9. For simplicity, assume that the probabilities of the birth of a boy and of a girl are the same (which is not quite so in reality). For a certain family, we know that they have exactly two children, and that at least of them is a boy. What is the probability that they have two boys?
10. (a) ^{CS} Write a program to generate a random graph with a given edge probability p and to find its connected components. For a given number n of vertices, determine experimentally at which value of p the random graph starts to be connected, and at which value of p it starts to have a "giant component" (a component with at least $\frac{n}{2}$ vertices, say).

    (b) **Can you find theoretical explanations for the findings in (a)? You may want to consult the book [12].

10.3 Random variables and their expectation

10.3.1 Definition. *Let* (Ω, P) *be a finite probability space. By a* random variable *on* Ω*, we mean any mapping* $f\colon \Omega \to \mathbf{R}$.

A random variable f thus assigns some real number $f(\omega)$ to each elementary event $\omega \in \Omega$. Let us give several examples of random variables.

10.3.2 Example (Number of 1s). If $\mathcal{C}_n$ is the probability space of all n-term sequences of 0s and 1s, we can define a random variable f_1 as follows: for a sequence s, $f_1(s)$ is the number of 1s in s.

10.3.3 Example (Number of surviving rabbits). Each of n hunters selects a rabbit at random from a group of n rabbits, aims a gun at it, and then all the hunters shoot at once. (We feel sorry for the rabbits but this is what really happens sometimes.) A random variable f_2 is the number of rabbits that survive (assuming that no hunter misses). Formally, the probability space here is the set of all mappings $\alpha\colon \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$, each of them having the probability n^{-n}, and $f_2(\alpha) = |\{1,2,\ldots,n\} \setminus \alpha(\{1,2,\ldots,n\})|$.

10.3.4 Example (Number of left maxima). On the probability space $\mathcal{S}_n$ of all permutations of the set $\{1, 2, \ldots, n\}$, we define a random variable f_3: $f_3(\pi)$ is the number of *left maxima* of a permutation π, i.e. the number of the i such that $\pi(i) > \pi(j)$ for all $j < i$. Imagine a long-jump contest, and assume for simplicity that each competitor has a very stable performance, i.e. always jumps the same distance, and these distances are different for different competitors (these, admittedly unrealistic, assumptions can be relaxed significantly). In the first series of jumps, n competitors jump in a random order. Then f_3 means the number of times the current longest jump changes during the series.

10.3.5 Example (Sorting algorithm complexity). This random variable is somewhat more complicated. Let A be some sorting algorithm, meaning that the input of A is an n-tuple $(x_1, x_2, \ldots, x_n)$ of numbers, and the output is the same numbers in a sorted order. Suppose that the number of steps made by algorithm A only depends on the ordering of the input numbers (so that we can imagine that the input is some permutation π of the set $\{1, 2, \ldots, n\}$). This condition is satisfied by many algorithms that only use pairwise comparisons of the input numbers for sorting; some of them are frequently used in practice. We define a random variable f_4 on the probability space $\mathcal{S}_n$: we let $f_4(\pi)$ be the number of steps made by algorithm A for the input sequence $(\pi(1), \pi(2), \ldots, \pi(n))$.

10.3.6 Definition. *Let (Ω, P) be a finite probability space, and let f be a random variable on it. The* expectation *of f is a real number denoted by $\mathbf{E}[f]$ and defined by the formula*

$$\mathbf{E}[f] = \sum_{\omega \in \Omega} P(\{\omega\}) f(\omega).$$

In particular, if all the elementary events $\omega \in \Omega$ have the same probability (as is the case in almost all of our examples), then the expectation of f is simply the arithmetic average of the values of f over all elements of Ω:

$$\mathbf{E}[f] = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} f(\omega).$$

The expectation can be thought of as follows: if we repeat a random choice of an elementary event ω from Ω many times, then the average of f over these random choices will approach $\mathbf{E}[f]$.

Example 10.3.2 (Number of 1s) continued. For an illustration, we compute the expectation of the random variable f_1, the number of 1s in an n-term random sequence of 0s and 1s, according to the definition. The random variable f_1 attains a value 0 for a single sequence (all 0s), value 1 for n sequences, ..., value k for $\binom{n}{k}$ sequences from $\mathcal{C}_n$. Hence

$$\begin{aligned}\mathbf{E}[f_1] &= \frac{1}{2^n} \sum_{s \in \{0,1\}^n} f_1(s) \\ &= \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k} k.\end{aligned}$$

As we will calculate in Example 12.1.1, the final sum equals $n2^{n-1}$, and so $\mathbf{E}[f_1] = \frac{n}{2}$. Since we expect that for n coin tosses, heads should occur about $\frac{n}{2}$ times, the result agrees with intuition.

The value of $\mathbf{E}[f_1]$ can be determined in a simpler way, by the following trick. For each sequence $s \in \mathcal{C}_n$ we consider the sequence $\bar{s}$ arising from s by exchanging all 0s for 1s and all 1s for 0s. We have $f_1(s) + f_1(\bar{s}) = n$, and so

$$\begin{aligned}\mathbf{E}[f_1] &= \frac{1}{2^n} \sum_{s \in \{0,1\}^n} f_1(s) = \frac{1}{2^n \cdot 2} \sum_{s \in \{0,1\}^n} (f_1(s) + f_1(\bar{s})) \\ &= 2^{-n-1} 2^n n = \frac{n}{2}\,.\end{aligned}$$

We now describe a method that often allows us to compute the expectation in a surprisingly simple manner (we saw that the calculation according to the definition can be quite laborious even in very simple cases). We need a definition and a simple theorem.

10.3.7 Definition. *Let $A \subseteq \Omega$ be an event in a probability space (Ω, P). By the* indicator *of the event A we understand the random variable $I_A \colon \Omega \to \{0,1\}$ defined in the following way:*

$$I_A(\omega) = \begin{cases} 1 & \text{for } \omega \in A \\ 0 & \text{for } \omega \notin A. \end{cases}$$

(So the indicator is just another name for the characteristic function of A.)

10.3.8 Observation. *For any event A, we have $\mathbf{E}[I_A] = P(A)$.*

Proof. By the definition of expectation we get

$$\mathbf{E}[I_A] = \sum_{\omega \in \Omega} I_A(\omega) P(\{\omega\}) = \sum_{\omega \in A} P(\{\omega\}) = P(A).$$

□

The following result almost doesn't deserve to be called a theorem since its proof from the definition is immediate (and we leave it to the reader). But we will find this statement extremely useful in the sequel.

10.3.9 Theorem (Linearity of expectation). *Let f, g be arbitrary random variables on a finite probability space (Ω, P), and let α be a real number. Then we have $\mathbf{E}[\alpha f] = \alpha \mathbf{E}[f]$ and $\mathbf{E}[f+g] = \mathbf{E}[f] + \mathbf{E}[g]$.* □

Let us emphasize that f and g can be totally arbitrary, and need not be independent in any sense or anything like that. (On the other hand, this nice behavior of expectation *only* applies to adding random variables and multiplying them by a constant. For instance, it is not true in general that $\mathbf{E}[fg] = \mathbf{E}[f]\,\mathbf{E}[g]$!) Let us continue with a few examples of how 10.3.7–10.3.9 can be utilized.

Example 10.3.2 (Number of 1s) continued again. We calculate $\mathbf{E}[f_1]$, the average number of 1s, in perhaps the most elegant way. Let the event A_i be "the ith coin toss gives heads", so A_i is the set of all n-term sequences with a 1 in the ith position. Obviously, $P(A_i) = \frac{1}{2}$ for all i. We note that for each sequence $s \in \{0,1\}^n$ we have $f_1(s) = I_{A_1}(s) + I_{A_2}(s) + \cdots + I_{A_n}(s)$ (this is just a rather complicated way to write down a trivial statement). By linearity of expectation and then using Observation 10.3.8 we obtain

$$\begin{aligned} \mathbf{E}[f_1] &= \mathbf{E}[I_{A_1}] + \mathbf{E}[I_{A_2}] + \cdots + \mathbf{E}[I_{A_n}] \\ &= P(A_1) + P(A_2) + \cdots + P(A_n) = \frac{n}{2}. \end{aligned}$$

□

Example 10.3.3 (Number of surviving rabbits) continued. We will compute $\mathbf{E}[f_2]$, the expected number of surviving rabbits.

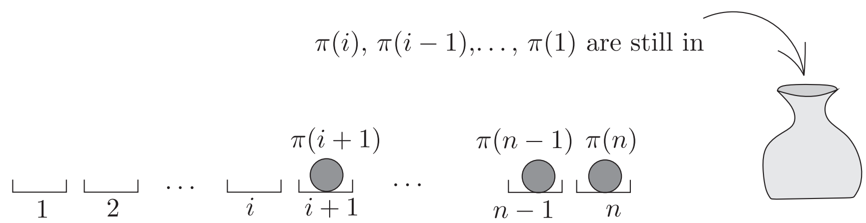

Fig. 10.2 A procedure for selecting a random permutation.

This time, let A_i be the event "the ith rabbit survives"; formally, A_i is the set of all mappings α that map no element to i. The probability that the jth hunter shoots the ith rabbit is $\frac{1}{n}$, and since the hunters select rabbits independently, we have $P(A_i) = (1 - 1/n)^n$. The remaining calculation is as in the preceding example:

$$\mathbf{E}\left[f_2\right] = \sum_{i=1}^{n} \mathbf{E}\left[I_{A_i}\right] = \sum_{i=1}^{n} P(A_i) = \left(1 - \frac{1}{n}\right)^n n \approx \frac{n}{e}$$

(since $(1 - 1/n)^n$ converges to e^{-1} for $n \to \infty$; see Exercise 3.5.2). About 37% of the rabbits survive on the average. □

Example 10.3.4 (Number of left maxima) continued. Now we will calculate the expected number of left maxima of a random permutation, $\mathbf{E}[f_3]$. Let us define A_i as the event "i is a left maximum of π", meaning that $A_i = \{\pi \in S_n\colon \pi(i) > \pi(j) \text{ for } j = 1, 2, \ldots, i-1\}$. We claim that $P(A_i) = \frac{1}{i}$. Perhaps the most intuitive way of deriving this is to imagine that the random permutation π is produced by the following method. We start with a bag containing the numbers $1, 2, \ldots, n$. We draw a number from the bag at random and declare it to be $\pi(n)$. Then we draw another random number from the bag which becomes $\pi(n-1)$ etc., as in Fig. 10.2. The value of $\pi(i)$ is selected at the moment the bag contains exactly i numbers. The probability that we choose the largest one of these i numbers for $\pi(i)$ (which is exactly the event A_i) thus equals $\frac{1}{i}$. The rest is again the same as in previous examples:

$$\mathbf{E}\left[f_3\right] = \sum_{i=1}^{n} \mathbf{E}\left[I_{A_i}\right] = \sum_{i=1}^{n} P(A_i) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\,.$$

The value of the sum of reciprocals on the right-hand side is roughly $\ln n$; see Section 3.4. □