

$$M_{X+Y}(t) = M_X(t) M_Y(t). \tag{B.14}$$

For random variables X and Y related through $Y = aX + b$, it holds that

$$M_Y(t) = e^{bt} M_X(at). \tag{B.15}$$

So the obvious generalization of (B.14) to a sum of several variables is at hand: if $X_1, X_2, \ldots, X_n$ are mutually independent random variables and $Y = c_1 X_1 +$

[1] Gaussian integrals with linear terms in the exponent can be handled by using the formulas

$$\int_{-\infty}^{\infty} e^{-ax^2/2+bx}\, dx = \sqrt{\frac{2\pi}{a}} e^{b^2/2a}, \qquad \int_{-\infty}^{\infty} e^{-ax^2/2+i\,bx}\, dx = \sqrt{\frac{2\pi}{a}} e^{-b^2/2a}.$$

$c_2 X_2 + \cdots + c_n X_n$ is their linear combination with real coefficients c_i, the moment-generating function of Y is equal to the product of moment-generating functions of individual variables X_i:

$$M_Y(t) = E\left[e^{tY}\right] = \prod_{i=1}^{n} E\left[e^{tX_i}\right] = M_{X_1}(c_1 t) M_{X_2}(c_2 t) \cdots M_{X_n}(c_n t). \tag{B.16}$$

Just as in (B.7) these recipes may not be read in reverse: $M_{X+Y}(t) = M_X(t) M_Y(t)$ does not necessarily mean that X and Y are independent.

Example The convolution problem from the Example on p. 148 can also be solved by generating functions. The moment-generating functions of X and Y are

$$M_X(t) = \sum_{n=-\infty}^{\infty} f_n e^{t x_n}, \qquad M_Y(t) = \sum_{n=-\infty}^{\infty} g_n e^{t y_n},$$

that is,

$$M_X(t) = 0.15\,e^{-3t} + 0.25\,e^{-t} + 0.1\,e^{2t} + 0.3\,e^{6t} + 0.2\,e^{8t},$$
$$M_Y(t) = 0.2\,e^{-2t} + 0.1\,e^{t} + 0.3\,e^{5t} + 0.4\,e^{8t}.$$

Since X and Y are mutually independent, the moment-generating function $M_Z(t)$ of their sum $Z = X + Y$ is the product of the individual moment-generating functions $M_X(t)$ and $M_Y(t)$:

$$\begin{aligned} M_Z(t) = M_X(t) M_Y(t) = \sum_n h_n e^{t z_n} \\ &= 0.03\,e^{-5t} + 0.05\,e^{-3t} + 0.015\,e^{-2t} + 0.045\,e^{0t} + 0.045\,e^{2t} \\ &\quad + 0.01\,e^{3t} + 0.135\,e^{4t} + 0.06\,e^{5t} + 0.04\,e^{6t} + 0.16\,e^{7t} + 0.02\,e^{9t} \\ &\quad + 0.04\,e^{10t} + 0.09\,e^{11t} + 0.06\,e^{13t} + 0.12\,e^{14t} + 0.08\,e^{16t}. \end{aligned}$$

All we need, then, is to read off the coefficient in front of e^{4t}, which is $h_4 = P(Z = 4) = 0.135$, and analogously for any other h_n. ◁

B.3 Characteristic Functions

Let X be a real (discrete or continuous) random variable and t a non-random real variable. The quantity

$$\phi_X(t) = E\left[e^{itX}\right], \quad t \in \mathbb{R}, \tag{B.17}$$

is called the *characteristic function* of the random variable X [5, 6]. In contrast to the moment-generating function a characteristic function exists regardless of the distribution of X, and its definition domain is the whole real axis. Any characteristic function satisfies

$$|\phi_X(t)| \le 1, \qquad \phi_X(0) = 1.$$

Besides, one has $\phi_X(t) = M_X(\mathrm{i}\,t)$ and $\phi_X(-\mathrm{i}\,t) = M_X(t)$, if M_X exists. If the distribution of X is discrete, with probability function $f_n = P(X = x_n)$, where $n = 0, 1, 2, \ldots$, it has the characteristic function

$$\phi_X(t) = \sum_{n=0}^{\infty} f_n \mathrm{e}^{\mathrm{i}\,t x_n}. \tag{B.18}$$

If the distribution is continuous, with probability density f_X, it corresponds to

$$\phi_X(t) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\,tx} f_X(x)\,\mathrm{d}x, \tag{B.19}$$

which is the usual Fourier transformation of f_X.

Example The Poisson distribution with the probability function

$$f_n = \frac{\lambda^n \mathrm{e}^{-\lambda}}{n!}, \qquad n = 0, 1, 2, \ldots$$

has the characteristic function

$$\phi_X(t) = \sum_{n=0}^{\infty} f_n \mathrm{e}^{\mathrm{i}\,tn} = \sum_{n=0}^{\infty} \frac{\lambda^n \mathrm{e}^{-\lambda}}{n!} \mathrm{e}^{\mathrm{i}\,tn} = \mathrm{e}^{-\lambda} \sum_{n=0}^{\infty} \frac{\left(\lambda \mathrm{e}^{\mathrm{i}\,t}\right)^n}{n!} = \mathrm{e}^{-\lambda} \mathrm{e}^{\lambda \mathrm{e}^{\mathrm{i}\,t}} = \mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}\,t}-1)}.$$

Calculate also the corresponding moment-generating function! ◁

Example The standard normal distribution (3.10) has the characteristic function

$$\phi_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\,tx} \mathrm{e}^{-x^2/2}\,\mathrm{d}x = \mathrm{e}^{-t^2/2},$$

while the equivalent for the non-standardized normal distribution (3.7) is

$$\phi_X(t) = \mathrm{e}^{\mathrm{i}\,\mu t - \sigma^2 t^2/2}, \tag{B.20}$$

where we have used the formula in (see footnote 1). ◁

The following important properties of characteristic functions are given without proof. If a and b are constants and $Y = aX + b$, it holds that

$$\phi_Y(t) = \mathrm{e}^{\mathrm{i}bt}\phi_X(at), \tag{B.21}$$

which is also seen from (B.15). If random variables $X_1, X_2, \ldots, X_n$ are mutually independent and $Y = c_1X_1 + c_2X_2 + \cdots + c_nX_n$ is their linear combination, then

$$\phi_Y(t) = \phi_{X_1}(c_1t)\phi_{X_2}(c_2t)\cdots\phi_{X_n}(c_nt). \tag{B.22}$$

This theorem also can not be reversed: having $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$ does not necessarily mean that X and Y are independent.

As the moment-generating functions, the characteristic functions, too, can be used to derive the statistical moments $E[X^n]$, $n = 0, 1, 2, \ldots$ Namely, if ϕ_X is at least p-times continuously differentiable at the origin, it holds that

$$E\big[X^n\big] = \frac{1}{\mathrm{i}^n}\left[\frac{\mathrm{d}^n\phi_X}{\mathrm{d}t^n}\right]\bigg|_{t=0}, \qquad n = 1, 2, \ldots, p.$$

There is a one-to-one correspondence between the characteristic function and the probability distribution: any two random variables X and Y have the same probability distribution precisely when $\phi_X = \phi_Y$, therefore

$$\phi_X = \phi_Y \quad \Leftrightarrow \quad X \sim Y.$$

Example Let us calculate the characteristic function of the binomial distribution

$$f_n = P(X = n) = \binom{N}{n} p^n q^{N-n}, \qquad n = 0, 1, 2, \ldots, N.$$

Imagine a Bernoulli (binomial) sequence of trials. To the jth trial in this sequence we assign a random variable Y_j with value 1 for a "good" event A (probability p), or value 0 for the complementary event $\overline{A}$ (probability $q = 1 - p$). Since the trials in the sequence are mutually independent, the same applies to the random variables Y_j. The variable X takes the value n if there were n occurrences of A in N trials: in this case precisely n variables Y_j have value 1, while the others are zero, hence $X = Y_1 + Y_2 + \cdots + Y_N$. For an individual Y_j we then use (B.18) to calculate

$$\phi_{Y_j}(t) = \sum_{k=0}^{1} P(Y_j = y_k)\mathrm{e}^{\mathrm{i}ty_k} = \underbrace{P(Y_j = 0)}_{q}\,\mathrm{e}^{\mathrm{i}t0} + \underbrace{P(Y_j = 1)}_{p}\,\mathrm{e}^{\mathrm{i}t1} = p\,\mathrm{e}^{\mathrm{i}t} + q.$$

By (B.22) we then obtain the characteristic function of the binomial distribution

$$\phi_X(t) = \big(\phi_{Y_j}(t)\big)^N = \big(p\mathrm{e}^{\mathrm{i}t} + q\big)^N, \tag{B.23}$$

which we shall use in the following. ◁

B.3.1 Proof of Laplace's Limit Theorem

Characteristic functions allow us to show why the *discrete* binomial distribution at large N can be approximated by the *continuous* normal distribution, as claimed in Sect. 5.4. One starts with a sequence of binomially distributed random variables $\{X_N\}$ $(N = 1, 2, 3, \ldots)$ with the probability functions

$$P(X_N = n) = \binom{N}{n} p^n q^{N-n}, \qquad n = 0, 1, 2, \ldots, N, \qquad N = 1, 2, 3, \ldots$$

By (B.23) each such distribution possesses the characteristic function

$$\phi(t; N) = \left(p\mathrm{e}^{\mathrm{i}\,t} + q\right)^N.$$

We introduce standardized random variables

$$Y_N = \frac{X_N - E[X_N]}{\sqrt{\mathrm{var}[X_N]}} = \frac{X_N - Np}{\sqrt{Npq}}$$

and denote the characteristic function of each of them by $\widetilde{\phi}(t; N)$. By (B.21) we get

$$\widetilde{\phi}(t; N) = \left(p\,\mathrm{e}^{\mathrm{i}\,qt/\sqrt{Npq}} + q\,\mathrm{e}^{-\mathrm{i}\,pt/\sqrt{Npq}}\right)^N. \tag{B.24}$$

The terms in the brackets can be expanded in a power series:

$$p\,\mathrm{e}^{\mathrm{i}\,qt/\sqrt{Npq}} \approx p + \mathrm{i}\,t\sqrt{\frac{pq}{N}} - \frac{qt^2}{2N} + \mathcal{O}\big(t^2/N\big),$$
$$q\,\mathrm{e}^{-\mathrm{i}\,pt/\sqrt{Npq}} \approx q - \mathrm{i}\,t\sqrt{\frac{pq}{N}} - \frac{pt^2}{2N} + \mathcal{O}\big(t^2/N\big).$$

Here for each t one has $\lim_{N\to\infty} N\,\mathcal{O}\big(t^2/N\big) = 0$. When this is inserted in (B.24), we get

$$\widetilde{\phi}(t; N) = \left(1 - \frac{t^2}{2N} + \mathcal{O}\left(\frac{t^2}{N}\right)\right)^N \sim \mathrm{e}^{-t^2/2}, \quad \text{when } N \to \infty. \tag{B.25}$$

The limit of the sequence of characteristic functions $\widetilde{\phi}(t; N)$ is thus a continuous function, which is just the characteristic function of the standardized normal distribution. The aid to the final result is the theorem (given without proof): "*If the sequence of characteristic functions $\{\phi_n(t)\}$ at any real t converges to the function $\phi(t)$ and if ϕ is continuous on an arbitrary small interval $(-T, T)$, the sequence $\{F_n(x)\}$ of corresponding distribution functions converges to the distribution function $F(x)$, whose characteristic function is precisely $\phi(t)$.*" This means that for any x one has

$$\lim_{N\to\infty} P(Y_N \le x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du,$$

so, at large N (and any x) also

$$P(X_N \le x) \approx \frac{1}{\sqrt{2\pi Npq}} \int_{-\infty}^{x} e^{-(u-Np)^2/(2Npq)}\, du.$$

Put in a more practical form: if the experiment outcome A has a probability p ($0 < p < 1$, $q = 1 - p$) of occurring and X is its frequency in N trials of this experiment, then for arbitrary real numbers a and b ($a < b$) it holds that

$$\lim_{N\to\infty} P\left(a \le \frac{X - Np}{\sqrt{Npq}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-x^2/2}\, dx.$$

Now we understand why, at large N, the binomial distribution could be approximated by the normal distribution with the same mean and variance as possessed by the given binomial distribution. This realization is known as the *Laplace's limit theorem* (in its integral form).

By the same token the general central limit theorem can be derived that applies to any probability distribution, as long as its first and second moments exist. The tool is always the same: we power-expand the characteristic function and analyze its behaviour in the $N \to \infty$ limit, which always has the form (B.25).

B.3.2 Inverting the Characteristic Function and Uniqueness of the Density

The characteristic function—as well as its closest relative, the moment-generating function—uniquely determine the probability distribution. In other words, the probability distribution and the characteristic functions offer equivalent description of statistical properties of a random variable. Both worlds are linked by the Fourier transformation: the inverse of (B.18) is

$$f_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_X(t)\, e^{-itn}\, dt \qquad \text{(discrete case)},$$

while the inverse of (B.19) is

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_X(t)\, e^{-itx}\, dt \qquad \text{(continuous case)}.$$

But one must realize that the distribution is *not* necessarily uniquely determined if all its moments are known. A well-known case [7] are the probability densities

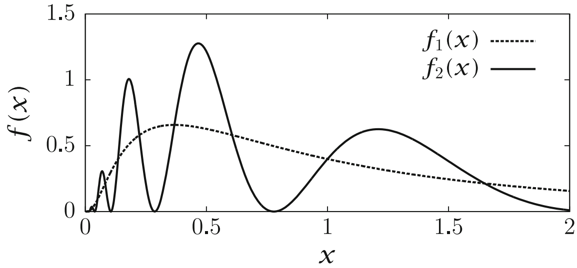


Fig. B.3 Example of different probability densities with identical moments. The function f_1 is the probability density of the *log-normal distribution:* is Y is a normally distributed continuous random variable and $X = \mathrm{e}^Y$, then X is log-normally distributed

$$f_1(x) = \frac{1}{\sqrt{2\pi x^2}}\,\mathrm{e}^{-(\log^2 x)/2}, \qquad x \geq 0,$$
$$f_2(x) = f_1(x)\big[1 + \sin(2\pi \log x)\big], \qquad x \geq 0, \tag{B.26}$$

which have very different functional dependencies (see Fig. B.3) yet identical moments, namely

$$E\big[X\big] = \sqrt{\mathrm{e}}, \quad E\big[X^2\big] = \mathrm{e}^2, \quad E\big[X^3\big] = \mathrm{e}^{9/2}, \ldots, E[X^n] = \mathrm{e}^{n^2/2}.$$

This is the so-called *indeterminate moment problem*, briefly outlined below.

If the variables X and Y have identical moments, their characteristic functions $\phi_X(t)$ and $\phi_Y(t)$ have identical expansions near the origin of the real axis. But having equal expansions does not say much about the equality of ϕ_X and ϕ_Y, as the expansion may not converge at all—the terms are calculable in principle, but they can not be summed: in the case just mentioned the convergence of the Taylor series of the characteristic function corresponding to the log-normal density (B.26),

$$\phi_X(t) = \sum_{n=0}^{\infty} a_n (\mathrm{i}\,t)^n, \qquad a_n = \frac{1}{n!}\,E\big[X^n\big],$$

is zero:

$$\rho = \left(\limsup_{n\to\infty} |a_n|^{1/n}\right)^{-1} = 0.$$

However, in specific cases the convergence *is* guaranteed (Theorem 9.6.2 in [8]): if one can find $\rho > 0$ such that near the origin, $|t| < \rho$, the expected value of $\mathrm{e}^{t|X|}$ is finite, i. e. $E[\mathrm{e}^{t|X|}] < \infty$, then $\phi_X(t)$ is absolutely convergent for $|t| < \rho$. Then one may conclude $E[\mathrm{e}^{\mathrm{i}tX}] = E[\mathrm{e}^{\mathrm{i}tY}] \Leftrightarrow X \sim Y$ or

$$\phi_X(t) = \phi_Y(t), \quad |t| < \rho \quad \Leftrightarrow \quad X \sim Y,$$

and under these conditions the equality of the moments of X and Y implies the equality of their probability distributions.

References

1. S. Širca, M. Horvat, *Computational Methods for Physicists* (Springer, Berlin, 2012)
2. J.K. Cavers, On the fast Fourier transform inversion of probability generating functions. J. Inst. Math. Appl. **22**, 275 (1978)
3. J. Abate, W. Whitt, Numerical inversion of probability generating functions. Oper. Res. Lett. **12**, 245 (1992)
4. J. Abate, W. Whitt, The Fourier-series method for inverting transforms of probability distributions. Queueing Syst. **10**, 5 (1992)
5. S. Kullback, An application of characteristic functions to the distribution problem of statistics. Ann. Math. Stat. **5**, 263 (1934)
6. E. Lukacs, Applications of characteristic functions in statistics. Indian J. Stat. A **25**, 175 (1963)
7. C. Berg, Indeterminate moment problem and the theory of entire functions. J. Comput. Appl. Math. **65**, 27 (1995).
8. T. Kawata, *Fourier Analysis in Probability Theory* (Academic Press, New York, 1972)

Appendix C
Random Number Generators

Abstract Methods of generating almost random numbers by means of computer algorithms are presented, starting from integer-based linear and non-linear congruential generators of uniformly distributed random numbers. They are followed by a discussion of methods to draw random numbers from arbitrary continuous distribution, and a brief mention of the ways to generate truly random numbers.

Statistical methods and numerical procedures often require us to use random samples or some kind of "source" of numbers that are as random as possible, that is, *pseudo-random*. The computer namely can not do anything "by chance", so in order to generate pseudo-random numbers we rely on *deterministic* processes of computing particular sequences that are only *seemingly* random [1]. Generating pseudo-random numbers—labeled simply 'random' in the following—is called *drawing*.

C.1 Uniformly Distributed Pseudo-Random Numbers

In order to generate uniformly distributed pseudo-random numbers one uses *uniform generators*. They are supposed to deliver uniform numbers $X \sim U(0, 1)$, distributed according to (3.1).

The sequences $\{x_i\}$ generated by a good uniform generator are expected to be *uncorrelated:* this means that the vectors of sub-sequences $(x_i, x_{i+1}, \ldots, x_{i+k})$ are as weakly correlated as possible, for each k separately. One also wishes for the sequence to possess a *long period:* it should not repeat itself too quickly. Besides, one would like the sequence $\{x_i\}$ to be *uniform and unbiased*, meaning that the same number of generated points fall in the same volume of space. An important request is a good uniformity of the distribution of the points $(x_i, x_{i+1}, \ldots, x_{i+k-1})$ in a k-dimensional hypercube, with k as large as possible: this is known as the *serial uniformity of the sequence*.

S. Širca, *Probability for Physicists*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-31611-6

Most uniform generators are devised in integer arithmetic. Such generators return numbers with equal probabilities on the interval $[0, m-1]$, where $m = 2^{32}$ or 2^{64}. Uniform generators are standard components of general libraries and tools, e.g. `rand()` in MATLAB and C/C++, `gsl_rng_rand` in GSL or `Random[]` in MATHEMATICA. Random integers $x_i \in \mathbb{Z}_m$ generated by an integer generator can be converted to uniformly distributed real numbers $\xi_i \sim U(0, 1)$ by using the transformations

$$\begin{aligned} \xi_i &= x_i/m && \text{approximately uniform in } [0, 1), \\ \xi_i &= x_i/(m-1) && [0, 1], \\ \xi_i &= (x_i+1)/m && (0, 1], \\ \xi_i &= (x_i+1/2)/m && (0, 1). \end{aligned}$$

If one uses floating-point arithmetic (precision 2^{-n}, mantissa length n), the numbers generated in this way have $b = \log_2 m$ random most significant bits, which is often not enough, and certainly less than n. An approximation of a real number ξ on the interval $[0, 1)$ with all bits random is obtained by independently drawing integers $\{x_i \in \mathbb{Z}_m\}_{i=1}^{h}$ and using the formula $\xi = x_1 m^{-1} + x_2 m^{-2} + \cdots + x_h m^{-h}$, where $(h-1)b < n < (h+1)b$.

C.1.1 Linear Congruential Generators

Classical random generators are based on the relation of congruence.[2] Congruential generators of numbers $x_i \in \mathbb{Z}_m = [0, m-1]$, where $i \in \mathbb{N}_0 = \{0, 1, \ldots\}$, are defined by the *transition function* ϕ and the relation

$$x_{i+1} \equiv \phi(x_i, x_{i-1}, \ldots, x_{i-k+1}) \mod m,$$

where k is the generator *order*. Thus ϕ is restricted to $\mathbb{Z}_m$ by the congruence relation modulo m. The initial state of the generator $\{x_0, x_1, \ldots, x_{k-1}\}$ is a unique function of the number called the *seed* by which the sequence is completely determined: a generator initialized by the same seed always delivers the same sequence of numbers. If ϕ is a linear function of parameters, one refers to *linear* generators, otherwise they are *non-linear*.

The simplest *linear congruential generator* (LCG) has the form

$$x_{i+1} \equiv (a x_i + c) \mod m, \tag{C.1}$$

[2]One has $\{x \equiv x \mod m\}$; the congruence relation is commutative, $\{x \equiv y \mod m\} \Leftrightarrow \{y \equiv x \mod m\}$, and transitive, $\{x \equiv y \mod m\} \wedge \{y \equiv z \mod m\} \Rightarrow \{x \equiv z \mod m\}$.

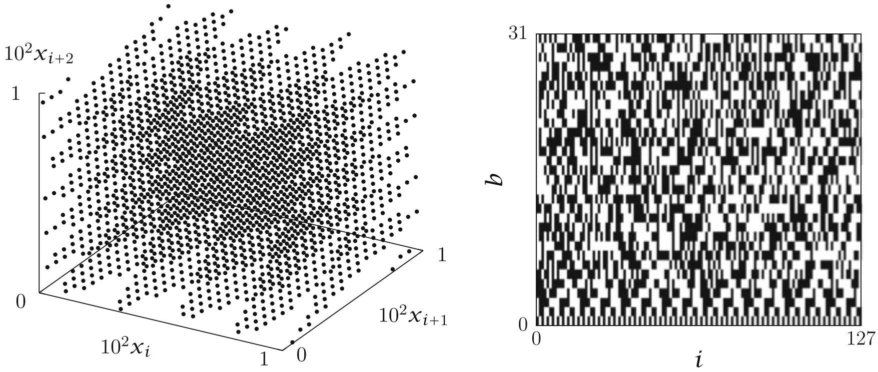


Fig. C.1 [Left] Zoom-in of the phase space $[0, 1]^3$ of points $2^{-31}(x_i, x_{i+1}, x_{i+2})$ of the sequence x_i obtained by the standard random generator from the `glibc` library with $x_0 = 12345$. [Right] The bits b of the random numbers x_i (*black* = 1, *white* = 0)

where a is the *multiplier* and c is the *carry* parameter, while x_0 is the seed. Since x_i are determined by x_{i-1} and can take only m different values, the period of a LCG is at most m for $c \neq 0$ and at most $m - 1$ for $c = 0$.

Example Take a LCG with $m = 31$, $a = 3$, $c = 0$ and $x_0 = 9$. Run the recurrence (C.1) a hundred times: we get $\{x_1, x_2, x_3, \ldots, x_{30}\} = \{27, 19, 26, \ldots, 9\}$, then again $\{x_{31}, x_{32}, x_{33}, \ldots, x_{60}\} = \{27, 19, 26, \ldots, 9\}$, and so on. The period of the generator is therefore only 30, but this is not the only problem. If subsequent pairs $(x_i, x_{i+1}) = (9, 27), (27, 19), (19, 26), \ldots$ are plotted on a graph—do it!—we realize that all points lie on straight lines with slope 3. That certainly does not appear to be random!

Is one better off by increasing m and a, and changing c? Take, for instance, $m = 2^{32}$, $a = 1103515245$ and $c = 123454$: this corresponds to the default generator in the 32-bit `glibc` library. Figure C.1 (left) shows the distribution of subsequent triplets (x_i, x_{i+1}, x_{i+2}). Obviously the points are arranged in planes and this deficiency of LCG persists at larger k as well: in general the points $m^{-1}(x_i, \ldots, x_{i+k-1})$ do not fill the entire k-dimensional hypercube, but rather lie on at most $(mk!)^{1/k}$ hyperplanes. Besides, the least significant bits are less random that the rest (Fig. C.1 (right)). A good generator ought to produce points on many hyperplanes and make all their bits random. In applications where such deficiencies are irrelevant, LCG-type generators are nevertheless put to good use, as they are supported by all programming languages, simple and fast. ◁

Further representatives of the LCG family are the generators of the Add-with-Carry (AWC), Subtract-with-Borrow (SWB) and Multiply-with-Carry (MWC) type:

$$\begin{aligned}
\text{AWC}: \; x_i &\equiv (x_{i-r} + x_{i-k} + c_{i-1}) \mod m, \quad c_i = \lfloor (x_{i-r} + x_{i-k} + c_{i-1})/m \rfloor, \\
\text{SWB}: \; x_i &\equiv (x_{i-r} - x_{i-k} - c_{i-1}) \mod m, \quad c_i = \lfloor (x_{i-r} - x_{i-k} - c_{i-1})/m \rfloor, \\
\text{MWC}: \; x_i &\equiv (a x_{i-r} + c_{i-1}) \mod m, \quad c_i = \lfloor (a x_{i-r} + c_{i-1})/m \rfloor.
\end{aligned}$$

The SWB algorithm is the basis of the `RANLUX` generator from the GSL library. *Multiple recursive generators* (MRG) are also in wide-spread use:

$$x_i \equiv (a_1 x_{i-1} + \cdots + a_k x_{i-k} + c_i) \mod m, \tag{C.2}$$

where $a_k \in \mathbb{Z}_m$ are constants. The MR generators usually exhibit much larger periods than simple LC generators. If m is a prime, the maximal period may be as high as $m^k - 1$. An example of such a generator of the fifth order is

$$x_i \equiv \left(107374182\, x_{i-1} + 104480\, x_{i-5}\right) \mod \left(2^{31} - 1\right).$$

C.1.2 Non-linear Congruential Generators

In general, non-linear generators are more random than linear ones, but they are also slower. Their main representatives are the *inversive congruential generators* (ICG) defined by the recurrence

$$x_i \equiv (a\overline{x}_{i-1} + b) \mod m,$$

where $1 \equiv (\overline{x}x) \mod m$, and the *explicit inversive congruential generators* (EICG) based on the relation

$$x_i \equiv \overline{a(i + i_0) + b} \mod m.$$

For prime modules m the generators of IC and EIC types generate points that avoid accumulation in planes, a behavior so typical of the LC generators, yet modular inversion is a numerically intensive procedure, while the filling of space tends to be slightly less uniform.

C.1.3 Generators Based on Bit Shifts

A completely different approach to generating random numbers is offered by *feedback shift register* generators. If the numbers x_i are written as n-plets of bits, the relation (C.2) can be written as

$$b_i \equiv (a_p b_{i-p} + a_{p-1} b_{i-p+1} + \cdots + a_1 b_{i-1}) \mod 2, \tag{C.3}$$

where all variables can take only values 0 or 1. It turns out that the recurrence (C.3) can be performed by shifting bits: an example is shown in Fig. C.2.

The evicted bit is then combined by the pattern of bits on its right by using a variety of logical operations. The recurrence (C.3) often has the form $b_i \equiv (b_{i-p} + b_{i-p+q}) \mod 2$ or $b_i \equiv b_{i-p} \oplus b_{i-p+q}$, where $\oplus$ is the exclusive "or" (adding 0 and 1 modulo

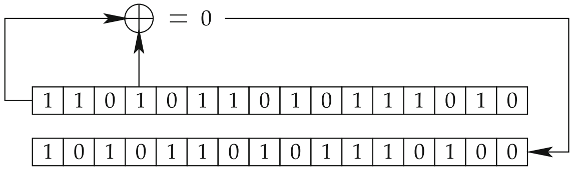


Fig. C.2 Example of bit shifts in a FSR-type generator. The pushed-out bit 1 at the extreme left and the bit 1 deeper in the register are combined by an exclusive "or" (XOR). The result 0 replaces the missing bit at the extreme right

2). For n-tuples x_i this means $x_i = x_{i-p} \oplus x_{i-p+q}$, where the operation $\oplus$ is performed bit-wise. This game can be continued: if x_i are interpreted as n-dimensional vectors, they can be multiplied by $n \times n$ matrices:

$$x_i \equiv x_{i-p} \oplus Ax_{i-p+q} .$$

Where is all this heading? The matrix A can be used to *twist* the bit n-tuples prior to being logically combined, thereby increasing the randomness of the generated x_i. Such "kneading" of bit samples is at the heart of the *Mersenne twister* generator [2] (algorithm MT19937), which we recommend for serious applications. It is implemented in 32-bit integer arithmetic, has been theoretically well explored and is accessible in standard packages and libraries. Its period is $2^{19937} - 1$ and is serially uniform for dimensions $k \in [1, 623]$. Its weakness is a somewhat lower randomness of subsequent bits between consecutive generated numbers.

C.1.4 Some Hints for Use of Random Generators

Any random number generator, no matter how sophisticated, has some deficiency, which is usually very specific. If we, as non-specialists, need a generator to be invoked many times in our code, we may consider the following guidelines.

Only choose a generator devised and tested by experts. The code should be as terse as possible and based on integer arithmetic in favor of greater speed. Use generators with long periods and high serial uniformity in as many dimensions as possible. If the generator is accessible in source code, incorporate it into the program, as modern compilers can link the code segments in the form of `inline` functions. Before using a generator, study its statistical properties and ascertain whether its deficiencies may jeopardize the correctness of the result. Perform each calculation by using different generators and different seeds.

C.2 Drawing from Arbitrary Continuous Distributions

A generator of random numbers with arbitrary distribution is obtained by transforming the numbers returned by a uniform generator. An exhaustive overview of the area is offered by the classical monograph [3]; here we present some cases of transformations to continuous distributions most commonly encountered in physics. For transformation to discrete distributions see Sect. C.2 in [4].

C.2.1 Uniform Distribution Within a Circle or Sphere

How do we draw points that are homogeneously distributed within a circle? Homogeneity means: the ratio of a tiny probability $\mathrm{d}P$ that the drawn point falls in a tiny surface element, to its area, $\mathrm{d}S = r\,\mathrm{d}r\,\mathrm{d}\phi$, is equal to the ratio of probability 1 that the point falls in the whole circle, to its area, πR^2:

$$\frac{\mathrm{d}P}{\mathrm{d}S} = \frac{\mathrm{d}P}{r\,\mathrm{d}r\,\mathrm{d}\phi} = \frac{1}{\pi R^2} \quad \stackrel{R=1}{\Longrightarrow} \quad \frac{\mathrm{d}P}{\mathrm{d}\big(r^2\big)\mathrm{d}\big(\phi/2\pi\big)} = 1.$$

Therefore we must draw *uniformly* in r^2 from 0 to 1 (not r from 0 to 1!) and in ϕ from 0 to 2π. We need two random numbers $U_1, U_2 \sim U[0, 1)$ and compute

$$\big(r_i, \phi_i\big) = \big(R\sqrt{U_1},\ 2\pi U_2\big).$$

In the three-dimensional case the circle area $S = \pi R^2$ needs to be replaced by the sphere volume $V = 4\pi R^3/3$, and the area element $\mathrm{d}S$ by the volume element $\mathrm{d}V = r^2\,\mathrm{d}r\,\mathrm{d}(\cos\theta)\,\mathrm{d}\phi$. Thus

$$\frac{\mathrm{d}P}{\mathrm{d}V} = \frac{\mathrm{d}P}{r^2\,\mathrm{d}r\,\mathrm{d}(\cos\theta)\,\mathrm{d}\phi} = \frac{3}{4\pi R^3} \quad \stackrel{R=1}{\Longrightarrow} \quad \frac{\mathrm{d}P}{\mathrm{d}\big(r^3\big)\mathrm{d}\big(\frac{1}{2}\cos\theta\big)\mathrm{d}\big(\phi/2\pi\big)} = 1.$$

Hence we must draw uniformly in r^3 from 0 to 1, uniformly in $\cos\theta$ from -1 to 1 (not θ from 0 to π!) and uniformly in ϕ from 0 to 2π. The three numbers $U_1, U_2, U_3 \sim U[0, 1)$ drawn according to these distributions define the point

$$\big(r_i,\ \theta_i,\ \phi_i\big) = \big(R\sqrt[3]{U_1},\ \arccos(2U_2 - 1),\ 2\pi U_3\big). \tag{C.4}$$

C.2.2 Uniform Distribution with Respect to Directions in $\mathbb{R}^3$ and $\mathbb{R}^d$

A uniform distribution over *directions* in space (usually $\mathbb{R}^3$) is called isotropic. Isotropy means that the ratio between the number of points $\mathrm{d}N$ on the small surface $\mathrm{d}S$ on the unit sphere to an infinitesimal solid angle $\mathrm{d}\Omega$, is equal to the ratio of the number of points N on the whole surface to the full solid angle $\Omega = 4\pi$. A frequent beginner's mistake is to uniformly draw the angles θ and ϕ according to $U(0, \pi)$ and $U(0, 2\pi)$, respectively, and compute $(x, y, z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. But this generates points that prefer to accumulate near the poles, as shown in Fig. C.3 (left). The correct way to draw is by recipe (C.4), where the radial coordinate is simply ignored. This results in a homogeneous surface distribution, as shown in Fig. C.3 (right).

The points $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)^\mathrm{T} \in \mathbb{R}^d$, uniformly distributed over the $(d-1)$-dimensional sphere $\boldsymbol{S}_{d-1} \in \mathbb{R}^d$, can be generated by independently drawing the components of the vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_d)^\mathrm{T}$ with probability density $N(0, 1)$ and normalizing it: $x_i = y_i / \|\boldsymbol{y}\|_2$, where $\|\boldsymbol{y}\|_2 = \left(\sum_{i=1}^d y_i^2\right)^{1/2}$.

C.2.3 Uniform Distribution Over a Hyperplane

The points $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)^\mathrm{T}$, $x_i > 0$, uniformly distributed over a hyperplane defined by the equation $\sum_{i=1}^d a_i x_i = b$ ($a_i > 0$, $b > 0$), are generated by independently drawing d components of the vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_d)^\mathrm{T}$ with exponential density $f(y) = \exp(-y)$ (see Table C.1) and calculating [5]

$$S = \sum_{i=1}^d a_i y_i, \qquad x_i = \frac{b}{S}\, a_i y_i.$$

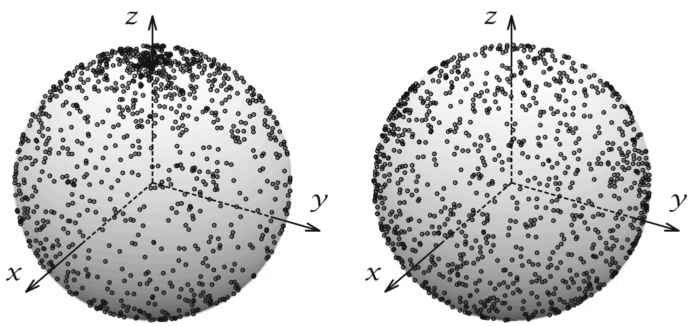


Fig. C.3 Generating an isotropic distribution in $\mathbb{R}^3$. [Left] Incorrect drawing by using $\theta_i = \pi\xi$, $\xi \sim U[0, 1)$. [Right] Correct drawing by using $\theta_i = \arccos(2\xi - 1)$

C.2.4 Transformation (Inverse) Method

Our knowledge of variable transformations from Sects. 2.7 and 2.10 can be used to generate random numbers according to an arbitrary continuous distribution. We know how uniform numbers $Y \sim U(0,1)$ can be generated; but as for arbitrary probability densities f_X and f_Y one has $|f_X(x)\,\mathrm{d}x| = |f_Y(y)\,\mathrm{d}y|$, this means that

$$f_X(x) = \frac{\mathrm{d}y}{\mathrm{d}x},$$

since $f_Y(y) = 1$. The solution of this equation is $y = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = F_X(x)$, where F_X is the distribution function of X. In other words,

$$x = F_X^{-1}(y), \qquad Y \sim U(0,1),$$

where F_X^{-1} is the *inverse* function of F_X (not its reciprocal value). Clearly we have obtained a tool to generate random variables distributed according to F_X (see Fig. C.4 (left)).

The transformation method is useful if the inverse F_X^{-1} is relatively easy to compute. The collection of such functions is quickly exhausted; some common examples are listed in Table C.1.

Example Let us construct a generator of dipole electro-magnetic radiation! The distribution of radiated power with respect to the solid angle is $\mathrm{d}P/\mathrm{d}\Omega \propto \sin^2\theta$,

$$f_\Theta(\theta) = \frac{\mathrm{d}P}{\mathrm{d}\theta} = \frac{3}{4}\sin^3\theta, \qquad 0 \le \theta \le \pi,$$

where the normalization constant has been determined by $C\int_0^\pi \sin^3\theta\,\mathrm{d}\theta = 1$. (The radiation is uniform in ϕ.) The corresponding distribution function is

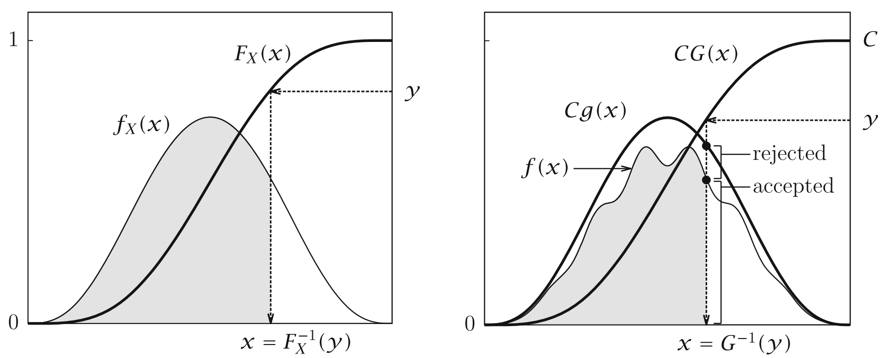


Fig. C.4 Generating random numbers according to arbitrary continuous distributions. [Left] Transformation (inverse of distribution function) method. [Right] Rejection method

Table C.1 Generating random numbers according to chosen probability distributions by the transformation method

Distribution	$f_X(x)$	$F_X(x)$	$X = F_X^{-1}(U)$
Exponential $(x \geq 0)$	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	$-\frac{1}{\lambda}\log U$
Normal $(-\infty < x < \infty)$	$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	$\frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$	$\sqrt{2}\mathrm{erf}^{-1}(2U - 1)$
Cauchy $(-\infty < x < \infty)$	$\frac{a}{\pi(a^2 + x^2)}$	$\frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x}{a}\right)$	$a \tan \pi U$
Pareto $(0 < b \leq x)$	$\frac{ab^a}{x^{a+1}}$	$1 - \left(\frac{b}{x}\right)^a$	$\frac{b}{U^{1/a}}$
Triangular on $[0, a]$ $(0 \leq x \leq a)$	$\frac{2}{a}\left(1 - \frac{x}{a}\right)$	$\frac{2}{a}\left(x - \frac{x^2}{2a}\right)$	$a\left(1 - \sqrt{U}\right)$
Rayleigh $(x \geq 0)$	$\frac{x}{\sigma} e^{-x^2/(2\sigma^2)}$	$1 - e^{-x^2/(2\sigma^2)}$	$\sigma\sqrt{-\log U}$

Note that drawing Y by the uniform distribution $U(0, 1)$ is equivalent to drawing by $1 - U(0, 1)$. For the normal distribution see also Sect. C.2.5

$$F_\Theta(\theta) = \int_0^\theta f_\Theta(\theta')\,d\theta' = \frac{3}{4}\left[\frac{\cos^3\theta}{3} - \cos\theta + \frac{2}{3}\right].$$

The desired distribution in θ is obtained by drawing the values x according to $U(0, 1)$ and calculating $\theta = F_\Theta^{-1}(x)$. The inverse of F_Θ is annoying but can be done. By substituting $t = \cos\theta$ the problem amounts to solving the cubic equation $t^3 - 3t + 2 = 4x$, for which explicit formulas exist. Alternatively, one can seek the solution of the equation $\mathcal{F}(\theta) = F_\Theta(\theta) - x = 0$. ◁

C.2.5 Normally Distributed Random Numbers

If U_1 and U_2 are independent random variables, distributed as $U_1 \sim U(0, 1]$ and $U_2 \sim [0, 1)$, their Box-Muller transformation [6]

$$X_1 = \sqrt{-2\log U_1}\,\cos(2\pi U_2), \qquad X_2 = \sqrt{-2\log U_1}\,\sin(2\pi U_2),$$

yields independent random variables X_1 and X_2, distributed according to the standard normal distribution $N(0, 1)$. The variables U_1 and U_2 define the length $R = \sqrt{-2\log U_1}$ and the directional angle $\theta = 2\pi U_2$ of a planar vector $(X_1, X_2)^T$. The numerically intensive calculation of trigonometric functions can be avoided by using Marsaglia's implementation (see [7], Chap. 7, Algorithm P):

repeat
  Independently draw u_1 by $U(0, 1]$ and u_2 by $U[0, 1)$;
  $\boldsymbol{v} = 2(u_1, u_2)^{\mathrm{T}} - (1, 1)^{\mathrm{T}}$;
  $s = |\boldsymbol{v}|^2$;
until $(s \geq 1 \vee s \neq 0)$;
$(x_1, x_2)^{\mathrm{T}} = \sqrt{-(2/s) \log s}\, \boldsymbol{v}$;

The drawn vector $\boldsymbol{v}$ on average uniformly covers the unit circle, while approximately $1-\pi/4 \approx 21.5\%$ generated points are rejected, so that for one pair (x_1, x_2) one needs to draw $2/(\pi/4) \approx 2.54$ uniform numbers.

Values of the random vector $\boldsymbol{X} \in \mathbb{R}^d$, distributed according to the multivariate probability density (4.23) with mean $\boldsymbol{\mu}$ and correlation matrix Σ are generated by independently drawing d components of the vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_d)^{\mathrm{T}}$ by the standardized normal distribution $N(0, 1)$ and computing

$$\boldsymbol{x} = L\boldsymbol{y} + \boldsymbol{\mu},$$

where L is the lower-triangular $d \times d$ matrix from the Cholesky decomposition of the correlation matrix, $\Sigma = LL^{\mathrm{T}}$.

C.2.6 Rejection Method

Suppose we wish to draw random numbers according to some complicated density f, while some very efficient way is at hand to generate the numbers according to another, simpler density g. We first try to find $C > 1$ such that f is bounded by Cg from above as tightly as possible (Fig. C.4 (right)), that is, to ensure $f(x) < Cg(x)$ for all x with C as close to 1 as possible. Then the random numbers Y distributed according to f can be generated by the procedure:

1. Generate the value x of random variable X according to density g.
2. Generate the value u of random variable U according to $U(0, 1)$.
3. If $u \leq f(x)/(Cg(x))$, assign $y = x$ (x is "accepted"), otherwise return to step 1 (x is "rejected").

Does this recipe really do what it is supposed to do? Let us define the event $B = \{U \leq f(X)/\big(Cg(X)\big)\}$. From the given recipe and the Figure it is clear that

$$P\big(B \mid X = x\big) = P\left(U \leq \frac{f(X)}{Cg(X)} \middle| X = x\right) = \frac{f(x)}{Cg(x)},$$

hence

$$P(B) = \int_{-\infty}^{\infty} P\big(B \mid X = x\big)\, g(x)\, \mathrm{d}x = \int_{-\infty}^{\infty} \frac{f(x)}{Cg(x)}\, g(x)\, \mathrm{d}x = \frac{1}{C} \int_{-\infty}^{\infty} f(x)\, \mathrm{d}x = \frac{1}{C}.$$

Now define the event $A = \{X \le x\}$. We must prove that the conditional distribution function for X, given condition B, is indeed F, that is, we must check

$$P(A|B) = P\left(X \le x \,\middle|\, U \le \frac{f(X)}{Cg(X)}\right) \stackrel{?}{=} F(x).$$

For this purpose we first calculate $P(B|A)$, where we exploit the definition of conditional probability (1.10) in the form $P(B|A) = P(AB)/P(A)$,

$$\begin{aligned}
P(B|A) = P\left(U \le \frac{f(X)}{Cg(X)}\,\middle|\, X \le x\right) &= \frac{P\big(U \le f(X)/\big(Cg(X)\big) \cap X \le x\big)}{P(X \le x)} \\
&= \int_{-\infty}^{x} \frac{P\big(U \le f(X)/\big(Cg(X)\big)\big|\, X = z \le x\big)}{P(X \le x)}\, g(z)\,\mathrm{d}z \\
&= \frac{1}{G(x)} \int_{-\infty}^{x} \frac{f(z)}{Cg(z)}\, g(z)\,\mathrm{d}z = \frac{1}{CG(x)} \int_{-\infty}^{x} f(z)\,\mathrm{d}z = \frac{F(x)}{CG(x)},
\end{aligned}$$

and then invoke the product formula (1.10) for the final result

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{F(x)}{CG(x)} \frac{G(x)}{1/C} = F(x).$$

Example For the Cauchy distribution with probability density (3.18) the distribution function and its inverse are easy to calculate:

$$F_X(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x, \qquad F_X^{-1}(t) = \tan\left[\pi\left(t - \frac{1}{2}\right)\right]. \tag{C.5}$$

To generate the values of a Cauchy-distributed variable X one could therefore resort to the transformation method by using in (C.5) a random variable U, uniformly distributed over $[-1/2, 1/2]$—or, due to symmetry, over $[0, 1]$—and calculating $X = \tan \pi U$ (third row of Table C.1). But since computing the tangent is slow, it is better to seek the values of X as the ratios between the projections of the points within the circle onto x and y axes. These points are uniformly distributed with respect to the angles. We use the algorithm

repeat
| Draw u_1 according to $U(-1, 1)$ and u_2 according to $U(0, 1)$.
until *(* $u_1^2 + u_2^2 > 1 \vee u_2 = 0$ *)*;
$x = u_1/u_2$;

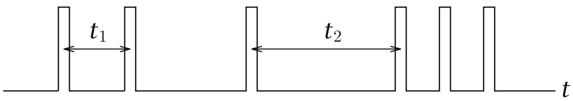
Note that the fraction of rejected points is $1 - \pi/4$ and that the accepted points (u_1, u_2) lie in the upper half of the unit circle. (Check this!) ◁

C.3 Generating *Truly* Random Numbers

If we wish to cast off the burden of the ‘pseudo’ attribute in our discussion and generate truly random numbers, we must also reach for a genuinely random process. An example of such process is the radioactive decay of atomic nuclei, which is exploited by the `HotBits` generator of random bit sequences [8]. The laboratory maintains a sample of radioactive cesium, decaying to an excited state of barium, electron and anti-neutrino with a decay time of 30.17 years:

$$^{137}\mathrm{Cs} \longrightarrow {}^{137}\mathrm{Ba}^* + \mathrm{e}^- + \overline{\nu}_\mathrm{e}.$$

The decay instant is defined by the detected electron. The time of the decay of any nucleus in the source is completely random, so the time difference between subsequent decays is also completely random. The apparatus measures the time differences between two *pairs* of decays, t_1 and t_2, as shown in the figure.



If $t_1 = t_2$ (within instrumental resolution), the measurement is discarded. If $t_1 < t_2$, the value 0 is recorded, and if $t_1 > t_2$, the value 1 is recorded. The sense of comparing t_1 to t_2 is reversed with each subsequent pair in order to avoid systematic errors in the apparatus or in the measurement that could bias one outcome against the other. The final result is a random bit sequence like

```
1111011100100001101110100010110001001100110110011100111100000001
0100001010011111110010111011110011010011011100001000101100011111 ...
```

The speed of generation depends on the activity of the radioactive source.

Example Imagine a descent along a binary tree (Fig. C.5) where each branch point represents a random step to the left ($n_i = 1$) with probability p or to the right ($n_i = 0$) with probability $1-p$. (The left-right decision can be made, for example, by “asking” the radioactive source discussed above.) The values n_i corresponding to the traversed branches are arranged in a k-digit binary number $B_k = \left(n_{k-1}n_{k-2}\ldots n_1 n_0\right)_2$ and suitably normalized,

$$X_k = N_k B_k = N_k \sum_{i=0}^{k-1} 2^i n_i, \qquad N_k = \left(2^k - 1\right)^{-1},$$

so that we ultimately end up with $0 \le X_k < 1$. What is the expected value of X_k in the decimal system (base 10)? The individual digits n_i take the values 0 or 1 with probabilities $P_i = p\delta_{i,1} + (1-p)\delta_{i,0}$. Obviously $E[n_i] = p$, hence

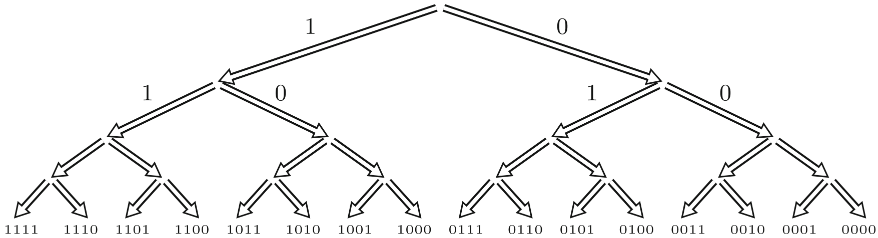


Fig. C.5 Binary tree used to generate a random k-digit binary number

$$E\big[X_k\big] = E\left[N_k \sum_{i=0}^{k-1} 2^i n_i\right] = N_k E[n_i] \sum_{i=0}^{k-1} 2^i = N_k p\left(2^k - 1\right) = p.$$

The variance of X_k is

$$\begin{aligned}\mathrm{var}[X_k] &= E\big[X_k^2\big] - E\big[X_k\big]^2 = N_k^2 \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} 2^{i+j}\left(\underbrace{E[n_i n_j]}_{p\delta_{i,j}} - \underbrace{E[n_i]E[n_j]}_{p^2\delta_{i,j}}\right) \\ &= N_k^2 p(1-p)\sum_{i=0}^{k-1} 4^i = N_k^2 p(1-p)\,\frac{4^k-1}{3} = \frac{p(1-p)}{3}\,\frac{2^k+1}{2^k-1}.\end{aligned}$$

We have thus devised a generator of *truly random* numbers, distributed according to $U[0,1)$. In particular, for $p = 1/2$ one indeed has $E[X_k] = 1/2$, while $\lim_{k\to\infty} \mathrm{var}\big[X_k\big] = 1/12$, as expected of a uniform distribution. ◁

References

1. P. L'Ecuyer, Uniform random number generators, in *Non-uniform random variate generation*, *International Encyclopedia of Statistical Science*, ed. by M. Lovric (Springer, Berlin, 2011)
2. M. Matsumoto, T. Nishimura, Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans. Model. Comput. Simul. **8**, 3 (1998)
3. L. Devroye, *Non-uniform Random Variate Generation* (Springer, Berlin, 1986)
4. S. Širca, M. Horvat, *Computational Methods for Physicists* (Springer, Berlin, 2012)
5. M. Horvat, The ensemble of random Markov matrices. J. Stat. Mech. **2009**, P07005 (2009)
6. G.E.P. Box, M.E. Muller, A note on the generation of random normal deviates. Ann. Math. Stat. **29**, 610 (1958)
7. D. Knuth, *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*, 3rd edn. (Addison-Wesley Professional, Reading 1998)
8. J. Walker, Hotbits; see http://www.fourmilab.ch/hotbits

Appendix D
Tables of Distribution Quantiles

Abstract Definite integrals of the normal distribution are given in tabular form, along with the most frequently used quantiles of the χ^2, t and F distributions.

Definite integrals of some distributions have awkward analytic expressions, so one may prefer to read them off from tables. Table D.1 lists the integrals of the standardized normal distribution (Fig. D.1 (top left)), Table D.2 contains the values of the erf function, and Table D.3 has the quantiles χ^2_p of the χ^2 distribution with ν degrees of freedom (Fig. D.1 (top right)). Table D.4 lists the quantiles t_p of the Student's t distribution with ν degrees of freedom (Fig. D.1 (bottom left)). Tables D.5 and D.6 contain the 95. percentile ($F_{0.95}$) and 99. percentile ($F_{0.99}$), respectively, of the F distribution with ν_1 and ν_2 degrees of freedom in the numerator and denominator, respectively (Fig. D.1 (bottom right)).

Note that the integral of the standardized normal distribution (Table D.1) and the value of the erf function (Table D.2) are related by

$$\frac{1}{\sqrt{2\pi}} \int_0^z \mathrm{e}^{-t^2/2} \mathrm{d}t = \frac{1}{2} \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right).$$

The distribution function of the standardized normal distribution is

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \mathrm{e}^{-t^2/2}\, \mathrm{d}t = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z \mathrm{e}^{-t^2/2}\, \mathrm{d}t = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)\right].$$

D.1 Calculating Quantiles with Mathematica

Arbitrary quantiles not given in the following tables can be calculated by interpolation or by resorting to a general tool like Mathematica [1]. For example, to obtain the 90. percentile of the χ^2 distribution with $\nu = 5$ degrees of freedom, the

S. Širca, *Probability for Physicists*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-31611-6

Table D.1 Integral of the standardized normal distribution (3.10) and (3.12) from 0 to z in steps of 0.01

z	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

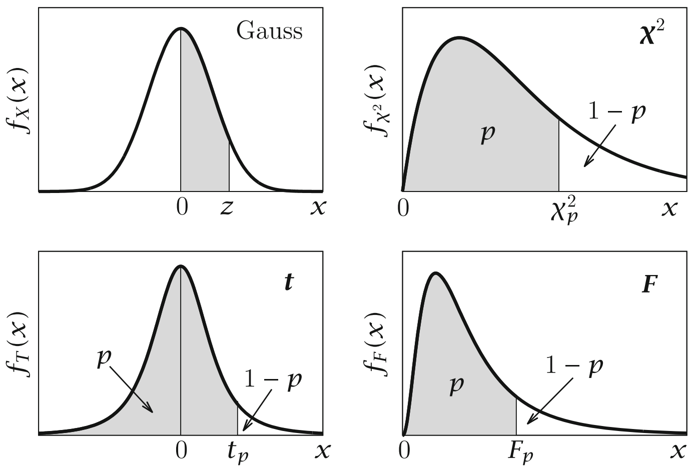


Fig. D.1 [Top left] Definite integral of the standard normal distribution (3.10) and (3.12) from 0 to z. [Top right] Definition of the pth quantile of the χ^2 distribution (3.21). [Bottom left] Definition of the pth quantile of the t distribution (3.22). [Bottom right] Definition of the pth quantile of the F distribution (3.23)

0.995th quantile of the Student's t distribution with $\nu = 1$ degree of freedom and the 95. percentile of the F distribution for $\nu_1 = \nu_2 = 10$ we issue the commands

```
Quantile[ChiSquareDistribution[5], 0.90],
Quantile[StudentTDistribution[1], 0.995],
Quantile[FRatioDistribution[10,10], 0.95],
```

which give (in the same order as above)

```
9.23636,
63.6567,
2.97824.
```

(Compare these values to entries in the corresponding Tables.) Definite integrals of all mentioned distributions can be obtained by commands of the form

```
NIntegrate[PDF[FRatioDistribution[7,9], x], {x, 0, 3.293}],
NIntegrate[PDF[FRatioDistribution[7,9], x], {x, 0, 3.70}],
NIntegrate[PDF[FRatioDistribution[7,9], x], {x, 0, 5.613}],
NIntegrate[PDF[FRatioDistribution[9,7], x], {x, 0, 1./3.70}].
```

Here we have only demonstrated a sample calculation of integrating the density of the F distribution with parameters required by the Example on p. 187: the four command lines listed above yield the values

Table D.2 Values of the erf function (3.8) from 0 to z in steps of 0.01

z	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0113	0.0226	0.0338	0.0451	0.0564	0.0676	0.0789	0.0901	0.1013
0.1	0.1125	0.1236	0.1348	0.1459	0.1569	0.1680	0.1790	0.1900	0.2009	0.2118
0.2	0.2227	0.2335	0.2443	0.2550	0.2657	0.2763	0.2869	0.2974	0.3079	0.3183
0.3	0.3286	0.3389	0.3491	0.3593	0.3694	0.3794	0.3893	0.3992	0.4090	0.4187
0.4	0.4284	0.4380	0.4475	0.4569	0.4662	0.4755	0.4847	0.4937	0.5027	0.5117
0.5	0.5205	0.5292	0.5379	0.5465	0.5549	0.5633	0.5716	0.5798	0.5879	0.5959
0.6	0.6039	0.6117	0.6194	0.6270	0.6346	0.6420	0.6494	0.6566	0.6638	0.6708
0.7	0.6778	0.6847	0.6914	0.6981	0.7047	0.7112	0.7175	0.7238	0.7300	0.7361
0.8	0.7421	0.7480	0.7538	0.7595	0.7651	0.7707	0.7761	0.7814	0.7867	0.7918
0.9	0.7969	0.8019	0.8068	0.8116	0.8163	0.8209	0.8254	0.8299	0.8342	0.8385
1.0	0.8427	0.8468	0.8508	0.8548	0.8586	0.8624	0.8661	0.8698	0.8733	0.8768
1.1	0.8802	0.8835	0.8868	0.8900	0.8931	0.8961	0.8991	0.9020	0.9048	0.9076
1.2	0.9103	0.9130	0.9155	0.9181	0.9205	0.9229	0.9252	0.9275	0.9297	0.9319
1.3	0.9340	0.9361	0.9381	0.9400	0.9419	0.9438	0.9456	0.9473	0.9490	0.9507
1.4	0.9523	0.9539	0.9554	0.9569	0.9583	0.9597	0.9611	0.9624	0.9637	0.9649
1.5	0.9661	0.9673	0.9684	0.9695	0.9706	0.9716	0.9726	0.9736	0.9745	0.9755
1.6	0.9763	0.9772	0.9780	0.9788	0.9796	0.9804	0.9811	0.9818	0.9825	0.9832
1.7	0.9838	0.9844	0.9850	0.9856	0.9861	0.9867	0.9872	0.9877	0.9882	0.9886
1.8	0.9891	0.9895	0.9899	0.9903	0.9907	0.9911	0.9915	0.9918	0.9922	0.9925
1.9	0.9928	0.9931	0.9934	0.9937	0.9939	0.9942	0.9944	0.9947	0.9949	0.9951
2.0	0.9953	0.9955	0.9957	0.9959	0.9961	0.9963	0.9964	0.9966	0.9967	0.9969
2.1	0.9970	0.9972	0.9973	0.9974	0.9975	0.9976	0.9977	0.9979	0.9980	0.9980
2.2	0.9981	0.9982	0.9983	0.9984	0.9985	0.9985	0.9986	0.9987	0.9987	0.9988
2.3	0.9989	0.9989	0.9990	0.9990	0.9991	0.9991	0.9992	0.9992	0.9992	0.9993
2.4	0.9993	0.9993	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995	0.9995	0.9996
2.5	0.9996	0.9996	0.9996	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
2.6	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9999
2.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
2.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	1.000	1.000	1.000
2.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

```
0.95001,
0.96385,
0.99001,
0.03615 = 1-0.96385.
```

The calculation for other distributions proceeds along the same lines.

Table D.3 Quantiles χ^2_p of the χ^2 distribution (3.21) with ν degrees of freedom for some typical (most commonly used) values of p from 0.005 to 0.999

ν	$\chi^2_{.005}$	$\chi^2_{.01}$	$\chi^2_{.025}$	$\chi^2_{.05}$	$\chi^2_{.1}$	$\chi^2_{.25}$	$\chi^2_{.5}$	$\chi^2_{.75}$	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.975}$	$\chi^2_{.99}$	$\chi^2_{.995}$	$\chi^2_{.999}$
1	0.000	0.0002	0.0010	0.0039	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88	10.8
2	0.010	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6	13.8
3	0.072	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8	16.3
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9	18.5
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7	20.5
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5	22.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3	24.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0	26.1
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6	27.9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2	29.6
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8	31.3
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3	32.9
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8	34.5
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3	36.1
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8	37.7
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3	39.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7	40.8
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2	42.3
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6	43.8
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0	45.3
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4	46.8
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8	48.3

(continued)

Table D.3 (continued)

ν	$\chi^2_{.005}$	$\chi^2_{.01}$	$\chi^2_{.025}$	$\chi^2_{.05}$	$\chi^2_{.1}$	$\chi^2_{.25}$	$\chi^2_{.5}$	$\chi^2_{.75}$	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.975}$	$\chi^2_{.99}$	$\chi^2_{.995}$	$\chi^2_{.999}$
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2	49.7
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6	51.2
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9	52.6
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	54.1
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	55.5
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	56.9
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	58.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	59.7
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8	73.4
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5	86.7
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0	99.6
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100	104	112
80	51.2	53.5	57.2	60.4	64.3	71.1	79.3	88.1	96.6	102	107	112	116	125
90	59.2	61.8	65.6	69.1	73.3	80.6	89.3	98.6	108	113	118	124	128	137
100	67.3	70.1	74.2	77.9	82.4	90.1	99.3	109	118	124	130	136	140	149

Table D.4 Quantiles t_p of the Student's t distribution (3.22) with ν degrees of freedom for some typical (most commonly used) values of p from 0.55 to 0.999

ν	$t_{0.55}$	$t_{0.60}$	$t_{0.70}$	$t_{0.75}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$	$t_{0.999}$
1	0.158	0.325	0.727	1.000	1.376	3.08	6.31	12.7	31.8	63.7	3183
2	0.142	0.289	0.617	0.816	1.061	1.89	2.92	4.30	6.96	9.92	70.7
3	0.137	0.277	0.584	0.765	0.978	1.64	2.35	3.18	4.54	5.84	22.2
4	0.134	0.271	0.569	0.741	0.941	1.53	2.13	2.78	3.75	4.60	13.0
5	0.132	0.267	0.559	0.727	0.920	1.48	2.02	2.57	3.36	4.03	9.68
6	0.131	0.265	0.553	0.718	0.906	1.44	1.94	2.45	3.14	3.71	8.02
7	0.130	0.263	0.549	0.711	0.896	1.41	1.89	2.36	3.00	3.50	7.06
8	0.130	0.262	0.546	0.706	0.889	1.40	1.86	2.31	2.90	3.36	6.44
9	0.129	0.261	0.543	0.703	0.883	1.38	1.83	2.26	2.82	3.25	6.01
10	0.129	0.260	0.542	0.700	0.879	1.37	1.81	2.23	2.76	3.17	5.69
11	0.129	0.260	0.540	0.697	0.876	1.36	1.80	2.20	2.72	3.11	5.45
12	0.128	0.259	0.539	0.695	0.873	1.36	1.78	2.18	2.68	3.05	5.26
13	0.128	0.259	0.538	0.694	0.870	1.35	1.77	2.16	2.65	3.01	5.11
14	0.128	0.258	0.537	0.692	0.868	1.35	1.76	2.14	2.62	2.98	4.99
15	0.128	0.258	0.536	0.691	0.866	1.34	1.75	2.13	2.60	2.95	4.88
16	0.128	0.258	0.535	0.690	0.865	1.34	1.75	2.12	2.58	2.92	4.79
17	0.128	0.257	0.534	0.689	0.863	1.33	1.74	2.11	2.57	2.90	4.71
18	0.127	0.257	0.534	0.688	0.862	1.33	1.73	2.10	2.55	2.88	4.65
19	0.127	0.257	0.533	0.688	0.861	1.33	1.73	2.09	2.54	2.86	4.59
20	0.127	0.257	0.533	0.687	0.860	1.33	1.72	2.09	2.53	2.85	4.54
21	0.127	0.257	0.532	0.686	0.859	1.32	1.72	2.08	2.52	2.83	4.49

(continued)

Table D.4 (continued)

ν	$t_{0.55}$	$t_{0.60}$	$t_{0.70}$	$t_{0.75}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$	$t_{0.999}$
22	0.127	0.256	0.532	0.686	0.858	1.32	1.72	2.07	2.51	2.82	4.45
23	0.127	0.256	0.532	0.685	0.858	1.32	1.71	2.07	2.50	2.81	4.42
24	0.127	0.256	0.531	0.685	0.857	1.32	1.71	2.06	2.49	2.80	4.38
25	0.127	0.256	0.531	0.684	0.856	1.32	1.71	2.06	2.49	2.79	4.35
26	0.127	0.256	0.531	0.684	0.856	1.31	1.71	2.06	2.48	2.78	4.32
27	0.127	0.256	0.531	0.684	0.855	1.31	1.70	2.05	2.47	2.77	4.30
28	0.127	0.256	0.530	0.683	0.855	1.31	1.70	2.05	2.47	2.76	4.28
29	0.127	0.256	0.530	0.683	0.854	1.31	1.70	2.05	2.46	2.76	4.25
30	0.127	0.256	0.530	0.683	0.854	1.31	1.70	2.04	2.46	2.75	4.23
40	0.126	0.255	0.529	0.681	0.851	1.30	1.68	2.02	2.42	2.70	4.09
60	0.126	0.254	0.527	0.679	0.848	1.30	1.67	2.00	2.39	2.66	3.96
120	0.126	0.254	0.526	0.677	0.845	1.29	1.66	1.98	2.36	2.62	3.84
∞	0.126	0.253	0.524	0.674	0.842	1.28	1.64	1.96	2.21	2.58	3.72

Table D.5 95. percentiles ($F_{0.95}$) of the F distribution (3.23); ν_1 degrees of freedom in the numerator and ν_2 in the denominator

	$\nu_1 = 1$	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
$\nu_2 = 1$	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69

(continued)

Table D.5 (continued)

	$\nu_1=1$	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.41	2.11	1.92	1.79	1.70	1.62	1.56	1.52	1.44	1.37	1.28	1.52	1.46	1.39	1.32	1.22	1.00

Table D.6 99. percentiles ($F_{0.99}$) of the F distribution (3.23); ν_1 degrees of freedom in the numerator and ν_2 in the denominator

	$\nu_1=1$	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
$\nu_2=1$	4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.5	99.0	99.2	99.3	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.4	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.6	14.4	14.2	14.0	13.9	13.8	13.8	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.8	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.3	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13

(continued)

Table D.6 (continued)

	$\nu_1=1$	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Reference

1. S. Wolfram, Wolfram MATHEMATICA. http://www.wolfram.com

Index

S. Širca, *Probability for Physicists*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-31611-6

O

P

Q

R