7

Expectation and variance

Random variables are complicated objects, containing a lot of information on the experiments that are modeled by them. If we want to summarize a random variable by a *single number*, then this number should undoubtedly be its *expected value*. The expected value, also called the *expectation* or *mean*, gives the center—in the sense of average value—of the distribution of the random variable. If we allow a second number to describe the random variable, then we look at its *variance*, which is a measure of spread of the distribution of the random variable.

7.1 Expected values

An oil company needs drill bits in an exploration project. Suppose that it is known that (after rounding to the nearest hour) drill bits of the type used in this particular project will last 2, 3, or 4 hours with probabilities 0.1, 0.7, and 0.2. If a drill bit is replaced by one of the same type each time it has worn out, how long could exploration be continued if in total the company would reserve 10 drill bits for the exploration job? What most people would do to answer this question is to take the weighted average

$$0.1 \cdot 2 + 0.7 \cdot 3 + 0.2 \cdot 4 = 3.1,$$

and conclude that the exploration could continue for 10×3.1, or 31 hours. This weighted average is what we call the *expected value* or *expectation* of the random variable X whose distribution is given by

$$\mathrm{P}(X = 2) = 0.1, \quad \mathrm{P}(X = 3) = 0.7, \quad \mathrm{P}(X = 4) = 0.2.$$

It might happen that the company is unlucky and that each of the 10 drill bits has worn out after two hours, in which case exploration ends after 20 hours. At the other extreme, they may be lucky and drill for 40 hours on these 10 bits. However, it is a mathematical fact that the conclusion about a 31-hour total drilling time is correct in the following sense: for a large number n of drill bits the total running time will be around n times 3.1 hours with high probability. In the example, where $n = 10$, we have, for instance, that drilling will continue for 29, 30, 31, 32, or 33 hours with probability more than 0.86, while the probability that it will last only for 20, 21, 22, 23, or 24 hours is less than 0.00006. We will come back to this in Chapters 13 and 14. This example illustrates the following definition.

DEFINITION. The *expectation* of a discrete random variable X taking the values $a_1, a_2, \ldots$ and with probability mass function p is the number

$$\mathrm{E}[X] = \sum_i a_i \mathrm{P}(X = a_i) = \sum_i a_i p(a_i).$$

We also call $\mathrm{E}[X]$ the *expected value* or *mean* of X. Since the expectation is determined by the probability distribution of X only, we also speak of the expectation or mean of the distribution.

QUICK EXERCISE 7.1 Let X be the discrete random variable that takes the values 1, 2, 4, 8, and 16, each with probability 1/5. Compute the expectation of X.

Looking at an expectation as a weighted average gives a more physical interpretation of this notion, namely as the center of gravity of weights $p(a_i)$ placed at the points a_i. For the random variable associated with the drill bit, this is illustrated in Figure 7.1.

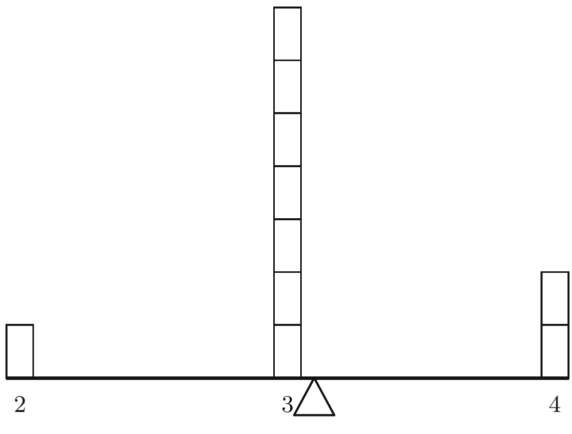

Fig. 7.1. Expected value as center of gravity.

This point of view also leads the way to how one should define the expected value of a continuous random variable. Let, for example, X be a continuous random variable whose probability density function f is zero outside the interval $[0,1]$. It seems reasonable to approximate X by the *discrete* random variable Y, taking the values

$$\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1$$

with as probabilities the masses that X assigns to the intervals $[\frac{k-1}{n}, \frac{k}{n}]$:

$$\mathrm{P}\left(Y = \frac{k}{n}\right) = \mathrm{P}\left(\frac{k-1}{n} \leq X \leq \frac{k}{n}\right) = \int_{(k-1)/n}^{k/n} f(x)\,\mathrm{d}x.$$

We have a good idea of the size of this probability. For large n, it can be approximated well in terms of f:

$$\mathrm{P}\left(Y = \frac{k}{n}\right) = \int_{k/n-1/n}^{k/n} f(x)\,\mathrm{d}x \approx \frac{1}{n} f\left(\frac{k}{n}\right).$$

The “center-of-gravity” interpretation suggests that the expectation $\mathrm{E}[Y]$ of Y should approximate the expectation $\mathrm{E}[X]$ of X. We have

$$\mathrm{E}[Y] = \sum_{k=1}^{n} \frac{k}{n} \mathrm{P}\left(Y = \frac{k}{n}\right) \approx \sum_{k=1}^{n} \frac{k}{n} f\left(\frac{k}{n}\right) \frac{1}{n}.$$

By the definition of a definite integral, for large n the right-hand side is close to

$$\int_0^1 x f(x)\,\mathrm{d}x.$$

This motivates the following definition.

DEFINITION. The *expectation* of a continuous random variable X with probability density function f is the number

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x.$$

We also call $\mathrm{E}[X]$ the *expected value* or *mean* of X. Note that $\mathrm{E}[X]$ is indeed the center of gravity of the mass distribution described by the function f:

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \frac{\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x}{\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x}.$$

This is illustrated in Figure 7.2.

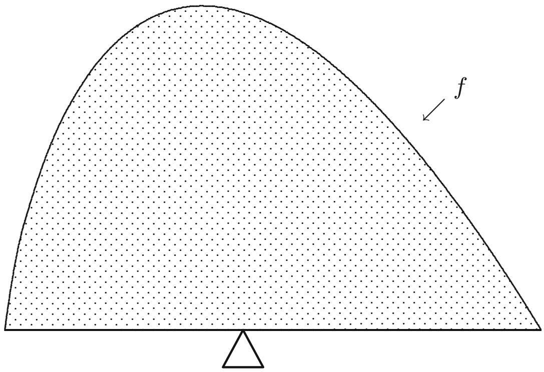

Fig. 7.2. Expected value as center of gravity, continuous case.

QUICK EXERCISE 7.2 Compute the expectation of a random variable U that is uniformly distributed over $[2, 5]$.

Remark 7.1 (The expected value may not exist!). In the definitions in this section we have been rather careless about the *convergence* of sums and integrals. Let us take a closer look at the integral $I = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$. Since a probability density function cannot take negative values, we have $I = I^- + I^+$ with $I^- = \int_{-\infty}^{0} x f(x)\,\mathrm{d}x$ a negative and $I^+ = \int_0^{\infty} x f(x)\,\mathrm{d}x$ a positive number. However, it may happen that I^- equals $-\infty$ or I^+ equals $+\infty$. If both $I^- = -\infty$ and $I^+ = +\infty$, then we say that the expected value *does not exist*. An example of a continuous random variable for which the expected value does not exist is the random variable with the Cauchy distribution (see also page 161), having probability density function

$$f(x) = \frac{1}{\pi(1+x^2)} \qquad \text{for } -\infty < x < \infty.$$

For this random variable

$$I^+ = \int_0^{\infty} x \cdot \frac{1}{\pi(1+x^2)}\,\mathrm{d}x = \left[\frac{1}{2\pi}\ln(1+x^2)\right]_0^{\infty} = +\infty,$$

$$I^- = \int_{-\infty}^{0} x \cdot \frac{1}{\pi(1+x^2)}\,\mathrm{d}x = \left[\frac{1}{2\pi}\ln(1+x^2)\right]_{-\infty}^{0} = -\infty.$$

If I^- is finite but $I^+ = +\infty$, then we say that the expected value is infinite. A distribution that has an infinite expectation is the Pareto distribution with parameter $\alpha = 1$ (see Exercise 7.11). The remarks we made on the integral in the definition of $\mathrm{E}[X]$ for continuous X apply similarly to the *sum* in the definition of $\mathrm{E}[X]$ for discrete random variables X.

7.2 Three examples

The geometric distribution

If you buy a lottery ticket every week and you have a chance of 1 in 10 000 of winning the jackpot, what is the *expected* number of weeks you have to buy tickets before you get the jackpot? The answer is: 10 000 weeks (almost two centuries!). The number of weeks is modeled by a random variable with a geometric distribution with parameter $p = 10^{-4}$.

THE EXPECTATION OF A GEOMETRIC DISTRIBUTION. Let X have a geometric distribution with parameter p; then

$$\mathrm{E}[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}.$$

Here $\sum_{k=1}^{\infty} kp(1-p)^{k-1} = 1/p$ follows from the formula $\sum_{k=1}^{\infty} kx^{k-1} = 1/(1-x)^2$ that has been derived in your calculus course. We will see a simple (probabilistic) way to obtain the value of this sum in Chapter 11.

The exponential distribution

In Section 5.6 we considered the chemical reactor example, where the residence time T, measured in minutes, has an $Exp(0.5)$ distribution. We claimed that this implies that the mean time a particle stays in the vessel is 2 minutes. More generally, we have the following.

THE EXPECTATION OF AN EXPONENTIAL DISTRIBUTION. Let X have an exponential distribution with parameter λ; then

$$\mathrm{E}[X] = \int_0^{\infty} x\lambda \mathrm{e}^{-\lambda x} \,\mathrm{d}x = \frac{1}{\lambda}.$$

The integral has been determined in your calculus course (with the technique of integration by parts).

The normal distribution

Here, using that the normal density integrates to 1 and applying the substitution $z = (x-\mu)/\sigma$,

$$\begin{aligned}\mathrm{E}[X] &= \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \,\mathrm{d}x = \mu + \int_{-\infty}^{\infty} (x-\mu) \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \,\mathrm{d}x \\ &= \mu + \sigma \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z^2} \,\mathrm{d}z = \mu,\end{aligned}$$

where the integral is 0, because the integrand is an odd function. We obtained the following rule.

THE EXPECTATION OF A NORMAL DISTRIBUTION. Let X be an $N(\mu, \sigma^2)$ distributed random variable. Then

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x = \mu.$$

7.3 The change-of-variable formula

Often one does not want to compute the expected value of a random variable X but rather of a function of X, as, for example, X^2. We then need to determine the distribution of $Y = X^2$, for example by computing the distribution function F_Y of Y (this is an example of the general problem of how distributions change under transformations—this topic is the subject of Chapter 8). For a concrete example, suppose an architect wants maximal variety in the sizes of buildings: these should be of the same width and depth X, but X is uniformly distributed between 0 and 10 meters. What is the distribution of the area X^2 of a building; in particular, will this distribution be (anything near to) uniform? Let us compute F_Y; for $0 \leq a \leq 100$:

$$F_Y(a) = \mathrm{P}\big(X^2 \leq a\big) = \mathrm{P}\big(X \leq \sqrt{a}\big) = \frac{\sqrt{a}}{10}.$$

Hence the probability density function f_Y of the area is, for $0 < y < 100$ meters squared, given by

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} \frac{\sqrt{y}}{10} = \frac{1}{20\sqrt{y}}. \tag{7.1}$$

This means that the buildings with small areas are heavily overrepresented, because f_Y explodes near 0—see also Figure 7.3, in which we plotted f_Y.

Surprisingly, this is not very visible in Figure 7.4, an example where we should believe our calculations more than our eyes. In the figure the locations of the buildings are generated by a Poisson process, the subject of Chapter 12. Suppose that a contractor has to make an offer on the price of the foundations of the buildings. The amount of concrete he needs will be proportional to the area X^2 of a building. So his problem is: what is the expected area of a building? With f_Y from (7.1) he finds

$$\mathrm{E}\left[X^2\right] = \mathrm{E}[Y] = \int_0^{100} y \cdot \frac{1}{20\sqrt{y}} \,\mathrm{d}y = \int_0^{100} \frac{\sqrt{y}}{20} \,\mathrm{d}y = \left[\frac{1}{20}\frac{2}{3} y^{\frac{3}{2}}\right]_0^{100} = 33\tfrac{1}{3} \ \mathrm{m}^2.$$

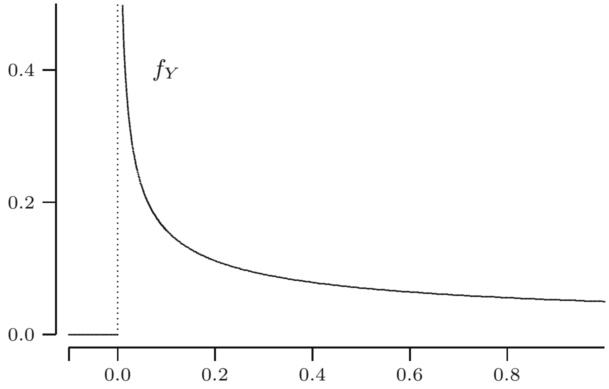

Fig. 7.3. The probability density of the square of a $U(0,10)$ random variable.

It is interesting to note that we really *need* to do this calculation, because the expected area is *not* simply the product of the expected width and the expected depth, which is $25\,\mathrm{m}^2$. However, there is a much easier way in which the contractor could have obtained this result. He could have argued that the value of the *area* is x^2 when x is the width, and that he should take the weighted average of *those* values, where the weight at width x is given by the value $f_X(x)$ of the probability density of X. Then he would have computed

$$\mathrm{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\,\mathrm{d}x = \int_0^{10} x^2 \cdot \frac{1}{10}\,\mathrm{d}x = \left[\frac{1}{30}x^3\right]_0^{10} = 33\tfrac{1}{3}\ \mathrm{m}^2.$$

It is indeed a mathematical theorem that this is *always* a correct way to compute expected values of functions of random variables.

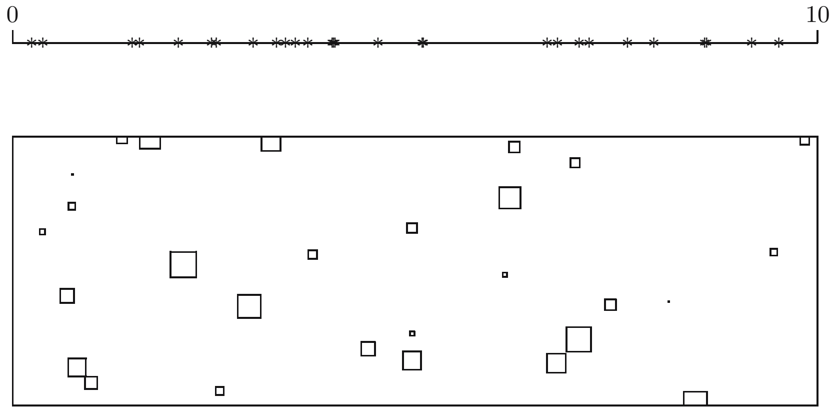

Fig. 7.4. Top: widths of the buildings between 0 and 10 meters. Bottom: corresponding buildings in a 100×300 m area.

THE CHANGE-OF-VARIABLE FORMULA. Let X be a random variable, and let $g : \mathbb{R} \to \mathbb{R}$ be a function.
If X is discrete, taking the values $a_1, a_2, \ldots$, then

$$\mathrm{E}[g(X)] = \sum_i g(a_i)\mathrm{P}(X = a_i)\,.$$

If X is continuous, with probability density function f, then

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x.$$

QUICK EXERCISE 7.3 Let X have a $Ber(p)$ distribution. Compute $\mathrm{E}\left[2^X\right]$.

An operation that occurs very often in practice is a change of units, e.g., from Fahrenheit to Celsius. What happens then to the expectation? Here we have to apply the formula with the function $g(x) = rx + s$, where r and s are real numbers. When X has a continuous distribution, the change-of-variable formula yields:

$$\begin{aligned}\mathrm{E}[rX + s] &= \int_{-\infty}^{\infty} (rx + s)f(x)\,\mathrm{d}x \\ &= r\int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x + s\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \\ &= r\mathrm{E}[X] + s.\end{aligned}$$

A similar computation with integrals replaced by sums gives the same result for discrete random variables.

7.4 Variance

Suppose you are offered an opportunity for an investment whose expected return is € 500. If you are given the extra information that this expected value is the result of a 50% chance of a € 450 return and a 50% chance of a € 550 return, then you would not hesitate to spend € 450 on this investment. However, if the expected return were the result of a 50% chance of a € 0 return and a 50% chance of a € 1000 return, then most people would be reluctant to spend such an amount. This demonstrates that the spread (around the mean) of a random variable is of great importance. Usually this is measured by the expected squared deviation from the mean.

DEFINITION. The *variance* $\mathrm{Var}(X)$ of a random variable X is the number

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right].$$

Note that the variance of a random variable is always positive (or 0). Furthermore, there is the question of existence and finiteness (cf. Remark 7.1). In practical situations one often considers the *standard deviation* defined by $\sqrt{\operatorname{Var}(X)}$, because it has the same dimension as $\mathrm{E}[X]$.

As an example, let us compute the variance of a normal distribution. If X has an $N(\mu, \sigma^2)$ distribution, then:

$$\begin{aligned}\operatorname{Var}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,\mathrm{d}x \\ &= \sigma^2 \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z.\end{aligned}$$

Here we substituted $z = (x-\mu)/\sigma$. Using integration by parts one finds that

$$\int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z = 1.$$

We have found the following property.

VARIANCE OF A NORMAL DISTRIBUTION. Let X be an $N(\mu, \sigma^2)$ distributed random variable. Then

$$\operatorname{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,\mathrm{d}x = \sigma^2.$$

QUICK EXERCISE 7.4 Let us call the two returns discussed above Y_1 and Y_2, respectively. Compute the variance and standard deviation of Y_1 and Y_2.

It is often not practical to compute $\operatorname{Var}(X)$ directly from the definition, but one uses the following rule.

AN ALTERNATIVE EXPRESSION FOR THE VARIANCE. For any random variable X,

$$\operatorname{Var}(X) = \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2.$$

To see that this rule holds, we apply the change-of-variable formula. Suppose X is a continuous random variable with probability density function f (the discrete case runs completely analogously). Using the change-of-variable formula, well-known properties of the integral, and $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$, we find

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\
&= \int_{-\infty}^{\infty} (x - \mathrm{E}[X])^2 f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(x^2 - 2x\mathrm{E}[X] + (\mathrm{E}[X])^2\right) f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x - 2\mathrm{E}[X] \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x + (\mathrm{E}[X])^2 \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \\
&= \mathrm{E}\left[X^2\right] - 2(\mathrm{E}[X])^2 + (\mathrm{E}[X])^2 \\
&= \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2.
\end{aligned}$$

With this rule we make two steps: first we compute $\mathrm{E}[X]$, then we compute $\mathrm{E}\left[X^2\right]$. The latter is called *the second moment* of X. Let us compare the computations, using the definition and this rule for the drill bit example. Recall that for this example X takes the values $2, 3$, and 4 with probabilities $0.1, 0.7$, and 0.2. We found that $\mathrm{E}[X]= 3.1$. According to the definition

$$\begin{aligned}
\mathrm{Var}(X) = \mathrm{E}\left[(X - 3.1)^2\right] &= 0.1 \cdot (2 - 3.1)^2 + 0.7 \cdot (3 - 3.1)^2 + 0.2 \cdot (4 - 3.1)^2 \\
&= 0.1 \cdot (-1.1)^2 + 0.7 \cdot (-0.1)^2 + 0.2 \cdot (0.9)^2 \\
&= 0.1 \cdot 1.21 + 0.7 \cdot 0.01 + 0.2 \cdot 0.81 \\
&= 0.121 + 0.007 + 0.162 \\
&= 0.29.
\end{aligned}$$

Using the rule is neater and somewhat faster:

$$\begin{aligned}
\mathrm{Var}(X) = \mathrm{E}\left[X^2\right] - (3.1)^2 &= 0.1 \cdot 2^2 + 0.7 \cdot 3^2 + 0.2 \cdot 4^2 - 9.61 \\
&= 0.1 \cdot 4 + 0.7 \cdot 9 + 0.2 \cdot 16 - 9.61 \\
&= 0.4 + 6.3 + 3.2 - 9.61 \\
&= 0.29.
\end{aligned}$$

What happens to the variance if we change units? At the end of the previous section we showed that $\mathrm{E}[rX + s] = r\mathrm{E}[X] + s$. This can be used to obtain the corresponding rule for the variance under change of units (see also Exercise 7.15).

Expectation and variance under change of units. For any random variable X and any real numbers r and s,

$$\mathrm{E}[rX + s] = r\mathrm{E}[X] + s, \quad \text{and} \quad \mathrm{Var}(rX + s) = r^2\mathrm{Var}(X)\,.$$

Note that the variance is insensitive to the shift over s. Can you understand why this must be true without doing any computations?

7.5 Solutions to the quick exercises

7.1 We have

$$\mathrm{E}[X] = \sum_i a_i \mathrm{P}(X = a_i) = 1 \cdot \frac{1}{5} + 2 \cdot \frac{1}{5} + 4 \cdot \frac{1}{5} + 8 \cdot \frac{1}{5} + 16 \cdot \frac{1}{5} = \frac{31}{5} = 6.2.$$

7.2 The probability density function f of U is given by $f(x) = 0$ outside $[2,5]$ and $f(x) = 1/3$ for $2 \leq x \leq 5$; hence

$$\mathrm{E}[U] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_2^5 \frac{1}{3} x\,\mathrm{d}x = \left[\frac{1}{6}x^2\right]_2^5 = 3\tfrac{1}{2}.$$

7.3 Using the change-of-variable formula we obtain

$$\begin{aligned}\mathrm{E}\left[2^X\right] &= \sum_i 2^{a_i} \mathrm{P}(X = a_i) \\ &= 2^0 \cdot \mathrm{P}(X = 0) + 2^1 \cdot \mathrm{P}(X = 1) \\ &= 1 \cdot (1 - p) + 2 \cdot p = 1 - p + 2p = 1 + p.\end{aligned}$$

You could also have noted that $Y = 2^X$ has a distribution given by $\mathrm{P}(Y = 1) = 1 - p, \mathrm{P}(Y = 2) = p$; hence

$$\mathrm{E}\left[2^X\right] = \mathrm{E}[Y] = 1 \cdot \mathrm{P}(Y = 1) + 2 \cdot \mathrm{P}(Y = 2) = 1 \cdot (1 - p) + 2 \cdot p = 1 + p.$$

7.4 We have

$$\mathrm{Var}(Y_1) = \tfrac{1}{2}(450 - 500)^2 + \tfrac{1}{2}(550 - 500)^2 = 50^2 = 2500,$$

so Y_1 has standard deviation € 50 and

$$\mathrm{Var}(Y_2) = \tfrac{1}{2}(0 - 500)^2 + \tfrac{1}{2}(1000 - 500)^2 = 500^2 = 250\,000,$$

so Y_2 has standard deviation € 500.

7.6 Exercises

7.1 ⊡ Let T be the outcome of a roll with a fair die.

a. Describe the probability distribution of T, that is, list the outcomes and the corresponding probabilities.

b. Determine $\mathrm{E}[T]$ and $\mathrm{Var}(T)$.

7.2 ⊡ The probability distribution of a discrete random variable X is given by

$$\mathrm{P}(X = -1) = \tfrac{1}{5}, \quad \mathrm{P}(X = 0) = \tfrac{2}{5}, \quad \mathrm{P}(X = 1) = \tfrac{2}{5}.$$

a. Compute $\mathrm{E}[X]$.

b. Give the probability distribution of $Y = X^2$ and compute $\mathrm{E}[Y]$ using the distribution of Y.

c. Determine $\mathrm{E}\left[X^2\right]$ using the change-of-variable formula. Check your answer against the answer in **b**.

d. Determine $\mathrm{Var}(X)$.

7.3 For a certain random variable X it is known that $\mathrm{E}[X] = 2$, $\mathrm{Var}(X) = 3$. What is $\mathrm{E}\left[X^2\right]$?

7.4 Let X be a random variable with $\mathrm{E}[X] = 2$, $\mathrm{Var}(X) = 4$. Compute the expectation and variance of $3 - 2X$.

7.5 ⊡ Determine the expectation and variance of the $Ber(p)$ distribution.

7.6 ⊞ The random variable Z has probability density function $f(z) = 3z^2/19$ for $2 \leq z \leq 3$ and $f(z) = 0$ elsewhere. Determine $\mathrm{E}[Z]$. Before you do the calculation: will the answer lie closer to 2 than to 3 or the other way around?

7.7 Given is a random variable X with probability density function f given by $f(x) = 0$ for $x < 0$, and for $x > 1$, and $f(x) = 4x - 4x^3$ for $0 \leq x \leq 1$. Determine the expectation and variance of the random variable $2X + 3$.

7.8 ⊡ Given is a continuous random variable X whose distribution function F satisfies $F(x) = 0$ for $x < 0$, $F(x) = 1$ for $x > 1$, and $F(x) = x(2 - x)$ for $0 \leq x \leq 1$. Determine $\mathrm{E}[X]$.

7.9 Let U be a random variable with a $U(\alpha, \beta)$ distribution.

a. Determine the expectation of U.

b. Determine the variance of U.

7.10 ⊡ Let X have an exponential distribution with parameter λ.

a. Determine $\mathrm{E}[X]$ and $\mathrm{E}\left[X^2\right]$ using partial integration.

b. Determine $\mathrm{Var}(X)$.

7.11 ⊡ In this exercise we take a look at the mean of a Pareto distribution.

a. Determine the expectation of a $Par(2)$ distribution.

b. Determine the expectation of a $Par(\frac{1}{2})$ distribution.

c. Let X have a $Par(\alpha)$ distribution. Show that $\mathrm{E}[X] = \alpha/(\alpha - 1)$ if $\alpha > 1$.

7.12 For which α is the variance of a $Par(\alpha)$ distribution finite? Compute the variance for these α.

7.13 Remember that we found on page 95 that the expected area of a building was $33\frac{1}{3}$ m^2, whereas the square of the expected width was only 25 m^2. This phenomenon is more general: show that for any random variable X one has $\mathrm{E}\left[X^2\right] \geq \left(\mathrm{E}\left[X\right]\right)^2$.
Hint: you might use that $\mathrm{Var}(X) \geq 0$.

7.14 Suppose we choose arbitrarily a point from the square with corners at (2,1), (3,1), (2,2), and (3,2). The random variable A is the area of the triangle with its corners at (2,1), (3,1), and the chosen point. (See also Exercise 5.9 and Figure 7.5.) Compute $\mathrm{E}\left[A\right]$.

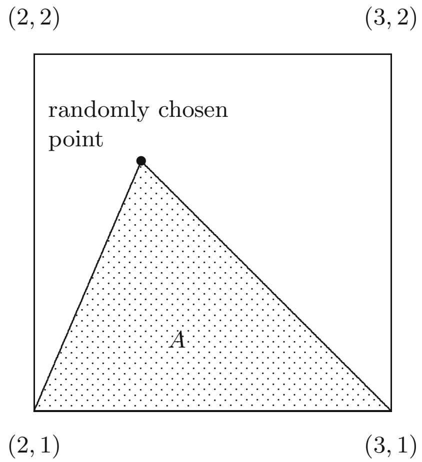

Fig. 7.5. A triangle in a 1×1 square.

7.15 ⊞ Let X be a random variable and r and s any real numbers. Use the change-of-units rule $\mathrm{E}\left[rX + s\right] = r\mathrm{E}\left[X\right] + s$ for the expectation to obtain **a** and **b**.

a. Show that $\mathrm{Var}(rX) = r^2\mathrm{Var}(X)$.
b. Show that $\mathrm{Var}(X + s) = \mathrm{Var}(X)$.
c. Combine parts **a** and **b** to show that

$$\mathrm{Var}(rX + s) = r^2\mathrm{Var}(X)\,.$$

7.16 ⊡ The probability density function f of the random variable X used in Figure 7.2 is given by $f(x) = 0$ outside $(0, 1)$ and $f(x) = -4x\ln(x)$ for $0 < x < 1$. Compute the position of the balancing point in the figure, that is, compute the expectation of X.

7.17 ⊞ Let U be a discrete random variable taking the values $a_1, \ldots, a_r$ with probabilities $p_1, \ldots, p_r$.

a. Suppose all $a_i \geq 0$, but that $\mathrm{E}\left[U\right]$=0. Show then

$$a_1 = a_2 = \cdots = a_r = 0.$$

In other words; $\mathrm{P}(U = 0) = 1$.

b. Suppose that V is a random variable taking the values $b_1, \ldots, b_r$ with probabilities $p_1, \ldots, p_r$. Show that $\mathrm{Var}(V) = 0$ implies

$$\mathrm{P}(V = \mathrm{E}[V]) = 1.$$

Hint: apply **a** with $U = (V - \mathrm{E}[V])^2$.