



Anyone writing a book will rarely follow a plan that was not revised several times during the process. This was definitely the case when this book was written. We have discussed many different versions before we arrived at the current format. In some of these versions mathematical terms like “convergence of functions” or “cardinality of sets” played an important role. At the end, we found a way to discuss the Brownian motion without using these terms explicitly. The obvious consequence could have been to simply drop this material.

Discussions with students and colleagues taught us that these topics can also be of use in several other areas of economics. Therefore we decided to leave the supplements in our book. The four subsequent sections can be read independently of each other. The entire chapter can be skipped for the understanding of the Brownian motion.

7.1 Cardinality of Sets

Imagine adding 0 to the set of scores on a dice as another element:

$$\{0, 1, 2, 3, 4, 5, 6\}.$$

Obviously, this set is larger than the original set: instead of six there exist now seven elements. With this simple fact in mind, one is inclined to conclude that this idea will also be applicable in the case of infinite sets. For example, if we compare the set \mathbb{N} of all natural numbers with the set \mathbb{Z} of integers, it seems reasonable to suppose that \mathbb{Z} is greater than \mathbb{N} .

However, one cannot prove whether such a proposition is correct or false by looking at the number of elements. This number is infinite in \mathbb{Z} as well as \mathbb{N} , and we had already realized that infinite is not a number that can be used to perform simple arithmetic operations such as addition or comparisons. Thus, one has to

create another concept if one wants to compare infinite sets. This boils down to cardinality.

If one looks at infinite sets, results dealing with finite sets seem to contradict common sense. First, one might think that the set of natural numbers is smaller than the set of integers since all negative values $-1, -2, \dots$ are missing. However, one can prove by a simple consideration that this conclusion is mistaken. Rather, it is shown that the set of integers is exactly as large as the set of natural numbers or both have “the same cardinality” which we will explain below. This underlines the fact that infinity must be handled very carefully. It is better not to rely on common sense or “intuition”!

The idea of cardinality is to employ a one-to-one relation when comparing two sets rather than counting their elements. Two sets are said to have the same cardinality (or are “equal in size”) only if there exists a one-to-one relation between all their elements.

With finite sets counting elements or using one-to-one relations lead to the same result. Figure 7.1 illustrates that the set with seven elements is greater than the set with six elements: one element from the set $\{0, 1, \dots, 6\}$ will never find a “partner.”

In the case of the two infinite sets, however, the outcome is surprising. This is demonstrated by the assignment in Fig. 7.2: each natural number is mapped to exactly one integer and this mapping is one-to-one. One can clearly observe that both every natural number and every integer appear exactly once. Those preferring formulas might use

$$f : \mathbb{N} \rightarrow \mathbb{Z}, \quad f(n) = \begin{cases} -\frac{n}{2}, & \text{if } n \text{ is even;} \\ \frac{n+1}{2}, & \text{if } n \text{ is odd.} \end{cases} \quad (7.1)$$

f is a function that obviously assigns an integer to each natural number n and f is also reversible in the sense that every integer in \mathbb{Z} is also captured.

The idea of cardinality will be further illustrated with another example.

Example 7.1 (Cantor’s Diagonal Argument) The set of nonnegative rational numbers \mathbb{Q}_+ has the same cardinality as the set of natural numbers. To show the equivalence it is necessary to prove—analogue to Fig. 7.2—that it is possible to uniquely assign all nonnegative rational numbers to natural numbers.

Fig. 7.1 The two finite sets $\{0, 1, \dots, 6\}$ and $\{1, 2, \dots, 6\}$ do not have equal cardinality

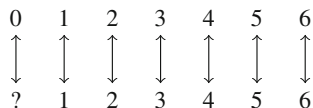
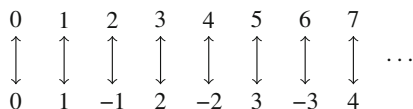


Fig. 7.2 The infinite sets \mathbb{N} and \mathbb{Z} have equal cardinality



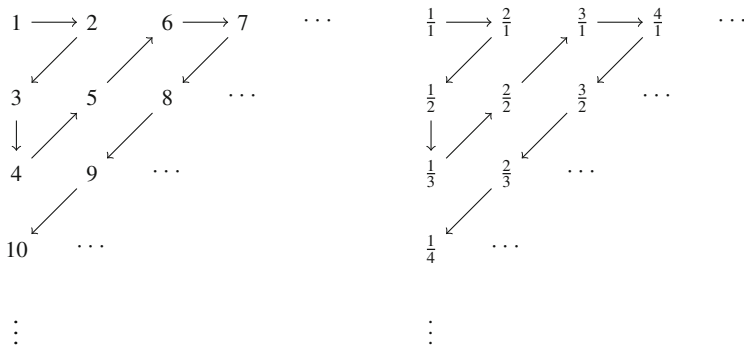


Fig. 7.3 Cantor’s diagonal argument to prove that \mathbb{N} and \mathbb{Q}_+ have equal cardinality

The rational numbers \mathbb{Q}_+ consist of all fractions $\frac{m}{n}$ with m and n being positive natural numbers. These rational numbers are now arranged in an infinite two-dimensional matrix as shown in Fig. 7.3.¹ The arrows shown illustrate how one may imagine the one-to-one correspondence between the natural and the rational numbers: the 1 is assigned to fraction $\frac{1}{1}$, the 2 to fraction $\frac{2}{1}$, the 3 to fraction $\frac{1}{2}$, the 4 to fraction $\frac{1}{3}$, the 5 to fraction $\frac{2}{2}$, and so on.

This procedure would create a one-to-one relation if there was not an annoying blemish. The right matrix contains too many elements. The rational numbers $\frac{1}{1}, \frac{2}{2}, \frac{3}{3}, \dots$ or $\frac{6}{17}, \frac{9}{34}, \frac{9}{51}, \dots$ are actually identical and do not represent different rational numbers at all. Therefore, they must not be assigned to different natural numbers. One has to make sure that they are accounted for only once. This is achieved by “thinning-out” the right matrix. All fractions $\frac{m}{n}$ consisting of m, n which are not coprime are deleted. In this case the diagonal construction is only carried out for values that are coprime. The formal proof is much more complicated due to this “thinning-out” and must—if one wants to be formally precise—be conducted with complete induction. However, we will not present the details of this proof.

A set whose cardinality corresponds to the cardinality of the natural numbers is called countable. In this sense natural numbers, integers and rational numbers are countable. Countable quantities are of great importance because they can appear as indices in sums and products. An expression of the form $\sum_{i \in A} a_i$ makes sense if and only if A is countable. If $A = \mathbb{N}$ one can even write $\lim_{n \rightarrow \infty} \sum_{i=1}^n a_i$ for this sum.

One could suspect that for all infinite sets it can be proven—with ingenious tricks—that they are countable. However, that is not the case and we will show for a very prominent set that it is larger than the set of natural numbers.

¹The idea of this proof goes back to the founder of set theory, Georg Ferdinand Ludwig Philipp Cantor (1845–1918, German mathematician).

Fig. 7.4 Cantor's diagonal argument to prove the uncountability of real numbers

$$\begin{array}{l} 0 . a_1 a_2 a_3 a_4 \dots \\ 0 . b_1 b_2 b_3 b_4 \dots \\ 0 . c_1 c_2 c_3 c_4 \dots \\ 0 . d_1 d_2 d_3 d_4 \dots \\ \vdots \end{array}$$

Example 7.2 (Uncountability) We prove that the set of real numbers \mathbb{R} has a different cardinality than the set of natural numbers. That is quite simple.

To this end we assume that someone claims being able to map the set of real numbers one-to-one to the set of natural numbers. This person would be able to list all real numbers one after the other. This would constitute a sequence of all real numbers. In particular, this person can name a unique predecessor and successor for each real number. We will show that at least one real number is still missing—which is a contradiction. This proves that the set of real numbers must be larger than the set of natural numbers.

In Fig. 7.4 we present the sequence of real numbers with their (possibly infinite) decimal representation which the above person claims to be complete, i.e., containing all real numbers. Instead of the decimals 0, 1, ..., 9 we use symbols $a_i, b_i, c_i, d_i, \dots$ for every real number.²

The missing number can be constructed very easily. We consider Fig. 7.4 as a matrix of numbers and focus on the diagonal (the diagonal elements are printed in red). Using the diagonal we form a new real number of the form $0 . z_1 z_2 z_3 z_4 \dots$. As first decimal z_1 of this new real number, a decimal must be selected such that it does not equal a_1 . The second decimal must fulfill the inequality $z_2 \neq b_2$, for the third decimal the inequality $z_3 \neq c_3$ must hold, and so on. The new real number formed in this way cannot match any of the numbers mentioned in our person's supposedly complete list. With each element of our person's list (at least) one decimal in the representation is different from our newly constructed number. We have found the missing number!

These considerations show that the set of real numbers can hardly be counted. It is said that the real numbers are *uncountable*. Therefore, it follows that an expression of the form $\sum_{i \in \mathbb{R}} a_i$ does not represent a mathematically meaningful term: each element i in an index set must have a unique predecessor and a unique successor, a situation impossible for the real numbers \mathbb{R} .

Example 7.2 shows that there exist infinite sets with different cardinalities. The set of real numbers \mathbb{R} is “larger” than \mathbb{N} , while the sets of natural numbers is “as large” as the sets \mathbb{Z} and \mathbb{Q}_+ . In mathematics this is indicated by appropriate symbols. The number of natural numbers is not indicated by the rather fuzzy infinity sign ∞

²Without loss of generality we can ignore all digits before the decimal point.

but by the symbol \aleph_0 .³ Since the cardinality of the real numbers is greater than \aleph_0 , the symbol \aleph_1 is used.

Concluding Remark Finally, we would like to draw the reader’s attention to an interesting issue. We have already shown that the set of natural numbers is smaller than the set of real numbers. Instead of the set of natural numbers, one could use their power set $\mathcal{P}(\mathbb{N})$, i.e., the set of all subsets of natural numbers. This power set contains the set of all even numbers, the set of all odd numbers, the set of all natural numbers less than 5, and so on. Without presenting the mathematical details, it can be shown that the power set has the same cardinality as the set of real numbers. On page 20 we had made it clear that for a finite set of n elements the number of subsets is just 2^n . This relationship is assigned to the symbols just introduced by writing the following equation:

$$2^{\aleph_0} = \aleph_1. \tag{7.2}$$

However, this symbolic notation should not be confused with real arithmetic operations. One must not write $\aleph_0 = \log_2(\aleph_1)$.

What do these considerations tell us? If mathematicians transfer as in (7.2) a symbolic notation from one subject area to another, one is tempted to use it in all its dimensions. Unfortunately, such practice cannot only be wrong but even be dangerous. We have already experienced this situation while discussing the notation of Brownian motion.

7.2 Continuous and Almost Nowhere Differentiable Functions

In order to discuss the Brownian motion thoroughly, it is useful to deal with remarkable features of functions. The paths of Brownian motion are continuous functions which one cannot differentiate at (almost) any point. Anyone wanting to handle such functions properly must recognize that the use of mathematical operations known from ordinary analysis is inadmissible. Compared to ordinary analysis dealing with Brownian paths can be considered as being “exotic.”

Non-mathematicians probably cannot imagine continuous functions that are not differentiable (almost) anywhere. We would like to assist this understanding by an example developed by Weierstraß.⁴ He also showed that in mathematics such functions are anything but rare. Prior to Weierstraß these functions had been

³The symbol \aleph is the first letter of the Hebrew alphabet and is pronounced aleph.

⁴Karl Theodor Wilhelm Weierstraß (1815–1897, German mathematician). In 1872 Weierstraß introduced this function in a lecture and claimed that Riemann had knowledge of such an example. However, no such reference has been found in Riemann’s inheritance. Around 1830 Bolzano found the first example of a function that could not be differentiated almost anywhere in a manuscript that was published only in 1922.

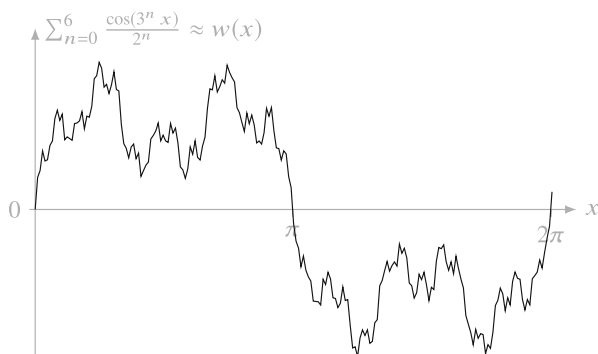


Fig. 7.5 Approximation of the Weierstraß function $w(x)$ using the first seven summands

regarded as “monster curves.”⁵ It was assumed that these functions were either only special cases or that the points where differentiation is not possible were indeed rare.

Weierstraß considered the function

$$w(x) = \sum_{n=0}^{\infty} \frac{\sin(3^n x)}{2^n}. \quad (7.3)$$

To give an idea of the appearance of this function, Fig. 7.5 shows only the first seven summands of a Taylor series.⁶ We concentrate on two characteristics of the Weierstraß function: first its continuity and second its differentiability.

Non-mathematicians state that a function is continuous if one can draw its path without interrupting the movement of the drawing pen. Although this is not a precise definition one may suspect that the Weierstraß function is continuous when looking at Fig. 7.5. Even with more precision the same result applies: the numerator of each fraction is at most 1 and the denominator grows exponentially. Therefore, the sum converges for each x . Furthermore, it also converges uniformly. This means that the difference between $\sum_{n=0}^m \frac{\sin(3^n x)}{2^n}$ and $w(x)$ going to zero can be estimated independently of x . In such cases the property of continuity of the summands $\frac{\sin(3^n x)}{2^n}$ also applies to the function $w(x)$.

The above considerations do not represent a complete proof but only give an indication of the evidence: the result is intuitively appealing. Looking at the definition of the function $w(x)$ the following observation is decisive. The numerator

⁵The French mathematician Charles Hermite (1822–1901) wrote in 1893 in a letter to Stieltjes: “I avert myself with horror and shock from this lamentable plague of functions that have no derivative at all.”

⁶The picture does not change very much if additional summands are added with the approximation error being reduced.

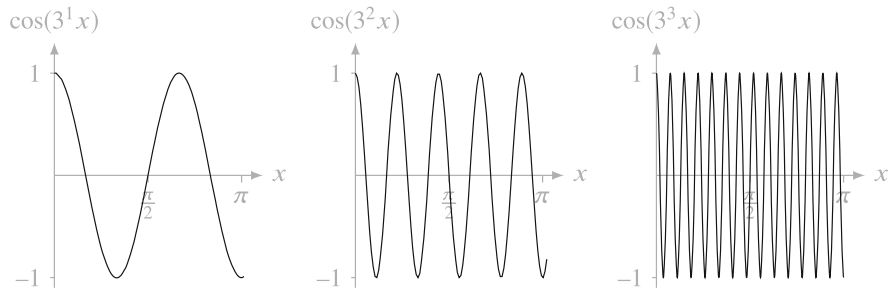


Fig. 7.6 Cosine functions $\cos(3^1 x)$, $\cos(3^2 x)$, and $\cos(3^3 x)$

of each additional summand exists in the interval $[-1, 1]$. On the other hand, the denominator of each new summand grows exponentially. Hence, each new summand (however it may behave) contributes only marginally to the change of the function value. Therefore, continuity is maintained at the limit.

Let us turn to the second characteristic of the function $w(x)$. Weierstraß was able to show that the function cannot be differentiated except for a few values x . While the proof is difficult, one can illustrate the result as follows: deriving the sum with respect to x one obtains⁷

$$\frac{dw(x)}{dx} = \lim_{N \rightarrow \infty} \sum_{n=0}^N \left(\frac{3}{2}\right)^n \cos(3^n x). \quad (7.4)$$

To examine this limit in more detail we first ignore the factor $\left(\frac{3}{2}\right)^n$ and draw several graphs of the function $\cos(3^n x)$ depending on n (see Fig. 7.6).

It can easily be seen that the frequency of the cosine function increases with every exponent n . Since the increasing fluctuations are multiplied by the factor $\left(\frac{3}{2}\right)^n$, their impact on the sum grows with n . Obviously, the sum can only converge for numbers x where the cosine function approaches zero. The zeros of these cosine functions are very thinly scattered.⁸ For all other x the sum diverges to plus or minus infinity and this represents the default case. Thus, the first derivative of this function is almost everywhere either minus or plus infinity. This implies that the function cannot be differentiated anywhere.

⁷We will change derivation and infinite summation in our calculation which is mathematically inadmissible under these circumstances. The following argument therefore does not constitute full proof.

⁸The set of those x has Lebesgue measure zero.

7.3 Convergence Terms

From numerous discussions with students and colleagues we learned that there is certainly interest in looking more closely at the issue of convergence of functions. When looking at convergence of numbers it is entirely irrelevant how to define convergence precisely. Regardless of the definition of convergence of numbers, all turn out to be equivalent. However, this is entirely different when dealing with sequences of functions. There are many different ways to define convergence with each option being fundamentally different from one another. While most non-mathematicians can imagine what a sequence of numbers is, the issue of dealing with a sequence of functions is very different.

To illustrate this phenomenon we use an analogy. Finding the shortest route from Berlin to San Francisco depends on the way the earth is looked at. Using a conventional map of the world it will be concluded that the shortest route of the two cities is always south of 53° North. However, when using a globe you will find that the shortest route is in fact via Greenland. This analogy is similar to the convergence concept for functions: there are not just one but several ways of defining the convergence of a sequence of functions. The results depend on the chosen convergence definition.

Convergence is important in the context of limits. To understand the applications, it is useful to realize how proofs are conducted in the theory of Lebesgue integration⁹: if one wants to prove that a certain property or a given proposition applies in general, one can make life easier to start by proving the correctness of the proposition for linear or piecewise linear functions. In order to show the general validity, one has to move from these simple functions to more general ones. To this end one has to consider the limit of a sequence of functions. A proposition applying to each (piecewise linear or simple) element of a function sequence will also apply to the limit of this sequence and thus to a general function. It should be noted it must not matter whether one integrates first and subsequently passes to the limit or vice versa. Integration and limit must be interchangeable:

$$\lim_n \int_{\Omega} \stackrel{!}{=} \int_{\Omega} \lim_n . \quad (7.5)$$

Let us look at random variables as an example of functions. For random variables expectation and variance are (Lebesgue) integrals.¹⁰ From (7.5) it should follow

$$\lim_{n \rightarrow \infty} E[Z_n] \stackrel{!}{=} E \left[\lim_{n \rightarrow \infty} Z_n \right] \quad (7.6)$$

⁹See page 71 ff.

¹⁰See page 80.

and

$$\lim_{n \rightarrow \infty} \text{Var}[Z_n] \stackrel{!}{=} \text{Var}\left[\lim_{n \rightarrow \infty} Z_n\right]. \quad (7.7)$$

Remember that Z_n is a random variable and thus a measurable function.

The above claims deserve two remarks: first, there is an exclamation mark above the equal signs. We need a definition of a limit such that right and left sides are identical. It is apparent that limit and expectation or limit and variance can be swapped. Second, consider the left side of Eq. (7.5) which represents limits of sequences of numbers since expected values and variances are numbers. The right side of Eq. (7.5) does not contain a sequence of numbers but a sequence of functions. While students of economics are aware of how to determine a limit of a sequence of numbers, they may not know what a sequence of a function is let alone how to determine its limit.

Before introducing two important concepts of convergence, namely pointwise convergence and mean square convergence,¹¹ we will start with sequences of numbers.

Sequences of Numbers In mathematical analysis, it is stated that a sequence of numbers converges to a limit if the numbers with a sufficiently large index will approach a particular value. For example, if you look at the sequence of numbers

$$s_n = a + \frac{1}{n} \quad \text{with } n = 1, 2, \dots, \quad (7.8)$$

we have

$$s_1 = a + 1, \quad s_2 = a + \frac{1}{2}, \quad s_3 = a + \frac{1}{3}, \quad (7.9)$$

and so on. By letting n increase the second summand decreases and approaches zero.¹² For $n \rightarrow \infty$ the summand can be neglected. Thus, the sequence converges to a which is written as

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \left(a + \frac{1}{n}\right) = a. \quad (7.10)$$

After exploring sequences of numbers we will now concentrate on sequences of functions.

¹¹In addition to these two types of convergence, there exist in mathematics a few others definitions that will not be discussed here.

¹²One easily realizes that, for example, the sequence $s_n = (-1)^n$ does not converge with increasing n . Such sequences are called divergent.

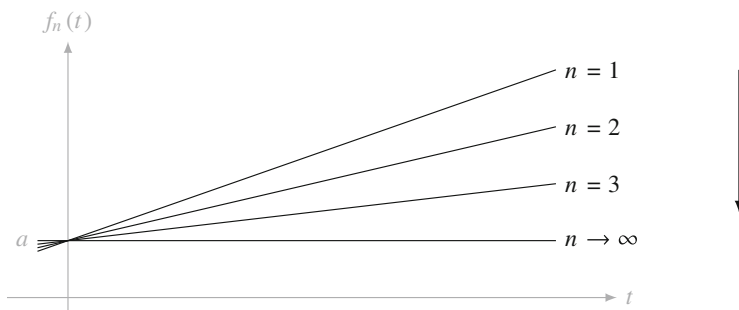


Fig. 7.7 What is the limit of a function sequence?

Sequences of Functions We look at the simple example

$$f_n(t) = a + \frac{t}{n}. \quad (7.11)$$

With increasing n one obtains

$$f_1(t) = a + t, \quad f_2(t) = a + \frac{t}{2}, \quad f_3(t) = a + \frac{t}{3}, \quad (7.12)$$

and so on. It seems clear that such a sequence of functions converges and how its limit is determined. In a sequence of numbers individual numerical values at the limit should converge to a certain value. With a sequence of functions it is quite plausible to expect that with increasing n a function “clings to a limit function.” In the above example the functions $f_n(t)$ are approaching the limit function $f(t) = a$. Figure 7.7 illustrates this vividly. With increasing n the influence of the term $\frac{t}{n}$ gets less and less significant in Eq. (7.11). The limit function takes the form $\lim_{n \rightarrow \infty} f_n(t) = a$.

Pointwise Convergence This definition can be regarded as a “natural” candidate based on the above example.

Definition 7.1 (Pointwise Convergence) Consider a sequence of functions of the form $f_n : \Omega \rightarrow \mathbb{R}$.

A sequence of functions f_n converges pointwise¹³ to a function f if and only if the following is valid¹⁴:

$$\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega) \quad \forall \omega \in \Omega. \quad (7.13)$$

¹³The noun is “pointwise convergence,” and the verb is “to converge pointwise.”

¹⁴The definition is easy to interpret: it is required here that for each value ω the sequence $f_n(\omega)$ converges against the number $f(\omega)$. So you concentrate on each value $f(\omega)$ and ignore the values $f(\omega \pm \delta)$ “next to it” when considering convergence.

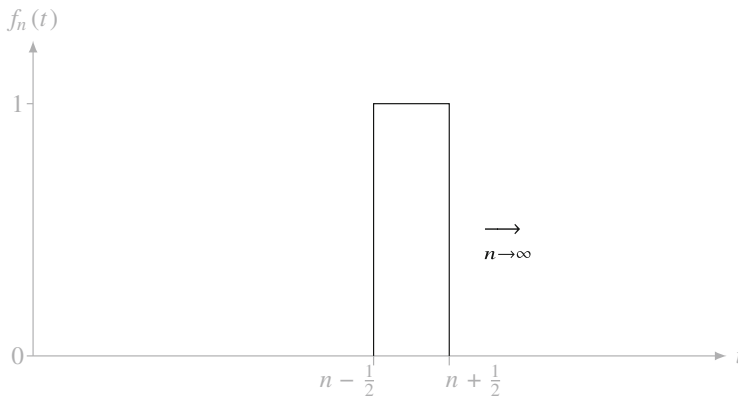


Fig. 7.8 An example with regard to the pointwise convergence

With this definition of convergence integration and limit can be swapped only under certain conditions.¹⁵

We will now present an example which demonstrates that the interchangeability of integration and limit is lost if one uses pointwise convergence. The expected value of the limit does not equal the limit of expectations.

Let us consider the state space $\Omega = \mathbb{R}$ and a function f_n which is zero on the real line except in the neighborhood of $n \in \mathbb{R}$. The area below the function should be exactly one. Figure 7.8 illustrates such a function that show a rectangle at index n . With increasing index the rectangle is moving to infinity.¹⁶

We look at this sequence of functions and apply the definition of pointwise convergence. Doing so we will show that the limit of this sequence is zero with the rectangle neither changing its form nor disappearing entirely. This might be surprising.

- The functions f_n converge pointwise to zero: consider a fixed value t . For t the following applies

$$\lim_{n \rightarrow \infty} f_n(t) = 0, \quad (7.14)$$

¹⁵Sufficient conditions are formulated in the theorem of monotone convergence. The theorem is due to Beppo Levi and can be found in any textbook on measure theory, for example, Rudin (1976), theorem 11.28.

¹⁶For example, consider $f_3(t)$ and $f_1(t)$. At $t = 1$ we have $f_3(t) = 0$ and $f_1(t) = 1$ and thus $f_3(t) \not\cong f_1(t)$.

because any index n will eventually be greater than t . This is why the following must hold:

$$\lim_{n \rightarrow \infty} f_n(t) = 0 \implies \int_{-\infty}^{\infty} \lim_{n \rightarrow \infty} f_n(t) dt = 0. \quad (7.15)$$

- On the other hand, the area under each function is 1 and therefore

$$\int_{-\infty}^{\infty} f_n(t) dt = \int_{-n}^n f_n(t) dt = n + \frac{1}{2} - \left(n - \frac{1}{2}\right) = 1, \quad (7.16)$$

and therefore

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(t) dt = \lim_{n \rightarrow \infty} 1 = 1. \quad (7.17)$$

Equations (7.15) and (7.17) show that one must not interchange integration and limit in the sequence of functions considered here. This conclusion can be expressed as

$$\lim_n \int \neq \int \lim_n. \quad (7.18)$$

For the reasons described above such a result is useless. We must therefore note that pointwise convergence is not an appropriate concept. Rather, it is advisable to find another concept of convergence which permits the interchangeability of integration and limit.

Mean Square Convergence This concept of convergence¹⁷ is used to ensure that expectation (i.e., expected value and variance) and limit can be interchanged. To this end we assume a measure space $(\Omega, \mathcal{F}, \mu)$. It is presupposed that there is a sequence of measurable functions f_n .

Mean square convergence measures the difference of a function (out of the sequence) and its limit. Mean square convergence is defined that the sequence converges if both the expectation and variance of this difference go to zero. The formal definition reads as follows.

Definition 7.2 (Mean Square Convergence) A sequence of measurable functions f_n converges in mean square to a function f

$$\lim_{n \rightarrow \infty} f_n = f, \quad (7.19)$$

¹⁷In the literature mean square convergence is also labeled as L^2 -convergence.

if and only if

$$\lim_{n \rightarrow \infty} \int_{\Omega} |f_n(\omega) - f(\omega)|^2 d\mu(\omega) = 0 \quad (7.20)$$

applies.

We will show that the mean square convergence ensures that integration and limit can be interchanged. For this we concentrate again on a probability measure, i.e., we consider random variables. We use the definition of mean square convergence and rely on the identity (5.36). Assume $\lim_{n \rightarrow \infty} f_n = f$. Thus we get from (7.20)

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \int_{\Omega} |f_n(\omega) - f(\omega)|^2 d\mu(\omega) \\ &= \lim_{n \rightarrow \infty} \left(\text{Var}[f_n - f] + E^2[f_n - f] \right) \\ &= \lim_{n \rightarrow \infty} \text{Var}[f_n - f] + \lim_{n \rightarrow \infty} E^2[f_n - f]. \end{aligned} \quad (7.21)$$

Since neither of the two summands can be negative, both $\lim_{n \rightarrow \infty} \text{Var}[f_n - f] = 0$ and $\lim_{n \rightarrow \infty} E^2[f_n - f] = 0$ apply. If the squared expectation is zero, $\lim_{n \rightarrow \infty} E[f_n - f] = 0$ must hold. The expectation is linear, and therefore $\lim_{n \rightarrow \infty} E[f_n] = E[f]$ is true. Thus $\lim_{n \rightarrow \infty} E[f_n] = E[\lim_{n \rightarrow \infty} f_n]$. That was what we had to show.

7.4 Conditional Expectations Are Random Variables

Finally, we want to draw the reader's attention to an aspect of conditional expectations that was originated by Kolmogoroff.¹⁸ So far we have realized that a conditional expectation is a real number that refers to an event A (the condition).¹⁹ The expectation depends on this event A . If we choose a different event, a different expectation will usually result. Therefore, Kolmogoroff has proposed that the conditional expectation should be interpreted as a random variable.²⁰

To understand this idea we need to remember how we had defined random variables. We wanted to perceive them as functions of elementary events. On page 83 we have shown that a random variable X can be characterized as a function

$$X : \Omega \rightarrow \mathbb{R} \quad (7.22)$$

¹⁸Andrei Nikolayevich Kolmogoroff (1903–1987), Russian mathematician.

¹⁹See page 80 ff.

²⁰See Kolmogoroff (1933), page 41 ff.

Table 7.1 States, cash flows CF_3 , and conditional expectations of the cash flows in the binomial model of Example 5.6

$\omega \in \Omega$	$CF_3 \in \mathbb{R}$	$E[CF_3 \mathcal{F}_2]$
uuu	140	} =135
uud	130	
udu	130	} =127.5
udd	125	
duu	130	} =127.5
dud	125	
ddu	125	} =82.5
ddd	40	

with its conditional expectation

$$E[X|\mathcal{F}] : \Omega \rightarrow \mathbb{R} \quad (7.23)$$

also being interpreted as a random variable. The following two examples will help to better understand this concept.

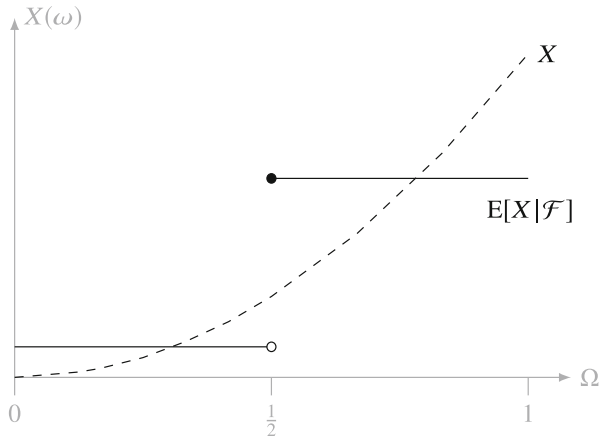
Example 7.3 (Binomial Model) With Table 7.1 we refer to Example 5.6 from page 83. While the first column of this table shows the states, the second column represents the cash flows CF_3 . The conditional expectation (at time $t = 2$) is given in the third column.

The σ -algebra \mathcal{F}_2 corresponds to the set of information that the decision-maker assumes today he will have available at the time $t = 2$. On the basis of this information the decision-maker forms his expectations. In Table 7.1 we have grouped by parentheses those states that cannot be discriminated at time $t = 2$. Let us call the combination of two such states a “box.” At time $t = 2$ he only knows which box he will be in but he cannot discriminate the states within the box.

Example 7.3 demonstrates the following: if a specific elementary event ω is given, the event $\{\omega\}$ and other elementary events are combined into a set A (the above-mentioned “box”). The set A contains only those elementary events that the decision-maker cannot discriminate from ω on the base of his information set given. In this example he was able to observe the uu node at $t = 2$ but did not (yet) know whether the state uuu or uud will occur at $t = 3$. The conditional expected value $E[X|\mathcal{F}]$ assigns the actual number $E[X|A]$ to the elementary event ω . To determine the conditional expected values, the payments associated with the elementary events are weighted with their respective probabilities of occurrence.

Example 7.4 (Share Price) To further deepen our reflections we consider a state space $\Omega = [0, 1]$. Each real number $\omega \in [0, 1]$ represents an elementary event. If we

Fig. 7.9 Illustration of the conditional expectation $E[X|\mathcal{F}]$



choose the Lebesgue measure²¹ λ with the corresponding σ -algebra, a probability space is generated since $\lambda(\Omega) = 1$ holds.

Let us consider the random variable

$$X(\omega) = \omega^2. \tag{7.24}$$

With the elementary event $\omega = \frac{1}{2}$ the random variable assumes the value $X(\omega) = \frac{1}{4}$. We present the path of this random variable in Fig. 7.9 as a dashed curve.

Let us determine the conditional expectation for the following σ -algebra

$$\mathcal{F} = \left\{ \emptyset, \left\{ \left[0, \frac{1}{2} \right) \right\}, \left\{ \left[\frac{1}{2}, 1 \right] \right\}, \{[0, 1]\} \right\}. \tag{7.25}$$

In this case the decision-maker cannot tell with certainty which specific elementary event $\omega \in [0, 1]$ is present; instead he receives only the information whether the elementary event is greater or less than $\frac{1}{2}$.²² This is all he knows. What is the conditional expectation of the random variable X ?

²¹See page 53.

²²For mathematical reasons, the second set in the σ -algebra must be a half-open interval. If we would add the set $[0, \frac{1}{2}]$ to the σ -algebra the intersection

$$\left[0, \frac{1}{2} \right] \cap \left[\frac{1}{2}, 1 \right] = \left\{ \frac{1}{2} \right\} \tag{7.26}$$

would also be measurable and the decision-maker could determine whether the state $\omega = \frac{1}{2}$ has occurred. But that would be more than we wanted to assume.

Concentrating on the first subinterval we get according to (5.37)²³ a conditional expectation of

$$E\left[X|\omega < \frac{1}{2}\right] = \frac{1}{\frac{1}{2}} \int_0^{\frac{1}{2}} X^2 d\lambda(\omega) = 2 \left[\frac{X^3}{3}\right]_0^{\frac{1}{2}} \quad (7.27)$$

and for the second subinterval

$$E\left[X|\omega > \frac{1}{2}\right] = \frac{1}{\frac{1}{2}} \int_{\frac{1}{2}}^1 X^2 d\lambda(\omega) = 2 \left[\frac{X^3}{3}\right]_{\frac{1}{2}}^1. \quad (7.28)$$

Thus, we can present the conditional expectation simply by

$$E[X|\mathcal{F}] = \begin{cases} \frac{1}{12}, & \text{if } \omega \in \left[0, \frac{1}{2}\right), \\ \frac{7}{12}, & \text{if } \omega \in \left[\frac{1}{2}, 1\right]. \end{cases} \quad (7.29)$$

Figure 7.9 shows the form of the conditional expectation which is a constant function with a jump at $\omega = \frac{1}{2}$.

As before we recognize the idea of conditional expectation. Beginning with an elementary event ω one must first determine the smallest set A which is part of the σ -algebra \mathcal{F} and also includes ω . The conditional expectation $E[X|A]$ is calculated using Eq. (5.37) and represents the value of the random variable $E[X|\mathcal{F}]$ at ω .

Finally, let us present the following rules for calculating for conditional expectations.

Expected value of known quantities If $X \in \mathcal{F}$ (it is also said that X is \mathcal{F} -measurable), then $E[X|\mathcal{F}] = X$ applies.

In order to illustrate the theorem imagine having to determine the conditional expectation of an uncertain quantity $X(\omega)$. However, the situation is such that the uncertain state ω can be derived directly from the observed value of the quantity X . Thus the observed quantity is not really uncertain, a result confirming the first theorem.

Further, if Z is \mathcal{F} -measurable and bounded, then $E[Z \cdot X|\mathcal{F}] = Z \cdot E[X|\mathcal{F}]$ holds.

Linearity For any numbers a, b the following is true: $E[aX + bY|\mathcal{F}] = aE[X|\mathcal{F}] + bE[Y|\mathcal{F}]$.

Since the conditional expectation represents a generalization of the classic (unconditional) expectation, the property of linearity remains valid. That is the substance of this theorem.

²³See page 83.

Monotonicity If $X \geq 0$, then $E[X|\mathcal{F}] \geq 0$ applies.

Since probabilities are nonnegative the expected value of nonnegative variables remains nonnegative. This applies to conditional expectations as well.

Limit almost everywhere If X_n is a monotonously growing sequence of random variables which converges to X almost everywhere and if X has a finite expectation, $\lim_{n \rightarrow \infty} E[X_n|\mathcal{F}] = E[X|\mathcal{F}]$ holds.

We had emphasized in Sect. 7.3 that the interchangeability of limit and expectation is of considerable importance in probability theory. This is one of the strengths of the concept of conditional expectation. Under certain conditions limit and expectation can be swapped using almost everywhere-convergence.

Iterated expectation If $\mathcal{F} \subset \mathcal{G}$, then $E[E[X|\mathcal{G}]|\mathcal{F}] = E[X|\mathcal{F}]$.

If iterated conditional expectations are to be calculated the inner expectation $E[X|\mathcal{G}]$ can be omitted.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

