Chapter 8
Sequences of Independent Random Variables. Limit Theorems

Abstract The chapter opens with proofs of Khintchin's (weak) Law of Large Numbers (Sect. 8.1) and the Central Limit Theorem (Sect. 8.2) the case of independent identically distributed summands, both using the apparatus of characteristic functions. Section 8.3 establishes general conditions for the Weak Law of Large Numbers for general sequences of independent random variables and also conditions for the respective convergence in mean. Section 8.4 presents the Central Limit Theorem in the triangular array scheme (the Lindeberg–Feller theorem) and its corollaries, illustrated by several insightful examples. After that, in Sect. 8.5 an alternative method of compositions is introduced and used to prove the Central Limit Theorem in the same situation, establishing an upper bound for the convergence rate for the uniform distance between the distribution functions in the case of finite third moments. This is followed by an extension of the above results to the multivariate case in Sect. 8.6. Section 8.7 presents important material not to be found in other textbooks: the so-called integro-local limit theorems on convergence to the normal distribution (the Stone–Shepp and Gnedenko theorems), including versions for sums of random variables depending on a parameter. These results will be of crucial importance in Chap. 9, when proving theorems on exact asymptotic behaviour of large deviation probabilities. The chapter concludes with Sect. 8.8 establishing integral, integro-local and local theorems on convergence of the distributions of scaled sums on independent identically distributed random variables to non-normal stable laws.

8.1 The Law of Large Numbers

Theorem 8.1.1 (Khintchin's Law of Large Numbers) *Let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of independent identically distributed random variables having a finite expectation $\mathbf{E}\xi_n = a$ and let $S_n := \xi_1 + \cdots + \xi_n$. Then*

$$\frac{S_n}{n} \xrightarrow{p} a \quad as \ n \to \infty.$$

The above assertion together with Theorems 6.1.6 and 6.1.7 imply the following.

Corollary 8.1.1 *Under the conditions of Theorem* 8.1.1, *as well as convergence of* S_n/n *in probability, convergence in mean also takes place*:

$$\mathbf{E}\left|\frac{S_n}{n} - a\right| \to 0 \quad as\ n \to \infty.$$

Note that the condition of independence of ξ_k and the very assertion of the theorem assume that all the random variables ξ_k are given on a common probability space.

From the physical point of view, the stated law of large numbers is the simplest ergodic theorem which means, roughly speaking, that for random variables their "time averages" and "space averages" coincide. This applies to an even greater extent to the strong law of large numbers, by virtue of which $S_n/n \to a$ with probability 1.

Under more strict assumptions (existence of variance) Theorem 8.1.1 was obtained in Sect. 4.7 as a consequence of Chebyshev's inequality.

Proof of Theorem 8.1.1 We have to prove that, for any $\varepsilon > 0$,

$$\mathbf{P}\left(\left|\frac{S_n}{n} - a\right| > \varepsilon\right) \to 0$$

as $n \to \infty$. The above relation is equivalent to the weak convergence of distributions $S_n/n \Longrightarrow \mathbf{I}_a$. Therefore, by the continuity theorem and Example 7.1.1 it suffices to show that, for any fixed t,

$$\varphi_{S_n/n}(t) \to e^{iat}.$$

The ch.f. $\varphi(t)$ of the random variable ξ_k has, in a certain neighbourhood of 0, the property $|\varphi(t) - 1| < 1/2$. Therefore for such t one can define the function $l(t) = \ln \varphi(t)$ (we take the principal value of the logarithm). Since ξ_n has finite expectation, the derivative

$$l'(0) = \frac{\varphi'(0)}{\varphi(0)} = ia$$

exists. For each fixed t and sufficiently large n, the value of $l(t/n)$ is defined and

$$\varphi_{S_n/n}(t) = \varphi^n(t/n) = e^{l(t/n)n}.$$

Since $l(0) = 0$, one has

$$e^{l(t/n)n} = \exp\left\{t\frac{l(t/n) - l(0)}{t/n}\right\} \to e^{l'(0)t} = e^{iat}$$

as $n \to \infty$. The theorem is proved.   □

8.2 The Central Limit Theorem for Identically Distributed Random Variables

Let, as before, $\{\xi_n\}$ be a sequence of independent identically distributed random variables. But now we assume, along with the expectation $\mathbf{E}\xi_n = a$, the existence of the variance $\text{Var}\,\xi_n = \sigma^2$. We retain the notation $S_n = \xi_1 + \cdots + \xi_n$ for sums of our random variables and $\Phi(x)$ for the normal distribution function with parameters $(0, 1)$. Introduce the sequence of random variables

$$\zeta_n = \frac{S_n - an}{\sigma\sqrt{n}}.$$

Theorem 8.2.1 *If* $0 < \sigma^2 < \infty$, *then* $\mathbf{P}(\zeta_n < x) \to \Phi(x)$ *uniformly in* x $(-\infty < x < \infty)$ *as* $n \to \infty$.

In such a case, the sequence $\{\zeta_n\}$ is said to be *asymptotically normal*.

It follows from $\zeta_n \Rightarrow \zeta \in \Phi_{0,1}$, $\zeta_n^2 \geq 0$, $\mathbf{E}\zeta_n^2 = \mathbf{E}\zeta^2 = 1$ and from Lemma 6.2.3 that the sequence $\{\zeta_n^2\}$ is uniformly integrable. Therefore, as well as the weak convergence $\zeta_n \Rightarrow \zeta$, $\zeta \in \Phi_{0,1}$ ($\mathbf{E}f(\zeta_n) \to \mathbf{E}f(\zeta)$ for any bounded continuous f), one also has convergence $\mathbf{E}f(\zeta_n) \to \mathbf{E}f(\zeta)$ for any continuous f such that $|f(x)| < c(1 + x^2)$ (see Theorem 6.2.3).

Proof of Theorem 8.2.1 The uniform convergence is a consequence of the weak convergence and continuity of $\Phi(x)$. Further, we may assume without loss of generality that $a = 0$, for otherwise we could consider the sequence $\{\xi'_n = \xi_n - a\}_{n=1}^{\infty}$ without changing the sequence $\{\zeta_n\}$. Therefore, to prove the required convergence, it suffices to show that $\varphi_{\zeta_n}(t) \to e^{-t^2/2}$ when $a = 0$. We have

$$\varphi_{\zeta_n}(t) = \varphi^n\left(\frac{t}{\sigma\sqrt{n}}\right), \quad \text{where } \varphi(t) = \varphi_{\xi_k}(t).$$

Since $\mathbf{E}\xi_n^2$ exists, $\varphi''(t)$ also exists and, as $t \to 0$, one has

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\varphi''(0) + o(t^2) = 1 - \frac{t^2\sigma^2}{2} + o(t^2). \tag{8.2.1}$$

Therefore, as $n \to \infty$,

$$\ln\varphi_{\zeta_n}(t) = n\ln\left[1 - \frac{\sigma^2}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + o\left(\frac{t^2}{n}\right)\right]$$

$$= n\left[-\frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right] = -\frac{t^2}{2} + o(1) \to -\frac{t^2}{2}.$$

The theorem is proved.   □

8.3 The Law of Large Numbers for Arbitrary Independent Random Variables

Now we proceed to elucidating conditions under which the law of large numbers and the central limit theorem will hold in the case when ξ_k are independent but not necessarily identically distributed. The problem will not become more complicated if, from the very beginning, we consider a more general situation where one is given an arbitrary series $\xi_{1,n}, \ldots, \xi_{n,n}$, $n = 1, 2, \ldots$ of independent random variables, where the distributions of $\xi_{k,n}$ may depend on n. This is the so-called *triangular array scheme*.

Put

$$\zeta_n := \sum_{k=1}^{n} \xi_{k,n}.$$

From the viewpoint of the results to follow, we can assume without loss of generality that

$$\mathbf{E}\xi_{k,n} = 0. \tag{8.3.1}$$

Assume that the following condition is met: as $n \to \infty$,

$$D_1 := \sum_{k=1}^{n} \mathbf{E} \min\left(|\xi_{k,n}|, |\xi_{k,n}|^2 \right) \to 0. \tag{$\mathbf{D}_1$}$$

Theorem 8.3.1 (The Law of Large Numbers) *If conditions* (8.3.1) *and* [$\mathbf{D}_1$] *are satisfied, then* $\zeta_n \Longrightarrow \mathbf{I}_0$ *or, which is the same,* $\zeta_n \xrightarrow{p} 0$ *as* $n \to \infty$.

Example 8.3.1 Assume $\xi_k = \xi_{k,n}$ do not depend on n, $\mathbf{E}\xi_k = 0$ and $\mathbf{E}|\xi_k|^s \le m_s < \infty$ for $1 < s \le 2$. For such s, there exists a sequence $b(n) = o(n)$ such that $n = o(b^s(n))$. Since, for $\xi_{k,n} = \xi_k/b(n)$,

$$\mathbf{E}\min\left(|\xi_{k,n}|, \xi_{k,n}^2\right) = \mathbf{E}\left[\left|\frac{\xi_k}{b(n)}\right|^2; |\xi_k| \le b(n)\right] + \mathbf{E}\left[\frac{|\xi_k|}{b(n)}; |\xi_k| > b(n)\right]$$

$$\le \mathbf{E}\left[\left|\frac{\xi_k}{b(n)}\right|^s; |\xi_k| \le b(n)\right] + \mathbf{E}\left[\left|\frac{\xi_k}{b(n)}\right|^s; |\xi_k| > b(n)\right]$$

$$= m_s b^{-s}(n),$$

we have

$$D_1 \le n m_s b^{-s}(n) \to 0,$$

and hence $S_n/b(n) \xrightarrow{p} 0$.

A more general sufficient condition (compared to $m_s < \infty$) for the law of large numbers is contained in Theorem 8.3.3 below. Theorem 8.1.1 is an evident corollary of that theorem.

Now consider condition $[\mathbf{D}_1]$ in more detail. It can clearly also be written in the form

$$D_1 = \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > 1\big) + \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|^2; |\xi_{k,n}| \le 1\big) \to 0.$$

Next introduce the condition

$$M_1 := \sum_{k=1}^{n} \mathbf{E}|\xi_{k,n}| \le c < \infty \tag{8.3.2}$$

and the condition

$$M_1(\tau) := \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > \tau\big) \to 0 \tag{M1}$$

for any $\tau > 0$ as $n \to \infty$. Condition $[\mathbf{M}_1]$ could be called a Lindeberg type condition (the Lindeberg condition $[\mathbf{M}_2]$ will be introduced in Sect. 8.4).

The following lemma explains the relationship between the introduced conditions.

Lemma 8.3.1 1. $\{[\mathbf{M}_1] \cap (3.2)\} \subset [\mathbf{D}_1]$. 2. $[\mathbf{D}_1] \subset [\mathbf{M}_1]$.

That is, conditions $[\mathbf{M}_1]$ and (8.3.2) imply $[\mathbf{D}_1]$, and condition $[\mathbf{D}_1]$ implies $[\mathbf{M}_1]$.

It follows from Lemma 8.3.1 that under condition (8.3.2), conditions $[\mathbf{D}_1]$ and $[\mathbf{M}_1]$ are equivalent.

Proof of Lemma 8.3.1 1. Let conditions (8.3.2) and $[\mathbf{M}_1]$ be met. Then, for

$$\tau \le 1, \qquad g_1(x) = \min\big(|x|, |x|^2\big),$$

one has

$$D_1 = \sum_{k=1}^{n} \mathbf{E} g_1(\xi_{k,n}) \le \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > \tau\big) + \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|^2; |\xi_{k,n}| \le \tau\big)$$

$$\le M_1(\tau) + \tau \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| \le \tau\big) \le M_1(\tau) + \tau M_1(0). \tag{8.3.3}$$

Since $M_1(0) = M_1 \le c$ and τ can be arbitrary small, we have $D_1 \to 0$ as $n \to \infty$.

2. Conversely, let condition $[\mathbf{D}_1]$ be met. Then, for $\tau \le 1$,

$$M_1(\tau) \le \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > 1\big)$$

$$+ \tau^{-1} \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|^2; \tau < |\xi_{k,n}| \le 1\big) \le \tau^{-1} D_1 \to 0 \qquad (8.3.4)$$

as $n \to \infty$ for any $\tau > 0$. The lemma is proved.   □

Let us show that condition [$\mathbf{M}_1$] (as well as [$\mathbf{D}_1$]) is essential for the law of large numbers to hold.

Consider the random variables

$$\xi_{k,n} = \begin{cases} 1 - \frac{1}{n} & \text{with probability } \frac{1}{n}, \\ -\frac{1}{n} & \text{with probability } 1 - \frac{1}{n}. \end{cases}$$

For them, $\mathbf{E}\xi_{k,n} = 0$, $\mathbf{E}|\xi_{k,n}| = \frac{2(n-1)}{n^2} \sim \frac{2}{n}$, $M_1 \le 2$, condition (8.3.2) is met, but $M_1(\tau) = \frac{n-1}{n} > \frac{1}{2}$ for $n > 2$, $\tau < 1/2$, and thus condition [$\mathbf{M}_1$] is not satisfied. Here the number ν_n of positive $\xi_{k,n}$, $1 \le k \le n$, converges in distribution to a random variable ν having the Poisson distribution with parameter $\lambda = 1$. The sum of the remaining $\xi_{k,n}$s is equal to $-\frac{(n-\nu_n)}{n} \xrightarrow{p} -1$. Therefore, $\zeta_n + 1 \Longleftrightarrow \mathbf{\Pi}_1$ and the law of large numbers does not hold.

Each of the conditions [$\mathbf{D}_1$] and [$\mathbf{M}_1$] imply the uniform smallness of $\mathbf{E}|\xi_{k,n}|$:

$$\max_{1 \le k \le n} \mathbf{E}|\xi_{k,n}| \to 0 \quad \text{as } n \to \infty. \qquad (8.3.5)$$

Indeed, equation [$\mathbf{M}_1$] means that there exists a sufficiently slowly decreasing sequence $\tau_n \to 0$ such that $M_1(\tau_n) \to 0$. Therefore

$$\max_{k \le n} \mathbf{E}|\xi_{k,n}| \le \max_{k \le n}\big[\tau_n + \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > \tau_n\big)\big] \le \tau_n + M_1(\tau_n) \to 0. \qquad (8.3.6)$$

In particular, (8.3.5) implies the negligibility of the summands $\xi_{k,n}$.

We will say that $\xi_{k,n}$ are *negligible*, or, equivalently, have property [$\mathbf{S}$], if, for any $\varepsilon > 0$,

$$\max_{k \le n} \mathbf{P}\big(|\xi_{k,n}| > \varepsilon\big) \to 0 \quad \text{as } n \to \infty. \qquad [\mathbf{S}]$$

Property [$\mathbf{S}$] could also be called *uniform convergence of $\xi_{k,n}$ in probability to zero*. Property [$\mathbf{S}$] follows immediately from (8.3.5) and Chebyshev's inequality. It also follows from stronger relations implied by [$\mathbf{M}_1$]:

$$\mathbf{P}\Big(\max_{k \le n} |\xi_{k,n}| > \varepsilon\Big) = \mathbf{P}\Big(\bigcup_{k \le n}\{|\xi_{k,n}| > \varepsilon\}\Big)$$
$$\le \sum_{k \le n} \mathbf{P}\big(|\xi_{k,n}| > \varepsilon\big) \le \varepsilon^{-1} \sum_{k \le n} \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > \varepsilon\big) \to 0. \quad [\mathbf{S}_1]$$

We now turn to proving the law of large numbers. We will give two versions of the proof. The first one illustrates the classical method of characteristic functions.

The second version is based on elementary inequalities and leads to a stronger assertion about convergence in mean.[1]

Here is the first version.

Proof of Theorem 8.3.1[2] Put

$$\varphi_{k,n}(t) := \mathbf{E}e^{it\xi_{k,n}}, \qquad \Delta_k(t) := \varphi_{k,n}(t) - 1.$$

One has to prove that, for each t,

$$\varphi_{\zeta_n}(t) = \mathbf{E}e^{it\zeta_n} = \prod_{k=1}^{n} \varphi_{k,n}(t) \to 1,$$

as $n \to \infty$. By Lemma 7.4.2

$$\left|\varphi_{\zeta_n}(t) - 1\right| = \left|\prod_{k=1}^{n} \varphi_{k,n}(t) - \prod_{k=1}^{n} 1\right| \le \sum_{k=1}^{n} |\Delta_k(t)|$$

$$= \sum_{k=1}^{n} \left|\mathbf{E}e^{it\xi_{k,n}} - 1\right| = \sum_{k=1}^{n} \left|\mathbf{E}\left(e^{it\xi_{k,n}} - 1 - it\xi_{k,n}\right)\right|.$$

By Lemma 7.4.1 we have (for $g_1(x) = \min(|x|, x^2)$)

$$\left|e^{itx} - 1 - itx\right| \le \min\left(2|tx|, t^2x^2/2\right) \le 2g_1(tx) \le 2h(t)g_1(t),$$

where $h(t) = \max(|t|, |t|^2)$. Therefore

$$\left|\varphi_{\zeta_n}(t) - 1\right| \le 2h(t) \sum_{k=1}^{n} \mathbf{E}g_1(\xi_{k,n}) = 2h(t)D_1 \to 0.$$

The theorem is proved. $\square$

The last inequality shows that $|\varphi_{\zeta_n}(t) - 1|$ admits a bound in terms of D_1. It turns out that $\mathbf{E}|\zeta_n|$ also admits a bound in terms of D_1. Now we will give the second version of the proof that actually leads to a stronger variant of the law of large numbers.

Theorem 8.3.2 *Under conditions* (8.3.1) *and* [**D**$_1$] *one has* $\mathbf{E}|\zeta_n| \to 0$ *(i.e.* $\zeta_n \xrightarrow{(1)} 0$).

---

[1]The second version was communicated to us by A.I. Sakhanenko.

[2]There exists an alternative "direct" proof of Theorem 8.3.1 using not ch.f.s but the so-called truncated random variables and estimates of their variances. However, because of what follows, it is more convenient for us to use here the machinery of ch.f.s.

The assertion of Theorem 8.3.2 clearly means *the uniform integrability of* $\{\zeta_n\}$; it implies Theorem 8.3.1, for

$$\mathbf{P}\big(|\zeta_n| > \varepsilon\big) \le \mathbf{E}|\zeta_n|/\varepsilon \to 0 \quad \text{as } n \to \infty.$$

Proof of Theorem 8.3.2 Put

$$\xi'_{k,n} := \begin{cases} \xi_{k,n} & \text{if } |\xi_{k,n}| \le 1, \\ 0 & \text{otherwise,} \end{cases}$$

and $\xi''_{k,n} := \xi_{k,n} - \xi'_{k,n}$. Then $\xi_{k,n} = \xi'_{k,n} + \xi''_{k,n}$ and $\zeta_n = \zeta'_n + \zeta''_n$ with an obvious convention for the notations ζ'_n, ζ''_n. By the Cauchy–Bunjakovsky inequality,

$$\mathbf{E}|\zeta_n| \le \mathbf{E}\big|\zeta'_n - \mathbf{E}\zeta'_n\big| + \mathbf{E}\big|\zeta''_n - \mathbf{E}\zeta''_n\big| \le \sqrt{\mathbf{E}\big(\zeta'_n - \mathbf{E}\zeta'_n\big)^2} + \mathbf{E}\big|\zeta''_n\big| + \big|\mathbf{E}\zeta''_n\big|$$

$$\le \sqrt{\sum \mathrm{Var}\big(\xi'_{k,n}\big)} + 2\sum \mathbf{E}\big|\xi''_{k,n}\big| \le \sqrt{\sum \mathbf{E}\big(\xi'_{k,n}\big)^2} + 2\sum \mathbf{E}\big|\xi''_{k,n}\big|$$

$$= \Big[\sum \mathbf{E}\big(\xi^2_{k,n}; |\xi_{k,n}| \le 1\big)\Big]^{1/2}$$

$$+ 2\sum \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > 1\big) \le \sqrt{D_1} + 2D_1 \to 0,$$

if $D_1 \to 0$. The theorem is proved. □

Remark 8.3.1 It can be seen from the proof of Theorem 8.3.2 that the argument will remain valid if we replace the independence of $\xi_{k,n}$ by the weaker condition that $\xi'_{k,n}$ are non-correlated. It will also be valid if $\xi'_{k,n}$ are only weakly correlated so that

$$\mathbf{E}\big(\zeta'_n - \mathbf{E}\zeta'_n\big)^2 \le c\sum \mathrm{Var}\big(\xi'_{k,n}\big), \quad c < \infty.$$

If $\{\xi_k\}$ is a given *fixed* (not dependent on n) sequence of independent random variables, $S_n = \sum_{k=1}^n \xi_k$ and $\mathbf{E}\xi_k = a_k$, then one looks at the applicability of the law of large numbers to the sequences

$$\xi_{k,n} = \frac{\xi_k - a_k}{b(n)}, \qquad \zeta_n = \sum \xi_{k,n} = \frac{1}{b(n)}\Big(S_n - \sum_{k=1}^n a_k\Big), \qquad (8.3.7)$$

where $\xi_{k,n}$ satisfy (8.3.1), and $b(n)$ is an unboundedly increasing sequence. In some cases it is natural to take $b(n) = \sum_{k=1}^n \mathbf{E}|\xi_k|$ if this sum increases unboundedly. Without loss of generality we can set $a_k = 0$. The next assertion follows from Theorem 8.3.2.

Corollary 8.3.1 *If, as $n \to \infty$,*

$$D_1 := \frac{1}{b(n)} \sum \mathbf{E}\min\big(|\xi_k|, \xi_k^2/b(n)\big) \to 0$$

or, for any $\tau > 0$,

$$M_1(\tau) = \frac{1}{b(n)} \sum \mathbf{E}\big(|\xi_k|; |\xi_k| > \tau b(n)\big) \to 0, \quad b(n) = \sum_{k=1}^{n} \mathbf{E}|\xi_k| \to \infty, \quad (8.3.8)$$

then $\zeta_n \xrightarrow{(1)} 0$.

Now we will present an important sufficient condition for the law of large numbers that is very close to condition (8.3.8) and which explains to some extent its essence. In addition, in many cases this condition is easier to check. Let $b_k = \mathbf{E}|\xi_k|$, $\overline{b}_n = \max_{k \leq n} b_k$, and, as before,

$$S_n = \sum_{k=1}^{n} \xi_k, \qquad b(n) = \sum_{k=1}^{n} b_k.$$

The following assertion is a direct generalisation of Theorem 8.1.1 and Corollary 8.1.1.

Theorem 8.3.3 *Let $\mathbf{E}\xi_k = 0$, the sequence of normalised random variables ξ_k/b_k be uniformly integrable and $\overline{b}_n = o(b(n))$ as $n \to \infty$. Then*

$$\frac{S_n}{b(n)} \xrightarrow{(1)} 0.$$

If $\overline{b}_n \leq b < \infty$ then $b(n) \leq bn$ and $\frac{S_n}{n} \xrightarrow{(1)} 0$.

Proof Since

$$\mathbf{E}\big(|\xi_k|; |\xi_k| > \tau b(n)\big) \leq b_k \mathbf{E}\left(\left|\frac{\xi_k}{b_k}\right|; \left|\frac{\xi_k}{b_k}\right| > \tau \frac{b(n)}{\overline{b}_n}\right) \qquad (8.3.9)$$

and $\frac{b(n)}{\overline{b}_n} \to \infty$, the uniform integrability of $\{\frac{\xi_k}{b_k}\}$ implies that the right-hand side of (8.3.9) is $o(b_k)$ uniformly in k (i.e. it admits a bound $\varepsilon(n)b_k$, where $\varepsilon(n) \to 0$ as $n \to \infty$ and does not depend on k). Therefore

$$M_1(\tau) = \frac{1}{b(n)} \sum_{k=1}^{n} \mathbf{E}\big(|\xi_k|; |\xi_k| > \tau b(n)\big) \to 0$$

as $n \to \infty$, and condition (8.3.8) is met. The theorem is proved. $\qquad\square$

Remark 8.3.2 If, in the context of the law of large numbers, we are interested in convergence in probability, only then can we generalise Theorem 8.3.3. In particular, *convergence*

$$\frac{S_n}{b(n)} \xrightarrow{p} 0$$

will still hold if a finite number of the summands ξ_k (e.g., for $k \leq l$, l being fixed) are completely arbitrary (they can even fail to have expectations) *and the sequence $\xi_k^* = \xi_{k+l}$, $k \geq 1$, satisfies the conditions of Theorem* 8.3.3, *where $b(n)$ is defined for the variables ξ_k^* and has the property $\frac{b(n-1)}{b(n)} \to 1$ as $n \to \infty$.*

This assertion follows from the fact that

$$\frac{S_n}{b(n)} = \frac{S_l}{b(n)} + \frac{S_n - S_l}{b(n-l)} \cdot \frac{b(n-l)}{b(n)}, \qquad \frac{S_l}{b(n)} \overset{p}{\longrightarrow} 0, \qquad \frac{b(n-l)}{b(n)} \to 1,$$

and by Theorem 8.3.3

$$\frac{S_n - S_l}{b(n-l)} \overset{p}{\longrightarrow} 0 \quad \text{as } n \to \infty.$$

Now we will show that the uniform integrability condition in Theorem 8.3.3 (as well as condition $M_1(\tau) \to 0$) is essential for convergence $\zeta_n \overset{p}{\to} 0$. Consider a sequence of random variables

$$\xi_j = \begin{cases} 2^s - 1 & \text{with probability } 2^{-s}, \\ -1 & \text{with probability } 1 - 2^{-s} \end{cases}$$

for $j \in I_s := (2^{s-1}, 2^s]$, $s = 1, 2, \ldots$; $\xi_1 = 0$. Then $\mathbf{E}\xi_j = 0$, $\mathbf{E}|\xi_j| = 2(1 - 2^{-s})$ for $j \in I_s$, and, for $n = 2^k$, one has

$$b(n) = \sum_{s=1}^{k} 2(1 - 2^{-s})|I_s|,$$

where $|I_s| = 2^s - 2^{s-1} = 2^{s-1}$ is the number of points in I_k. Hence, as $k \to \infty$,

$$b(n) \sim 2\left[(1 - 2^{-k})2^{k-1} + (1 - 2^{-k+1})2^{k-2} + \cdots\right]$$
$$\sim 2^k + 2^{k-1} + \ldots \sim 2^{k+1} = 2n.$$

Observe that the uniform integrability condition is clearly not met here. The distribution of the number $\nu^{(s)}$ of jumps of magnitude $2^s - 1$ on the interval I_s converges, as $s \to \infty$, to the Poisson distribution with parameter $1/2 = \lim_{s \to \infty} 2^{-s}|I_s|$, while the distribution of $2^{-s}(S_{2^s} - S_{2^{s-1}})$ converges to the distribution of $\nu - 1/2$, where $\nu \in \mathbf{\Pi}_{1/2}$. Hence, assuming that $n = 2^k$, and partitioning the segment $[2, n]$ into the intervals $(2^{s-1}, 2^s]$, $s = 1, \ldots, k$, we obtain that the distribution of S_n/n converges, as $k \to \infty$, to the distribution of

$$\frac{S_n}{n} = 2^{-k}\sum_{s=1}^{k} \frac{S_{2^s} - S_{2^{s-1}}}{2^s} 2^s \Rightarrow \sum_{l=0}^{\infty}(\nu_l - 1/2)2^{-l} =: \zeta,$$

where ν_l, $l = 0, 1, \ldots$, are independent copies of ν. Clearly, $\zeta \not\equiv 0$, and so convergence $\frac{S_n}{n} \xrightarrow{P} 0$ fails to take place.

Let us return to arbitrary $\xi_{k,n}$. In order for [$\mathbf{D}_1$] to hold it suffices that the following condition is met: *for some s, $2 \geq s > 1$,*

$$\sum_{k=1}^{n} \mathbf{E}|\xi_{k,n}|^s \to 0. \qquad [\mathbf{L}_s]$$

This assertion is evident, since $g_1(x) \leq |x|^s$ for $2 \geq s > 1$. Conditions [$\mathbf{L}_s$] could be called the modified Lyapunov conditions (cf. the Lyapunov condition [$\mathbf{L}_s$] in Sect. 8.4).

To prove Theorem 8.3.2, we used the so-called "truncated versions" $\xi'_{k,n}$ of the random variables $\xi_{k,n}$. Now we will consider yet another variant of the law of large numbers, in which *conditions* are expressed in terms of truncated random variables.

Denote by $\xi^{(N)}$ the result of truncation of the random variable ξ at level N:

$$\xi^{(N)} = \max\left[-N, \min(N, \xi)\right].$$

Theorem 8.3.4 *Let the sequence of random variables $\{\xi_k\}$ in* (8.3.7) *satisfy the following condition*: *for any given $\varepsilon > 0$, there exist N_k such that*

$$\frac{1}{b(n)} \sum_{k=1}^{n} \mathbf{E}\left|\xi_k - \xi_k^{(N_k)}\right| < \varepsilon, \qquad \frac{1}{b(n)} \sum_{k=1}^{n} N_k < N < \infty.$$

Then the sequence $\{\zeta_n\}$ converges to zero in mean: $\zeta_n \xrightarrow{(1)} 0$.

Proof Clearly $a_k^{(N_k)} := \mathbf{E}\xi_k^{(N_k)} \to a_k$ as $N_k \to \infty$ and $|a_k^{(N_k)}| \leq N_k$. Further, we have

$$\mathbf{E}|\zeta_n| = \frac{1}{b(n)}\mathbf{E}\left|\sum(\xi_k - a_k)\right| \leq \frac{1}{b(n)}\sum \mathbf{E}|\xi_k - \xi_k^{(N_k)}|$$

$$+ \mathbf{E}\left|\sum \frac{\xi_k^{(N_k)} - a_k^{(N_k)}}{b(n)}\right| + \frac{1}{b(n)}\sum|a_k^{(N_k)} - a_k|.$$

Here the second term on the right-hand side converges to zero, since the sum under the expectation satisfies the conditions of Theorem 8.3.1 and is bounded. But the first and the last terms do not exceed ε. Since the left-hand side does not depend on ε, we have $\mathbf{E}|\zeta_n| \to 0$ as $n \to \infty$. $\qquad \square$

Corollary 8.3.2 *If $b(n) = n$ and, for sufficiently large N and all $k \leq n$,*

$$\mathbf{E}\left|\xi_k - \xi_k^{(N)}\right| < \varepsilon,$$

then $\zeta_n \xrightarrow{(1)} 0$.

The corollary follows from Theorem 8.3.4, since the conditions of the corollary clearly imply the conditions of Theorem 8.3.4.

It is obvious that, for identically distributed ξ_k, the conditions of Corollary 8.3.2 are always met, and we again obtain a generalisation of Theorem 8.1.1 and Corollary 8.1.1.

If $\mathbf{E}|\xi_k|^r < \infty$ for $r \geq 1$, then we can also establish in a similar way that

$$\frac{S_n}{n} \xrightarrow{(r)} a.$$

Remark 8.3.3 Condition [**D**$_1$] (or [**M**$_1$]) is not necessary for convergence $\zeta_n \xrightarrow{p} 0$ even when (8.3.2) and (8.3.5) hold, as the following example demonstrates. Let $\xi_{k,n}$ assume the values $-n$, 0, and n with probabilities $1/n^2$, $1 - 2/n^2$, and $1/n^2$, respectively. Here $\zeta_n \xrightarrow{p} 0$, since $\mathbf{P}(\zeta_n \neq 0) \leq \mathbf{P}(\bigcup\{\xi_{k,n} \neq 0\}) \leq 2/n \to 0$, $\mathbf{E}|\xi_{k,n}| = 2/n \to 0$ and $M_1 = \sum \mathbf{E}|\xi_{k,n}| = 2 < \infty$. At the same time, $\sum \mathbf{E}(|\xi_{k,n}|; |\xi_{k,n}| \geq 1) = 2 \not\to \infty$, so that conditions [**D**$_1$] and [**M**$_1$] are not satisfied.

However, if we require that

$$\xi_{k,n} \geq -\varepsilon_{k,n}, \qquad \varepsilon_{k,n} \geq 0,$$

$$\max_{k \leq n} \varepsilon_{k,n} \to 0, \qquad \sum_{k=1}^{n} \varepsilon_{k,n} \leq c < \infty, \tag{8.3.10}$$

then condition [**D**$_1$] will become necessary for convergence $\zeta_n \xrightarrow{p} 0$.

Before proving that assertion we will establish several auxiliary relations that will be useful in the sequel. As above, put $\Delta_k(t) := \varphi_{k,n}(t) - 1$.

Lemma 8.3.2 *One has*

$$\sum_{k=1}^{n} |\Delta_k(t)| \leq |t| M_1.$$

If condition [**S**] *holds, then for each t, as $n \to \infty$,*

$$\max_{k \leq n} |\Delta_k(t)| \to 0.$$

If a random variable ξ with $\mathbf{E}\xi = 0$ is bounded from the left: $\xi > -c$, $c > 0$, then $\mathbf{E}|\xi| \leq 2c$.

Proof By Lemma 7.4.1,

$$\left|\Delta_k(t)\right| \leq \mathbf{E}\left|e^{it\xi_{k,n}} - 1\right| \leq |t|\,\mathbf{E}|\xi_{k,n}|, \qquad \sum \left|\Delta_k(t)\right| \leq |t| M_1.$$

Further,

$$\left|\Delta_k(t)\right| \leq \mathbf{E}\big(\left|e^{it\xi_{k,n}} - 1\right|; |\xi_{k,n}| \leq \varepsilon\big) + \mathbf{E}\big(\left|e^{it\xi_{k,n}} - 1\right|; |\xi_{k,n}| > \varepsilon\big)$$

$$\leq |t|\varepsilon + 2\mathbf{P}\big(|\xi_{k,n}| > \varepsilon\big).$$

Since ε is arbitrary here, the second assertion of the lemma now follows from condition [S].

Put

$$\xi^+ := \max(0; \xi) \geq 0, \qquad \xi^- := -\big(\xi - \xi^+\big) \geq 0.$$

Then $\mathbf{E}\xi = \mathbf{E}\xi^+ - \mathbf{E}\xi^- = 0$ and $\mathbf{E}|\xi| = \mathbf{E}\xi^+ + \mathbf{E}\xi^- = 2\mathbf{E}\xi^- \leq 2c$. The lemma is proved. $\square$

From the last assertion of the lemma it follows that (8.3.10) implies (8.3.2) and (8.3.5).

Lemma 8.3.3 *Let conditions* [S] *and* (8.3.2) *be satisfied. A necessary and sufficient condition for convergence* $\varphi_{\zeta_n}(t) \to \varphi(t)$ *is that*

$$\sum_{k=1}^{n} \Delta_k(t) \to \ln\varphi(t).$$

Proof Observe that

$$\operatorname{Re}\Delta_k(t) = \operatorname{Re}\big(\varphi_{k,n}(t) - 1\big) \leq 0, \quad \left|e^{\Delta_k(t)}\right| \leq 1,$$

and therefore, by Lemma 7.4.2,

$$\left|\varphi_{z_n}(t) - e^{\sum \Delta_k(t)}\right| = \left|\prod_{k=1}^{n}\varphi_{k,n}(t) - \prod_{k=1}^{n}e^{\Delta_k(t)}\right|$$

$$\leq \sum_{k=1}^{n}\left|\varphi_{k,n}(t) - e^{\Delta_k(t)}\right| = \sum_{k=1}^{n}\left|e^{\Delta_k(t)} - 1 - \Delta_k(t)\right|$$

$$\leq \frac{1}{2}\sum_{k=1}^{n}\Delta_k^2(t) \leq \frac{1}{2}\max_k\left|\Delta_k(t)\right|\sum_{k=1}^{n}\left|\Delta_k(t)\right|.$$

By Lemma 8.3.2 and conditions [S] and (8.3.2), the expression on the left-hand side converges to 0 as $n \to \infty$. Therefore, if $\varphi_{z_n}(t) \to \varphi(t)$ then $\exp\{\sum\Delta_k(t)\} \to \varphi(t)$, and vice versa. The lemma is proved. $\square$

The next assertion complements Theorem 8.3.1.

Theorem 8.3.5 *Assume that relations* (8.3.1) *and* (8.3.10) *hold. Then condition* [$\mathbf{D}_1$] *(or condition* [$\mathbf{M}_1$]*) is necessary for the law of large numbers.*

Proof If the law of large numbers holds then $\varphi_{z_n}(t) \to 1$ and, hence by Lemma 8.3.3 (recall that (8.3.10) implies (8.3.2), (8.3.5) and [$\mathbf{S}$])

$$\sum_{k=1}^{n} \Delta_k(t) = \sum_{k=1}^{n} \mathbf{E}\big(e^{it\xi_{k,n}} - 1 - it\xi_{k,n}\big) \to 0.$$

Moreover, by Lemma 7.4.1

$$\sum_{k=1}^{n} \mathbf{E}\big(\big|e^{it\xi_{k,n}} - 1 - it\xi_{k,n}\big|; |\xi_{k,n}| \le \varepsilon_{k,n}\big)$$

$$\le \frac{1}{2} \sum_{k=1}^{n} \mathbf{E}\big(|xi_{k,n}|^2; |\xi_{k,n}| \le \varepsilon_{k,n}\big) \le \sum_{k=1}^{n} \varepsilon_{k,n}^2 \le \max_k \varepsilon_{k,n} \sum_{k=1}^{n} \varepsilon_{k,n} \to 0.$$

Therefore, if the law of large numbers holds, then by virtue of (8.3.10)

$$\sum_{k=1}^{n} \mathbf{E}\big(e^{it\xi_{k,n}} - 1 - it\xi_{k,n}; \xi_{k,n} > \varepsilon_{k,n}\big) \to 0.$$

Consider the function $\alpha(x) = (e^{ix} - 1)/ix$. It is not hard to see that the inequality $|\alpha(x)| \le 1$ proved in Lemma 7.4.1 is strict for $x > \varepsilon > 0$, and hence there exists a $\delta(\tau) > 0$ for $\tau > 0$ such that $\mathrm{Re}(1 - \alpha(x)) \ge \delta(\tau)$ for $x > \tau$. This is equivalent to $\mathrm{Im}(1 + ix - e^{ix}) \ge \delta(\tau)x$, so that

$$x \le \frac{1}{\delta(\tau)} \mathrm{Im}\big(1 + ix - e^{ix}\big) \quad \text{for } x > \tau.$$

From this we find that

$$\mathbf{E}_1(\tau) = \sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|; |\xi_{k,n}| > \tau\big) = \sum_{k=1}^{n} \mathbf{E}(\xi_{k,n}; \xi_{k,n} > \tau)$$

$$\le \frac{1}{\delta(\tau)} \mathrm{Im} \sum_{k=1}^{n} \mathbf{E}\big(1 + i\xi_{k,n} - e^{i\xi_{k,n}}; \xi_{k,n} > \varepsilon_{k,n}\big) \to 0.$$

Thus condition [$\mathbf{M}_1$] holds. Together with relation (8.3.2), that follows from (8.3.10), this condition implies [$\mathbf{D}_1$]. The theorem is proved. $\square$

There seem to exist some conditions that are wider than (8.3.10) and under which condition $[\mathbf{D}_1]$ is necessary for convergence $\zeta_n \xrightarrow{(1)} 0$ in mean (condition (8.3.10) is too restrictive).

8.4 The Central Limit Theorem for Sums of Arbitrary Independent Random Variables

As in Sect. 8.3, we consider here a triangular array of random variables $\xi_{1,n}, \ldots, \xi_{n,n}$ and their sums

$$\zeta_n = \sum_{k=1}^{n} \xi_{k,n}. \tag{8.4.1}$$

We will assume that $\xi_{k,n}$ have finite second moments:

$$\sigma_{k,n}^2 := \mathrm{Var}(\xi_{k,n}) < \infty,$$

and suppose, without loss of generality, that

$$\mathbf{E}\xi_{k,n} = 0, \qquad \sum_{k=1}^{n} \sigma_{k,n}^2 = \mathrm{Var}(\zeta_n) = 1. \tag{8.4.2}$$

We introduce the following condition: *for some $s > 2$,*

$$D_2 := \sum_{k=1}^{n} \mathbf{E}\min\left(\xi_{k,n}^2, |\xi_{k,n}|^s\right) \to 0 \quad \text{as } n \to \infty, \tag{$\mathbf{D}_2$}$$

which is to play an important role in what follows. Our arguments related to condition $[\mathbf{D}_2]$ and also to conditions $[\mathbf{M}_2]$ and $[\mathbf{L}_s]$ to be introduced below will be quite similar to the ones from Sect. 8.3 that were related to conditions $[\mathbf{D}_1]$, $[\mathbf{M}_1]$ and $[\mathbf{L}_s]$.

We also introduce *the Lindeberg condition*: for any $\tau > 0$, as $n \to \infty$,

$$M_2(\tau) := \sum_{k=1}^{n} \mathbf{E}\left(|\xi_{k,n}|^2; |\xi_{k,n}| > \tau\right) \to 0. \tag{$\mathbf{M}_2$}$$

The following assertion is an analogue of Lemma 8.3.1.

Lemma 8.4.1 1. $\{[\mathbf{M}_2] \cap (4.2)\} \subset [\mathbf{D}_2]$. 2. $[\mathbf{D}_2] \subset [\mathbf{M}_2]$.

That is, conditions $[\mathbf{M}_2]$ and (8.4.2) imply $[\mathbf{D}_2]$, and condition $[\mathbf{D}_2]$ implies $[\mathbf{M}_2]$.

From Lemma 8.4.1 it follows that, under condition (8.4.2), conditions $[\mathbf{D}_2]$ and $[\mathbf{M}_2]$ are equivalent.

Proof of Lemma 8.4.1 1. Let conditions [$\mathbf{M}_2$] and (8.4.2) be met. Put

$$g_2(x) := \min\left(x^2, |x|^s\right), \quad s > 2.$$

Then (cf. (8.3.3), (8.3.4); $\tau \le 1$)

$$D_2 = \sum_{k=1}^{n} \mathbf{E}g_2(\xi_{k,n}) \le \sum_{k=1}^{n} \mathbf{E}\left(\xi_{k,n}^2; |\xi_{k,n}| > \tau\right) + \sum_{k=1}^{n} \mathbf{E}\left(|\xi_{k,n}|^s; |\xi_{k,n}| \le \tau\right)$$

$$\le M_2(\tau) + \tau^{s-2} M_2(0) = M_2(\tau) + \tau^{s-2}.$$

Since τ is arbitrary, we have $D_2 \to 0$ as $n \to \infty$.

2. Conversely, suppose that [$\mathbf{D}_2$] holds. Then

$$M_2(\tau) \le \sum_{k=1}^{n} \mathbf{E}\left(\xi_{k,n}^2; |\xi_{k,n}| > 1\right) + \frac{1}{\tau^{s-2}} \sum_{k=1}^{n} \left(|\xi_{k,n}|^s; \tau < |\xi_{k,n}| \le 1\right) \le \frac{1}{\tau^{s-2}} D_2 \to 0$$

for any $\tau > 0$, as $n \to \infty$. The lemma is proved. $\square$

Lemma 8.4.1 also implies that *if* (8.4.2) *holds, then condition* [$\mathbf{D}_2$] *is "invariant" with respect to $s > 2$.*

Condition [$\mathbf{D}_2$] can be stated in a more general form:

$$\sum_{k=1}^{n} \mathbf{E}\xi_{k,n}^2 h\left(|\xi_{k,n}|\right) \to 0,$$

where $h(x)$ is any function for which $h(x) > 0$ for $x > 0$, $h(x) \uparrow$, $h(x) \to 0$ as $x \to 0$, and $h(x) \to c < \infty$ as $x \to \infty$. All the key properties of condition [$\mathbf{D}_2$] will then be preserved. The Lindeberg condition clarifies the meaning of condition [$\mathbf{D}_2$] from a somewhat different point of view. In Lindeberg's condition, $h(x) = I_{(\tau,\infty)}$, $\tau \in (0, 1)$. A similar remark may be made with regard to conditions [$\mathbf{D}_1$] and [$\mathbf{M}_1$] in Sect. 8.3.

In a way similar to what we did in Sect. 8.3 when discussing condition [$\mathbf{M}_1$], one can easily verify that condition [$\mathbf{M}_2$] implies convergence (see (8.3.6))

$$\max_{k \le n} \mathrm{Var}(\xi_{k,n}) \to 0 \tag{8.4.3}$$

and the negligibility of $\xi_{k,n}$ (property [$\mathbf{S}$]). Moreover, one obviously has the inequality

$$M_1(\tau) \le \frac{1}{\tau} M_2(\tau).$$

For a given fixed (independent of n) sequence $\{\xi_k\}$ of independent random variables,

$$S_n = \sum_{k=1}^{\infty} \xi_k, \qquad \mathbf{E}\xi_k = a_k, \qquad \mathrm{Var}(\xi_k) = \sigma_k^2, \tag{8.4.4}$$

one considers the asymptotic behaviour of the normed sums

$$\zeta_n = \frac{1}{B_n}\left(S_n - \sum_{k=1}^{\infty} a_k\right), \qquad B_n^2 = \sum_{k=1}^{\infty} \sigma_k^2, \tag{8.4.5}$$

that are clearly also of the form (8.4.1) with $\xi_{k,n} = (\xi_k - a_k)/B_n$.

Conditions [$\mathbf{D}_1$] and [$\mathbf{M}_2$] for ξ_k will take the form

$$D_2 = \frac{1}{B_n^2} \sum_{k=1}^{\infty} \mathbf{E} \min\left((\xi_k - a_k)^2, \frac{|\xi_k - a_k|^s}{B_n^{s-2}}\right) \to 0, \quad s > 2;$$

$$M_2(\tau) = \frac{1}{B_n^2} \sum_{k=1}^{\infty} \mathbf{E}\big((\xi_k - a_k)^2; |\xi_k - a_k| > \tau B_n\big) \to 0, \quad \tau > 0. \tag{8.4.6}$$

Theorem 8.4.1 (The Central Limit Theorem) *If the sequences of random variables $\{\xi_{k,n}\}_{k=1}^{\infty}$, $n = 1, 2, \ldots$, satisfy conditions (8.4.2) and [$\mathbf{D}_2$] (or [$\mathbf{M}_2$]) then, as $n \to \infty$, $\mathbf{P}(\zeta_n < x) \to \Phi(x)$ uniformly in x.*

Proof It suffices to verify that

$$\varphi_{\zeta_n}(t) = \prod_{k=1}^{\infty} \varphi_{k,n}(t) \to e^{-t^2/2}.$$

By Lemma 7.4.2,

$$\left|\varphi_{\zeta_n}(t) - e^{-t^2/2}\right| = \left|\prod_{k=1}^{n} \varphi_{k,n}(t) - \prod_{k=1}^{n} e^{-t^2\sigma_{k,n}^2/2}\right|$$

$$\leq \sum_{k=1}^{n}\left|\varphi_{k,n}(t) - e^{-t^2\sigma_{k,n}^2/2}\right|$$

$$\leq \sum_{k=1}^{n}\left|\varphi_{k,n}(t) - 1 + \frac{1}{2}t^2\sigma_{k,n}^2\right|$$

$$+ \sum_{k=1}^{n}\left|e^{-t^2\sigma_{k,n}^2/2} - 1 + \frac{1}{2}t^2\sigma_{k,n}^2\right|. \tag{8.4.7}$$

Since by Lemma 7.4.1, for $s \leq 3$,

$$\left|e^{ix} - 1 - ix + \frac{x^2}{2}\right| \leq \min\left(x^2, \frac{|x^3|}{6}\right) \leq g_2(x)$$

(see the definition of the function g_2 in the beginning of the proof of Lemma 8.4.1), the first sum on the right-hand side of (8.4.7) does not exceed

$$\sum_{k=1}^{\infty}\left|\mathbf{E}\left(e^{it\xi_{k,n}} - 1 - it\xi_{k,n} + \frac{1}{2}t^2\xi_{k,n}^2\right)\right|$$

$$\leq \sum_{k=1}^{\infty}\mathbf{E}g_2\left(|t\xi_{k,n}|\right) \leq h(t)\sum_{k=1}^{\infty}\mathbf{E}g_2\left(|\xi_{k,n}|\right) \leq h(t)D_2 \to 0,$$

where $h(t) = \max(t^2, |t|^3)$. The last sum in (8.4.7) (again by Lemma 7.4.1) does not exceed (see (8.4.2) and (8.4.3))

$$\frac{t^4}{8}\sum_{k=1}^{n}\sigma_{k,n}^4 \leq \frac{t^4}{8}\max_k\sigma_{k,n}^2\sum_{k=1}^{n}\sigma_{k,n}^2 \leq \frac{t^4}{8}\max_k\sigma_{k,n}^2 \to 0 \quad \text{as } n \to \infty.$$

The theorem is proved. □

If we change the second relation in (8.4.2) to $\mathbf{E}\zeta_n \to \sigma^2 > 0$, then, introducing the new random variables $\xi_{k,n}' = \xi_{k,n}/\sqrt{\text{Var}\,\zeta_n}$ and using continuity theorems, it is not hard to obtain from Theorem 8.4.1 (see e.g. Lemma 6.2.2), the following assertion, which sometimes proves to be more useful in applications than Theorem 8.4.1.

Corollary 8.4.1 *Assume that* $\mathbf{E}\xi_{k,n} = 0$, $\text{Var}(\zeta_n) \to \sigma^2 > 0$, *and condition* [**D₂**] *(or* [**M₂**]*) is satisfied. Then* $\zeta_n \Longrightarrow \mathbf{\Phi}_{0,\sigma_2}$.

Remark 8.4.1 A sufficient condition for [**D₂**] and [**M₂**] is provided by the more restrictive *Lyapunov condition*, the verification of which is sometimes easier. Assume that (8.4.2) holds. For $s > 2$, the quantity

$$L_s := \sum_{k=1}^{n}\mathbf{E}|\xi_{k,n}|^2$$

is called the *Lyapunov fraction of the s-th order*. The condition

$$L_s \to 0 \quad \text{as } n \to \infty \qquad\qquad [\mathbf{L}_s]$$

is called the *Lyapunov condition*.

The quantity L_s is called a *fraction* since for $\xi_{k,n} = (\xi_k - a)/B_n$ (where $a_k = \mathbf{E}\xi_k$, $B_n^2 = \sum_{k=1}^{n}\text{Var}(\xi_k)$ and ξ_k do not depend on n), it has the form

$$L_s = \frac{1}{B_n^s}\sum_{k=1}^{n}\mathbf{E}|\xi_k - a_k|^s.$$

If the ξ_k are identically distributed, $a_k = a$, $\mathrm{Var}(\xi_k) = \sigma^2$, and $\mathbf{E}|\xi_k - a|^s = \mu < \infty$, then

$$L_s = \frac{\mu}{\sigma^s n^{(s-2)/2}} \to 0.$$

The sufficiency of the Lyapunov condition follows from the obvious inequalities $g_2(x) \leq |x|^s$ for any s, $D_2 \leq L_s$.

In the case of (8.4.4) and (8.4.5) we can give a sufficient condition for the integral limit theorem that is very close to the Lindeberg condition [$\mathbf{M}_2$]; the former condition elucidates to some extent the essence of the latter (cf. Theorem 8.3.3), and in many cases it is easier to verify. Put $\overline{\sigma}_n = \max_{k \leq n} \sigma_k$. Theorem 8.4.1 implies the following assertion which is a direct extension of Theorem 8.2.1

Theorem 8.4.2 *Let conditions* (8.4.4) *and* (8.4.5) *be satisfied, the sequence of normalised random variables* ξ_k^2/σ_k^2 *be uniformly integrable and* $\overline{\sigma}_n = o(B_n)$ *as* $n \to \infty$. *Then* $\zeta_n \Longrightarrow \mathbf{\Phi}_{0,1}$.

Proof of Theorem 8.4.2 repeats, to some extent, the proof of Theorem 8.3.3. For simplicity assume that $a_k = 0$. Then

$$\mathbf{E}\big(\xi_k^2; |\xi_k| > \tau B_n\big) \leq \sigma_k^2 \mathbf{E}\bigg(\frac{\xi_k^2}{\sigma_k^2}; \bigg|\frac{\xi_k}{\sigma_k}\bigg| > \tau \frac{B_n}{\overline{\sigma}_n}\bigg), \qquad (8.4.8)$$

where $B_n/\overline{\sigma}_n \to \infty$. Hence, it follows from the uniform integrability of $\{\frac{\xi_k^2}{\sigma_k^2}\}$ that the right-hand side of (8.4.8) is $o(\sigma_k^2)$ uniformly in k. This means that

$$M_2(\tau) = \frac{1}{B_n^2} \sum_{k=1}^n \mathbf{E}\big(\xi_k^2; |\xi_k| > \tau B_n\big) \to 0$$

as $n \to \infty$ and condition (8.4.6) (or condition [$\mathbf{M}_2$]) is satisfied. The theorem is proved. $\square$

Remark 8.4.2 We can generalise the assertion of Theorem 8.4.2 (cf. Remark 8.3.3). In particular, *convergence* $\zeta_n \Longrightarrow \mathbf{\Phi}_{0,1}$ *still takes place if a finite number of summands* ξ_k *(e.g., for* $k \leq l$, *l being fixed) are completely arbitrary, and the sequence* $\xi_k^* := \xi_{k+l}, k \geq 1$, *satisfies the conditions of Theorem* 8.4.2, *in which we put* $\sigma_k^2 = \mathrm{Var}(\xi_k^*)$, $B_n^2 = \sum_{k=1}^n \sigma_k^2$, *and it is also assumed that* $\frac{B_{n-1}}{B_n} \to 1$ *as* $n \to \infty$.

This assertion follows from the fact that

$$\frac{S_n}{B_n} = \frac{S_l}{B_n} + \frac{S_n - S_l}{B_{n-l}} \cdot \frac{B_{n-l}}{B_n},$$

where $\frac{S_l}{B_n} \xrightarrow{p} 0$, $\frac{B_{n-l}}{B_n} \to 1$ and, by Theorem 8.4.2, $\frac{S_n - S_l}{B_{n-l}} \Longrightarrow \mathbf{\Phi}_{0,1}$ as $n \to \infty$.

Remark 8.4.3 The uniform integrability condition that was used in Theorem 8.4.2 can be used for the triangular array scheme as well. In this more general case the uniform integrability should mean the following: *the sequences* $\eta_{1,n}, \ldots, \eta_{n,n}$, $n = 1, 2, \ldots$, *in the triangular array scheme are uniformly integrable if there exists a function* $\varepsilon(N) \downarrow 0$ *as* $N \uparrow \infty$ *such that, for all* n,

$$\max_{j \leq n} \mathbf{E}\big(|\eta_{j,n}|; |\eta_{j,n}| > N\big) \leq \varepsilon(N).$$

It is not hard to see that, with such an interpretation of uniform integrability, the assertion of Theorem 8.4.2 holds true for the triangular array scheme as well provided that the sequence $\{\frac{\xi_{j,n}^2}{\sigma_{j,n}^2}\}$ is uniformly integrable and $\max_{j \leq n} \sigma_{j,n} = o(1)$ as $n \to \infty$.

Example 8.4.1 We will clarify the difference between the Lindeberg condition and uniform integrability of $\{\frac{\xi_k^2}{\sigma_k^2}\}$ in the following example. Let η_k be independent bounded identically distributed random variables, $\mathbf{E}\,\eta_k = 0$, $\mathbf{D}\eta_k = 1$ and $g(k) > \sqrt{2}$ be an arbitrary function. Put

$$\xi_k := \begin{cases} \eta_k & \text{with probability } 1 - 2g^{-2}(k), \\ \pm g(k) & \text{with probability } g^{-2}(k). \end{cases}$$

Then clearly $\mathbf{E}\xi_k = 0$, $\sigma_k^2 := \mathbf{D}\xi_k = 3 - 2g^{-2}(k) \in (2, 3)$ and $B_n^2 \in (2n, 3n)$. The uniform integrability of $\{\frac{\xi_k^2}{\sigma_k^2}\}$, or the uniform integrability of $\{\xi_k^2\}$ which means the same in our case, excludes the case where $g(k) \to \infty$ as $k \to \infty$. The Lindeberg condition is wider and allows the growth of $g(k)$, except for the case where $g(k) > c\sqrt{k}$. If $g(k) = o(\sqrt{k})$, then the Lindeberg condition is satisfied because, for any fixed $\tau > 0$,

$$\mathbf{E}\big(\xi^2; |\xi_k| > \tau\sqrt{k}\big) = 0$$

for all large enough k.

Remark 8.4.4 Let us show that condition $[\mathbf{M}_2]$ (or $[\mathbf{D}_2]$) is essential for the central limit theorem. Consider random variables

$$\xi_{k,n} = \begin{cases} \pm\frac{1}{\sqrt{2}} & \text{with probability } \frac{1}{n}, \\ 0 & \text{with probability } 1 - \frac{2}{n}. \end{cases}$$

They satisfy conditions (8.4.2), $[\mathbf{S}]$, but not the Lindeberg condition as $M_2(\tau) = 1$ for $\tau < \frac{1}{\sqrt{2}}$. The number ν_k of non-zero summands converges in distribution to a random variable ν having the Poisson distribution with parameter 2. Therefore, ζ_n will clearly converge in distribution not to the normal law, but to $\sum_{j=1}^{\nu} \gamma_j$, where γ_j are independent and take values ± 1 with probability $1/2$.

Note also that conditions $[\mathbf{D}_2]$ or $[\mathbf{M}_2]$ are not necessary for convergence of the distributions of ζ_n to the normal distribution. Indeed, consider the following example: $\xi_{1,n} \Subset \mathbf{\Phi}_{0,1}$, $\xi_{2,n} = \cdots = \xi_{n,n} = 0$. Conditions (8.4.2) are clearly met, $\mathbf{P}(\zeta_n < x) = \Phi(x)$, but the variables $\xi_{k,n}$ are not negligible and therefore do not satisfy conditions $[\mathbf{D}_2]$ and $[\mathbf{M}_2]$.

If, however, as well as convergence $\zeta_n \Longrightarrow \mathbf{\Phi}_{0,1}$ we require that the $\xi_{k,n}$ are negligible, then conditions $[\mathbf{D}_2]$ and $[\mathbf{M}_2]$ become necessary.

Theorem 8.4.3 *Suppose that the sequences of independent random variables* $\{\xi_{k,n}\}_{k=1}^{n}$ *satisfy conditions (8.4.2) and* $[\mathbf{S}]$. *Then condition* $[\mathbf{D}_1]$ *(or* $[\mathbf{M}_2]$*) is necessary and sufficient for convergence* $\zeta_n \Longrightarrow \mathbf{\Phi}_{0,1}$.

First note that the assertions of Lemmas 8.3.2 and 8.3.3 remain true, up to some inessential modifications, if we substitute conditions (8.3.2) and $[\mathbf{S}]$ with (8.4.2) and $[\mathbf{S}]$.

Lemma 8.4.2 *Let conditions (8.4.2) and* $[\mathbf{S}]$ *hold. Then* $(\Delta_k(t) = \varphi_{k,n}(t) - 1)$

$$\max_{k \le n} |\Delta_k(t)| \to 0, \quad \sum |\Delta_k(t)| \le \frac{t^2}{2},$$

and the assertion of Lemma 8.3.3, that the convergence (8.3.10) is necessary and sufficient for convergence $\varphi_{\zeta_n}(t) \to \varphi(t)$, *remain completely true.*

Proof We can retain all the arguments in the proofs of Lemmas 8.3.2 and 8.3.3 except for one place where $\sum |\Delta_k(t)|$ is bounded. Under the new conditions, by Lemma 7.4.1, we have

$$|\Delta_k(t)| = |\varphi_{k,n}(t) - 1 - it\mathbf{E}\xi_{k,n}| \le \mathbf{E}|e^{it\xi_{k,n}} - 1 - it\xi_{k,n}| \le \frac{t^2}{2}\mathbf{E}\xi_{k,n}^2,$$

so that

$$\sum |\Delta_k(t)| \le \frac{t^2}{2}.$$

No other changes in the proofs of Lemmas 8.3.2 and 8.3.3 are needed. $\square$

Proof of Theorem 8.4.3 Sufficiency is already proved. To prove necessity, we make use of Lemma 8.4.1. If $\varphi_{\zeta_n}(t) \to e^{-t^2/2}$, then by virtue of that lemma, for $\Delta_k(t) = \varphi_{k,n}(t) - 1$, one has

$$\sum_{k=1}^{n} \Delta_k(t) \to \ln\varphi(t) = -\frac{t^2}{2}.$$

For $t = 1$ the above relation can be written in the form

$$R_n := \sum_{k=1}^{n} \mathbf{E}\left(e^{i\xi_{k,n}} - 1 - i\xi_{k,n} + \frac{1}{2}\xi_{k,n}^2\right) \to 0. \tag{8.4.9}$$

Put $\alpha(x) := (e^{ix} - 1 - ix)/x^2$. It is not hard to see that the inequality $|\alpha(x)| \leq 1/2$ proved in Lemma 7.4.1 is strict for $x \neq 0$, and

$$\sup_{|x| \geq \tau} |\alpha(x)| < \frac{1}{2} - \delta(\tau),$$

where $\delta(\tau) > 0$ for $\tau > 0$. This means that, for $|x| \geq \tau > 0$,

$$\mathrm{Re}\left[\alpha(x) + \frac{1}{2}\right] \geq \delta(\tau) > 0, \qquad x^2 \leq \frac{1}{\delta(\tau)} \mathrm{Re}\left(e^{ix} - 1 - ix + \frac{x^2}{2}\right),$$

$$\mathbf{E}\left(\xi_{k,n}^2; |\xi_{k,n}| > \tau\right) \leq \frac{1}{\delta(\tau)} \mathrm{Re}\,\mathbf{E}\left(e^{i\xi_{k,n}} - 1 - i\xi_{k,n} + \frac{\xi_{k,n}^2}{2}\right),$$

and hence by virtue of (8.4.9), for any $\tau > 0$,

$$M_2(\tau) \leq \frac{1}{\delta(\tau)}|R_n| \to 0$$

as $n \to \infty$. The theorem is proved.   □

Corollary 8.4.2 *Assume that* (8.4.2) *holds and*

$$\max_{k \leq n} \mathrm{Var}(\xi_{k,n}) \to 0. \tag{8.4.10}$$

Then a necessary and sufficient condition for convergence $\zeta_n \Longrightarrow \mathbf{\Phi}_{0,1}$ *is that*

$$\eta_n := \sum_{k=1}^{n} \xi_{k,n}^2 \Longrightarrow \mathbf{I}_1$$

(or that $\eta_n \overset{p}{\to} 1$*).*

Proof Let $\eta_n \Longrightarrow \mathbf{I}_1$. The random variables $\xi_{k,n}' = \xi_{k,n}^2 - \sigma_{k,n}^2$ satisfy, by virtue of (8.4.10), condition (8.3.10) and satisfy the law of large numbers:

$$\sum_{k=1}^{n} \xi_{k,n}' = \eta_n - 1 \overset{p}{\to} 0.$$

Therefore, by Theorem 8.3.5, the $\xi_{k,n}'$ satisfy condition $[\mathbf{M}_1]$: for any $\tau > 0$,

$$\sum_{k=1}^{n} \mathbf{E}\left(\left|\xi_{k,n}^2 - \sigma_{k,n}^2\right|; \left|\xi_{k,n}^2 - \sigma_{k,n}^2\right| > \tau\right) \to 0. \tag{8.4.11}$$

But by (8.4.10) this condition is clearly equivalent to condition $[\mathbf{M}_2]$ for $\xi_{k,n}$, and hence $\zeta_n \Longrightarrow \mathbf{\Phi}_{0,1}$.

Conversely, if $\zeta_n \Longrightarrow \Phi_{0,1}$, then $[\mathbf{M}_2]$ holds for $\xi_{k,n}$ which implies (8.4.11). Since, moreover,

$$\sum_{k=1}^{n} \mathbf{E}\left|\xi_{k,n}'\right| \leq 2 \sum_{k=1}^{n} \mathrm{Var}(\xi_{k,n}) = 2,$$

relation (8.3.2) holds for $\xi_{k,n}'$, and by Theorem 8.3.1

$$\sum_{k=1}^{n} \xi'_{k,n} = \eta_n - 1 \xrightarrow{p} 0.$$

The corollary is proved. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Example 8.4.2 Let ξ_k, $k = 1, 2, \ldots$, be independent random variables with distributions

$$\mathbf{P}\big(\xi_k = k^\alpha\big) = \mathbf{P}\big(\xi_k = -k^\alpha\big) = \frac{1}{2}.$$

Evidently, ξ_k can be represented as $\xi_k = k^\alpha \eta_k$, where $\eta_k \stackrel{d}{=} \eta$ are independent,

$$\mathbf{P}(\eta = 1) = \mathbf{P}(\eta = -1) = \frac{1}{2}, \qquad \mathrm{Var}(\eta) = 1, \qquad \sigma_k^2 = \mathrm{Var}(\xi_k) = k^{2\alpha}.$$

Let us show that, for all $\alpha \geq -1/2$, the random variables S_n/B_n are asymptotically normal. Since

$$\frac{\xi_k^2}{\sigma_k^2} \stackrel{d}{=} \eta^2$$

are uniformly integrable, by Theorem 8.4.2 it suffices to verify the condition

$$\overline{\sigma}_n = \max_{k \leq n} \sigma_k = o(B_n).$$

In our case $\overline{\sigma}_n = \max(1, n^{2\alpha})$ and, for $\alpha > -1/2$,

$$B_n^2 = \sum_{k=1}^{n} k^{2\alpha} \sim \int_0^n x^{2\alpha}\, d\alpha = \frac{n^{2\alpha+1}}{2\alpha + 1}.$$

For $\alpha = -1/2$, one has

$$B_n^2 = \sum_{k=1}^{n} k^{-1} \sim \ln n.$$

Clearly, in these cases $\overline{\sigma}_n = o(B_n)$ and the asymptotical normality of S_n/n holds.

If $\alpha < -1/2$ then the sequence B_n converges, condition $\overline{\sigma}_n = 1 = o(B_n)$ is not satisfied and the asymptotical normality of S_n/B_n fails to take place.

Note that, for $\alpha = -1/2$, the random variable

$$S_n = \sum_{k=1}^{n} \frac{\eta_k}{\sqrt{k}}$$

will be "comparable" with $\sqrt{\ln n}$ with a high probability, while the sums

$$\sum_{k=1}^{n} \frac{(-1)^k}{\sqrt{k}}$$

converge to a constant.

A rather graphical and well-known illustration of the above theorems is the scattering of shells when shooting at a target. The fact is that the trajectory of a shell is influenced by a large number of independent factors of which the individual effects are small. These are deviations in the amount of gun powder, in the weight and size of a shell, variations in the humidity and temperature of the air, wind direction and velocities at different altitudes and so on. As a result, the deviation of a shell from the aiming point is described by the normal law with an amazing accuracy.

Similar observations could be made about errors in measurements when their accuracy is affected by many "small" factors. (There even exists a theory of errors of which the crucial element is the central limit theorem.)

On the whole, the central limit theorem has a lot of applications in various areas. This is due to its *universality and robustness* under small deviations from the assumptions of the theorem, and its relatively *high accuracy* even for moderate values of n. The first two noted qualities mean that:

(1) the theorem is applicable to variables $\xi_{k,n}$ with *any distributions* so long as the variances of $\xi_{k,n}$ exist and are "negligible";

(2) the presence of a "moderate" dependence[3] between $\xi_{k,n}$ does not change the normality of the limiting distribution.

To illustrate the accuracy of the normal approximation, consider the following example. Let $F_n(x) = \mathbf{P}(S_n/\sqrt{n} < x)$ be the distribution function of the normalised sum S_n of independent variables ξ_k uniformly distributed over $[-\sqrt{3}, \sqrt{3}]$, so that $\mathrm{Var}(\xi_k) = 1$. Then it turns out that already for $n = 5$ (!) the maximum of $|F_n(x) - \Phi(x)|$ over the whole axis of x-values does not exceed 0.006 (the maximum is attained near the points $x = \pm 0.7$).

And still, despite the above circumstances, one has to be careful when applying the central limit theorem. For instance, one cannot expect high accuracy from the normal approximation when estimating probabilities of rare events, say when studying large deviation probabilities (this issue has already been discussed in Sect. 5.3).

---

[3]There exist several conditions characterising admissible dependence of $\xi_{k,n}$. Such considerations are beyond the scope of the present book, but can be found in the special literature. See e.g. [20].

After all, the theorem only ensures the smallness of the difference

$$\left|\Phi(x) - \mathbf{P}(\zeta < x)\right| \tag{8.4.12}$$

for large n. Suppose we want to use the normal approximation to find an x_0 such that the event $\{\zeta_n > x_0\}$ would occur on average once in 1000 trials (a problem of this sort could be encountered by an experimenter who wants to ensure that, in a single experiment, such an event will not occur). Even if the difference (8.4.12) does not exceed 0.02 (which can be a good approximation) then, using the normal approximation, we risk making a serious error. It can turn out, say, that $1 - \Phi(x_0) = 10^{-3}$ while $\mathbf{P}(\zeta < x) \approx 0.02$, and then the event $\{\zeta_n > x_0\}$ will occur much more often (on average, once in each 50 trials).

In Chap. 9 we will consider the problem of large deviation probabilities that enables one to handle such situations. In that case one looks for a function $P(n, x)$ such that $\mathbf{P}(\zeta < x)/P(n, x) \to 1$ as $n \to \infty$, $x \to \infty$. The function $P(n, x)$ turns out to be, generally speaking, different from $1 - \Phi(x)$. We should note however that using the approximation $P(n, x)$ requires more restrictive conditions on $\{\xi_{k,n}\}$.

In Sect. 8.7 we will consider the so-called *integro-local* and *local limit theorems* that establish convergence of the density of ζ_n to that of the normal law and enables one to estimate probabilities of rare events of another sort—say, of the form $\{a < \zeta_n < b\}$ where a and b are close to each other.

8.5[*] Another Approach to Proving Limit Theorems. Estimating Approximation Rates

The approach to proving the principal limit theorems for the distributions of sums of random variables that we considered in Sects. 8.1–8.4 was based on the use of ch.f.s. However, this is by far not the only method of proof of such assertions. Nowadays there exist several rather simple proofs of both the laws of large numbers and the central limit theorem that do not use the apparatus of ch.f.s. (This, however, does not belittle that powerful, well-developed, and rather universal tool.) Moreover, these proofs sometimes enable one to obtain more general results. As an illustration, we will give below a proof of the central limit theorem that extends, in a certain sense, Theorems 8.4.1 and 8.4.3 and provides an estimate of the convergence rate (although not the best one).

Along with the random variables $\xi_{k,n}$ in the triangular array scheme under assumption (8.4.2), consider mutually independent and independent of the sequence $\{\xi_{k,n}\}_{k=1}^{n}$ random variables $\eta_{k,n} \Subset \Phi_{0,\sigma_{k,n}^2}$, $\sigma_{k,n} := \mathrm{Var}(\xi_{k,n})$, so that

$$\eta_n := \sum_{k=1}^{n} \eta_{k,n} \Subset \Phi_{0,1}.$$

Set[4]

$$\mu_{k,n} := \mathbf{E}|\xi_{k,n}|^3, \qquad \nu_{k,n} := \mathbf{E}|\eta_{k,n}|^3 = c_3\sigma_{k,n}^3 \le c_3\mu_{k,n},$$

$$\mu_{k,n}^0 := \int |x|^3 |d(F_{k,n}(x) - \Phi_{k,n}(x))| \le \mu_{k,n} + \nu_{k,n},$$

$$L_3 := \sum_{k=1}^n \mu_{k,n}, \qquad N_3 := \sum_{k=1}^n \nu_{k,n}, \qquad L_3^0 := \sum_{k=1}^n \mu_{k,n}^0 \le L_3 + N_3 \le (1 + c_3)L_3.$$

Here $F_{k,n}$ and $\Phi_{k,n}$ are the distribution functions of $\xi_{k,n}$ and $\eta_{k,n}$, respectively. The quantities L_3 and N_3 are the third order Lyapunov fractions for the sequences $\{\xi_{k,n}\}$ and $\{\eta_{k,n}\}$. The quantities $\mu_{k,n}^0$ are called the third order *pseudomoments* and L_s^0 the Lyapunov fractions for pseudomoments. Clearly, $N_3 \le c_3 L_3 \to 0$, provided that the Lyapunov condition holds. As we have already noted, for $\xi_{k,n} = (\xi_k - a_k)/B_n$, where $a_k = \mathbf{E}\xi_k$, $B_n^2 = \sum_1^n \text{Var}(\xi_k)$, and ξ_k do not depend on n, one has

$$L_3 = \frac{1}{B_n^3} \sum_{k=1}^n \mu_k, \qquad \mu_k = \mathbf{E}|\xi_k - a_k|^3.$$

If, moreover, ξ_k are identically distributed, then

$$L_3 = \frac{\mu_1}{\sigma^3 \sqrt{n}}.$$

Our first task here is to estimate the closeness of $\mathbf{E}f(\zeta_n)$ to $\mathbf{E}f(\eta_n)$ for sufficiently smooth f. This problem could be of independent interest. Assume that f belongs to the class C_3 of all bounded functions with uniformly continuous and bounded third derivatives: $\sup_x |f^{(3)}(x)| \le f_3$.

Theorem 8.5.1 *If $f \in C_3$ then*

$$\left|\mathbf{E}f(\zeta_n) - \mathbf{E}f(\eta_n)\right| \le \frac{f_3 L_3^0}{6} \le \frac{f_3}{6}(L_3 + N_3). \qquad (8.5.1)$$

Proof Put, for $1 < l \le n$,

$$X_l := \xi_{1,n} + \cdots + \xi_{l-1,n} + \eta_{l,n} + \cdots + \eta_{n,n},$$

$$Z_l := \xi_{1,n} + \cdots + \xi_{l-1,n} + \eta_{l+1,n} + \cdots + \eta_{n,n},$$

$$X_1 := \eta_n, \qquad X_{n+1} = \zeta_n.$$

Then

$$X_{l+1} = Z_l + \xi_{l,n}, \qquad X_l = Z_l + \eta_{l,n}, \qquad (8.5.2)$$

---

[4]If $\eta \in \Phi_{0,1}$ then $c_3 = \mathbf{E}|\eta|^3 = \frac{2}{\sqrt{2\pi}} \int_0^\infty x^3 e^{-x^2/2} dx = \frac{4}{\sqrt{2\pi}} \int_0^\infty t e^{-t} dt = \frac{4}{\sqrt{2\pi}}.$

$$f(\zeta_n) - f(\eta_n) = \sum_{l=1}^{n}\big[f(X_{l+1}) - f(X_l)\big]. \tag{8.5.3}$$

Now we will make use of the following lemma.

Lemma 8.5.1 *Let $f \in C_3$ and Z, ξ and η be independent random variables with*

$$\mathbf{E}\xi = \mathbf{E}\eta = a, \qquad \mathbf{E}\xi^2 = \mathbf{E}\eta^2 = \sigma^2, \qquad \mu^0 = \int |x^3|\big|d\big(F_\xi(x) - F_\eta(x)\big)\big| < \infty.$$

Then

$$\big|\mathbf{E}f(Z+\xi) - \mathbf{E}f(Z+\eta)\big| \le \frac{f_3\mu^0}{6}. \tag{8.5.4}$$

Applying this lemma to (8.5.3), we get

$$\big|\mathbf{E}\big[f(X_{l+1}) - f(X_1)\big]\big| \le \frac{f_3\mu^0}{6}$$

which after summation gives (8.5.1). The theorem is proved.   □

Thus to complete the argument proving Theorem 8.5.1 it remains to prove Lemma 8.5.1.

Proof of Lemma 8.5.1 Set $g(x) := \mathbf{E}f(Z + x)$. It is evident that g, being the result of the averaging of f, has all the smoothness properties of f and, in particular, $|g'''(x)| \le f_3$. By virtue of the independence of Z, ξ and η, we have

$$\mathbf{E}f(Z+\xi) - \mathbf{E}f(Z+\eta) = \int g(x)\,d\big(F_\xi(x) - F_\eta(x)\big). \tag{8.5.5}$$

For the integrand, we make use of the expansion

$$g(x) = g(0) + xg'(0) + \frac{x^2}{2}g''(0) + \frac{x^3}{6}g'''(\theta_x), \qquad \theta_x \in [0, x].$$

Since the first and second moments of ξ coincide with those of η, we obtain for the right-hand side of (8.5.5) the bound

$$\left|\frac{1}{6}x^3 g'''(\theta_x)\,d\big(F_\xi(x) - F_\eta(x)\big)\right| \le \frac{f_3\mu^0}{6}.$$

The lemma is proved.   □

Remark 8.5.1 In exactly the same way one can establish the representation

$$\big|\mathbf{E}f(\zeta_n) - \mathbf{E}f(\eta_n)\big| \le \frac{g'''(0)}{6}\sum_{k=1}^{n}\mathbf{E}\big(\xi_{k,n}^3 - \eta_{k,n}^3\big) + \frac{f_4 L_4^0}{24}, \tag{8.5.6}$$

under obvious conventions for the notations f_4 and L_4^0. This bound can improve upon (8.5.1) if the differences $\mathbf{E}(\xi_{k,n}^3 - \eta_{k,n}^3)$ are small. If, for instance, $\xi_{k,n} = (\xi_k - a)/(\sigma\sqrt{n})$, ξ_k are identically distributed, and the third moments of $\xi_{k,n}$ and $\eta_{k,n}$ coincide, then on the right-hand side of (8.5.6) we will have a quantity of the order $1/n$.

Theorem 8.5.1 extends Theorem 8.4.1 in the case when $s = 3$. The extension is that, to establish convergence $\zeta_n \Longrightarrow \Phi_{0,1}$, one no longer needs the negligibility of $\xi_{k,n}$. If, for example, $\xi_{1,n} \subseteqq \Phi_{0,1/2}$ (in that case $\mu_{1,n}^0 = 0$) and $L_3^0 \to 0$, then $\mathbf{E}f(\zeta_n) \to \mathbf{E}f(\eta)$, $\eta \subseteqq \Phi_{0,1}$, for any f from the class C_3. Since C_3 is a distribution determining class (see Chap. 6), it remains to make use of Corollary 6.3.2.

We can strengthen the above assertion.

Theorem 8.5.2 *For any $x \in \mathbb{R}$,*

$$\left|\mathbf{P}(\zeta_n < x) - \Phi(x)\right| \le c\left(L_3^0\right)^{1/4}, \tag{8.5.7}$$

where c is an absolute constant.

Proof Take an arbitrary function $h \in C_3$, $0 \le h \le 1$, such that $h(x) = 1$ for $x \le 0$ and $h(x) = 0$ for $x \ge 1$, and put $h_3 = \sup_x |h'''(x)|$. Then, for the function $f(x) = h((x-t)/\varepsilon)$, we will have $f_3 = \sup_x |f'''(x)| \le h_3/\varepsilon^3$, and by Theorem 8.5.1

$$\mathbf{P}(\zeta_n < t) \le \mathbf{E}f(\zeta_n) \le \mathbf{E}f(\eta) + \frac{f_3 L_3^0}{6}$$

$$\le \mathbf{P}(\eta < t + \varepsilon) + \frac{h_3 L_3^0}{6\varepsilon^3} \le \mathbf{P}(\eta < t) + \frac{\varepsilon}{\sqrt{2\pi}} + \frac{h_3 L_3^0}{6\varepsilon^3}.$$

The last inequality holds since the maximum of the derivative of the normal distribution function $\Phi(t) = \mathbf{P}(\eta < t)$ is equal to $1/\sqrt{2\pi}$. Establishing in the same way the converse inequality and putting $\varepsilon = (L_3^0)^{1/4}$, we arrive at (8.5.7). The theorem is proved. □

The bound in Theorem 8.5.2 is, of course, not the best one. And yet inequality (8.5.7) shows that we will have a good normal approximation for $\mathbf{P}(\zeta_n < x)$ in the large deviations range (i.e. for $|x| \to \infty$) as well—at least for those x for which

$$\left(1 - \Phi\left(|x|\right)\right)\left(L_3^0\right)^{-1/4} \to \infty \tag{8.5.8}$$

as $n \to \infty$. Indeed, in that case, say, for $x = |x| > 0$,

$$\left|\frac{\mathbf{P}(\zeta_n \ge x)}{1 - \Phi(x)} - 1\right| \le \frac{c(L_3^0)^{1/4}}{1 - \Phi(x)} \to 0.$$

Since by L'Hospital's rule

$$1 - \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} \, dt \sim \frac{1}{\sqrt{2\pi}x} e^{-x^2/2} \quad \text{as } x \to \infty,$$

(8.5.8) holds for $|x| < c_1 \sqrt{-\ln L_3^0}$ with an appropriately chosen constant c_1.

In Chap. 20 we will obtain an extension of Theorems 8.5.1 and 8.5.2.

The problem of refinements and approximation rate bounds in the central limit theorem and other limit theorems is one of the most important in probability theory, because solving it will tell us how precise and efficient the applications of these theorems to practical problems will be. First of all, one has to find the true order of the decay of

$$\Delta_n = \sup_x \left| \mathbf{P}(\zeta_n < x) - \Phi(x) \right|$$

in n (or, say, in L_3 in the case of non-identically distributed variables). There exist at least two approaches to finding sharp bounds for Δ_n. The first one, the so-called *method of characteristic functions*, is based on the unimprovable bound for the closeness of the ch.f.s

$$\left| \ln \varphi_{\zeta_n}(t) + \frac{t^2}{2} \right| < cL_3$$

that the reader can obtain by him/herself, using Lemma 7.4.1 and somewhat modifying the argument in the proof of Theorem 8.4.1. The principal technical difficulties here are in deriving, using the inversion formula, the same order of smallness for Δ_n.

The second approach, the so-called *method of compositions*, has been illustrated in the present section in Theorem 8.5.1 (the idea of the method is expressed, to a certain extent, by relation (8.5.3)). It will be using just that method that we will prove in Appendix 5 the following general result (Cramér–Berry–Esseen):

Theorem 8.5.3 *If $\xi_{k,n} = (\xi_k - a_k)/B_n$, where ξ_k do not depend on n, then*

$$\sup_x \left| \mathbf{P}(\zeta_n < x) - \Phi(x) \right| \leq cL_3,$$

where c is an absolute constant.

In the case of identically distributed ξ_k the right-hand side of the above inequality becomes $c\mu_1/(\sigma^3 \sqrt{n})$. It was established that in this case $(2\pi)^{-1/2} < c < 0.4774$, while in the case of non-identically distributed summands $c < 0.5591$.[5]

One should keep in mind that the above theorems and the bounds for the constant c are universal and therefore hold under the most unfavourable conditions (from the point of view of the approximation). In real problems, the convergence rate is usually much better.

---

[5] See [33].

8.6 The Law of Large Numbers and the Central Limit Theorem in the Multivariate Case

In this section we assume that $\xi_{1,n}, \ldots, \xi_{n,n}$ are random vectors in the triangular array scheme,

$$\mathbf{E}\xi_{k,n} = 0, \qquad \zeta_n = \sum_{k=1}^{n} \xi_{k,n}.$$

The law of large numbers $\zeta_n \xrightarrow{p} 0$ follows immediately from Theorem 8.3.1, if we assume that the components of $\xi_{k,n}$ satisfy the conditions of that theorem. Thus we can assume that Theorem 8.3.1 was formulated and proved for vectors.

Dealing with the central limit theorem is somewhat more complicated. Here we will assume that $\mathbf{E}|\xi_{k,n}|^2 < \infty$, where $|x|^2 = (x, x)$ is square of the norm of x. Let

$$\sigma_{k,n}^2 := \mathbf{E}\xi_{k,n}^T \xi_{k,n}, \qquad \sigma_n^2 := \sum_{k=1}^{n} \sigma_{k,n}^2$$

(the superscript T denotes transposition, so that $\xi_{k,n}^T$ is a column vector).

Introduce the condition

$$\sum_{k=1}^{n} \mathbf{E}\min\big(|\xi_{k,n}|^2, |\xi_{k,n}|^s\big) \to 0, \quad s > 2, \qquad \qquad [\mathbf{D}_2]$$

and the Lindeberg condition

$$\sum_{k=1}^{n} \mathbf{E}\big(|\xi_{k,n}|^2; |\xi_{k,n}| > \tau\big) \to 0 \qquad \qquad [\mathbf{M}_2]$$

as $n \to \infty$ for any $\tau > 0$. As in the univariate case, we can easily verify that conditions $[\mathbf{D}_2]$ and $[\mathbf{M}_2]$ are equivalent provided that $\mathrm{tr}\,\sigma_n^2 := \sum_{j=1}^{d} (\sigma_n^2)_{jj} < c < \infty$.

Theorem 8.6.1 *If $\sigma_n^2 \to \sigma^2$, where σ^2 is a positive definite matrix, and condition $[\mathbf{D}_2]$ (or $[\mathbf{M}_2]$) is met, then*

$$\zeta_n \Longrightarrow \Phi_{0,\sigma^2}.$$

Corollary 8.6.1 ("The conventional" central limit theorem) *If $\xi_1, \xi_2, \ldots$ is a sequence of independent identically distributed random vectors, $\mathbf{E}\xi_k = 0$, $\sigma^2 = \mathbf{E}\xi_k^T \xi_k$ and $S_n = \sum_{k=1}^{n} \xi_k$ then, as $n \to \infty$,*

$$\frac{S_n}{\sqrt{n}} \Longrightarrow \Phi_{0,\sigma^2}.$$

This assertion is a consequence of Theorem 8.6.1, since the random variables $\xi_{k,n} = \xi_k/\sqrt{n}$ satisfy its conditions.

Proof of Theorem 8.6.1 Consider the characteristic functions

$$\varphi_{k,n}(t) := \mathbf{E}e^{i(t,\xi_{k,n})}, \qquad \varphi_n(t) := \mathbf{E}e^{i(t,\zeta_n)} = \prod_{k=1}^{n} \varphi_{k,n}(t).$$

In order to prove the theorem we have to verify that, for any t, as $n \to \infty$,

$$\varphi_n(t) \to \exp\left\{-\frac{1}{2}t\sigma^2 t^T\right\}.$$

We make use of Theorem 8.4.1. We can interpret $\varphi_{k,n}(t)$ and $\varphi_n(t)$ as the ch.f.s

$$\varphi_{k,n}^{\theta}(v) = \mathbf{E}\exp(iv\xi_{k,n}^{\theta}), \qquad \varphi_n^{\theta}(v) = \mathbf{E}\exp(iv\zeta_n^{\theta})$$

of the random variables $\xi_{k,n}^{\theta} = (\xi_{k,n}, \theta)$, $\zeta_n^{\theta} = (\zeta_n, \theta)$, where $\theta = t/|t|$, $v = |t|$. Let us show that the scalar random variables $\xi_{k,n}^{\theta}$ satisfy the conditions of Theorem 8.4.1 (or Corollary 8.4.1) for the univariate case. Clearly,

$$\mathbf{E}\xi_{k,n}^{\theta} = 0, \qquad \sum_{k=1}^{n} \mathbf{E}\left(\xi_{k,n}^{\theta}\right)^2 = \sum_{k=1}^{n} \mathbf{E}(\xi_{k,n}, \theta)^2 = \theta\sigma_n^2\theta^T \to \theta\sigma^2\theta^T > 0.$$

That condition [**D**$_2$] is satisfied follows from the obvious inequalities

$$(\xi_{k,n}, \theta)^2 = \left|\xi_{k,n}^{\theta}\right|^2 \le |\xi_{k,n}|^2, \qquad \sum_{k=1}^{n} \mathbf{E}\,g_2\left(\xi_{k,n}^{\theta}\right) \le \sum_{k=1}^{n} \mathbf{E}g_2\left(|\xi_{k,n}|\right),$$

where $g_2(x) = \min(x^2, |x|^s)$, $s > 2$. Thus, for any v and θ (i.e., for any t), by Corollary 8.4.1 of Theorem 8.4.1

$$\varphi_n(t) = \mathbf{E}\exp\left\{iv\zeta_n^{\theta}\right\} \to \exp\left\{-\frac{1}{2}v^2\theta\sigma^2\theta^T\right\} = \exp\left\{-\frac{1}{2}t\sigma^2 t^T\right\}.$$

The theorem is proved. $\square$

Theorem 8.6.1 does not cover the case where the entries of the matrix σ_n^2 grow unboundedly or behave in such away that the rank of the limiting matrix σ^2 becomes less than the dimension of the vectors $\xi_{k,n}$. This can happen when the variances of different components of $\xi_{k,n}$ have different orders of decay (or growth). In such a case, one should consider the transformed sums $\zeta_n' = \zeta_n\sigma_n^{-1}$ instead of ζ_n. Theorem 8.6.1 is actually a consequence of the following more general assertion which, in turn, follows from Theorem 8.6.1.

Theorem 8.6.2 *If the random variables $\xi_{k,n}' = \xi_{k,n}\sigma_n^{-1}$ satisfy condition* [**D**$_2$] *(or* [**M**$_2$]*) then $\zeta_n' \Longrightarrow \Phi_{0,E}$, where E is the identity matrix.*

8.7 Integro-Local and Local Limit Theorems for Sums of Identically Distributed Random Variables with Finite Variance

Theorem 8.2.1 from Sect. 8.2 is called the *integral* limit theorem. To understand the reasons for using such a name, one should compare this assertion with (more accurate) limit theorems of another type, that describe the asymptotic behaviour of the *densities* of the distributions of S_n (if any) or the asymptotics of the probabilities of sums S_n hitting a fixed interval. It is natural to call the theorems for densities *local* theorems. Theorems similar to Theorem 8.2.1 can be obtained from the local ones (if the densities exist) by integrating, and it is natural to call them *integral theorems*. Assertions about the asymptotics of the probabilities of S_n hitting an interval are "intermediate" between the local and integral theorems, and it is natural to call them *integro-local theorems*. In the literature, such statements are often also referred to as *local*, apparently because they describe the probability of the localisation of the sum S_n in a given interval.

8.7.1 Integro-Local Theorems

Integro-local theorems describe the asymptotics of

$$\mathbf{P}\big(S_n \in [x, x + \Delta)\big)$$

as $n \to \infty$ for a fixed $\Delta > 0$. Probabilities of this type for increasing Δ (or for $\Delta = \infty$) can clearly be obtained by summing the corresponding probabilities for fixed Δ.

We will derive integro-local and local theorems with the inversion formulas from Sect. 8.7.2.

For the sake of brevity, put

$$\Delta[x] = [x, x + \Delta)$$

and denote by $\phi(x) = \phi_{0,1}(x)$ the density of the standard normal distribution. Below we will restrict ourselves to the investigation of the sums $S_n = \xi_1 + \cdots + \xi_n$ of independent identically distributed random variables $\xi_k \stackrel{d}{=} \xi$.

Theorem 8.7.1 (The Stone–Shepp integro-local theorem) *Let ξ be a non-lattice random variable, $\mathbf{E}\xi = 0$ and $\mathbf{E}\xi^2 = \sigma^2 < \infty$. Then, for any fixed $\Delta > 0$, as $n \to \infty$,*

$$\mathbf{P}\big(S_n \in \Delta[x]\big) = \frac{\Delta}{\sigma\sqrt{n}} \phi\left(\frac{x}{\sigma\sqrt{n}}\right) + o\left(\frac{1}{\sqrt{n}}\right), \qquad (8.7.1)$$

where the remainder term $o(1/\sqrt{n})$ is uniform in x.

Remark 8.7.1 Since relation (8.7.1) is valid for any fixed Δ, it will also be valid when $\Delta = \Delta_n \to 0$ slowly enough as $n \to \infty$. If $\Delta = \Delta_n$ grows then the asymptotics of $\mathbf{P}(S_n \in \Delta[x))$ can be obtained by summing the right-hand sides of (8.7.1) for, say, $\Delta = 1$ (if $\Delta_n \to \infty$ is integer-valued). Thus the integral theorem follows from the integro-local one but not vice versa.

Remark 8.7.2 By virtue of the properties of densities (see Sect. 3.2), the right-hand side of representation (8.7.1) has the same form as if the random variable $\zeta_n = S_n/(\sigma\sqrt{n})$ had the density $\phi(v) + o(1)$, although the existence of the density of S_n (or ζ_n) is not assumed in the theorem.

Proof of Theorem 8.7.1 First prove the theorem under the simplifying assumption that condition

$$\limsup_{|t|\to\infty}|\varphi(t)| < 1 \tag{8.7.2}$$

is satisfied (the Cramér condition on the ch.f.). Property 11 of ch.f.s (see Sect. 8.7.1) implies that this condition is always met if the distribution of the sum S_m, for some $m \geq 1$, has a positive absolutely continuous component. The proof of Theorem 8.7.1 in its general form is more complicated and will be given at the end of this section, in Sect. 8.7.3.

In order to use the inversion formula (7.2.8), we employ the "smoothing method" and consider, along with S_n, the sums

$$Z_n = S_n + \eta_\delta, \tag{8.7.3}$$

where $\eta_\delta \in \mathbf{U}_{-\delta,0}$. Since the ch.f. $\varphi_{\eta_\delta}(t)$ of the random variable η_δ, being equal to

$$\varphi_{\eta_\delta}(t) = \frac{1 - e^{-it\delta}}{it\delta}, \tag{8.7.4}$$

possesses the property that the function $\varphi_{\eta_\delta}(t)/t$ is integrable at infinity, for the increments of the distribution function $G_n(x)$ of the random variable Z_n (its ch.f. divided by t is integrable, too) we can use formula (7.2.8):

$$G_n(x + \Delta) - G_n(x) = \mathbf{P}(Z_n \in \Delta[x)) = \frac{1}{2\pi}\int e^{-itx}\frac{1 - e^{-it\Delta}}{it}\,\varphi^n(t)\varphi_{\eta_\delta}(t)\,dt$$

$$= \frac{\Delta}{2\pi}\int e^{-itx}\varphi^n(t)\widehat{\varphi}(t)\,dt, \tag{8.7.5}$$

where $\widehat{\varphi}(t) = \varphi_{\eta_\delta}(t)\varphi_{\eta_\Delta}(t)$ (cf. (7.2.8)) is the ch.f. of the sum of independent random variables η_δ and η_Δ. We obtain that the difference $G_n(x + \Delta) - G_n(x)$, up to the factor Δ, is nothing else but the value of the density of the random variable $S_n + \eta_\delta + \eta_\Delta$ at the point x.

Split the integral on the right-hand side of (8.7.5) into the two subintegrals: one over the domain $|t| < \gamma$ for some $\gamma < 1$, and the other—over the complementary

domain. Put $x = v\sqrt{n}$ and consider first

$$I_1 := \int_{|t|<\gamma} e^{-itv\sqrt{n}}\varphi^n(t)\widehat{\varphi}(t)\,dt = \frac{1}{\sqrt{n}} \int_{|u|<\gamma\sqrt{n}} e^{-iuv}\varphi^n\left(\frac{u}{\sqrt{n}}\right)\widehat{\varphi}\left(\frac{u}{\sqrt{n}}\right)du.$$

Without loss of generality we can assume $\sigma = 1$, and by (8.2.1) obtain that

$$1 - \varphi(t) = \frac{t^2}{2} + o(t^2),$$

$$\ln\varphi(t) = \ln\left[1 - (1 - \varphi(t))\right] = -\frac{t^2}{2} + o(t^2) \quad \text{as } t \to 0. \qquad (8.7.6)$$

Hence

$$n\ln\varphi\left(\frac{u}{\sqrt{n}}\right) = -\frac{u^2}{2} + h_n(u), \qquad (8.7.7)$$

where $h_n(u) \to 0$ for any fixed u as $n \to \infty$. Moreover, for γ small enough, in the domain $|u| < \gamma\sqrt{n}$ we have

$$\left|h_n(u)\right| \le \frac{u^2}{6},$$

so the right-hand side of (8.7.7) does not exceed $-u^2/3$. Now we can rewrite I_1 in the form

$$I_1 = \frac{1}{\sqrt{n}} \int_{|u|<\gamma\sqrt{n}} \exp\left\{-iuv - \frac{u^2}{2} + h_n(u)\right\}\widehat{\varphi}\left(\frac{u}{\sqrt{n}}\right)du, \qquad (8.7.8)$$

where $|\widehat{\varphi}(u/\sqrt{n})| \le 1$ and $\widehat{\varphi}(u/\sqrt{n}) \to 1$ for any fixed u as $n \to \infty$. Therefore, by virtue of the dominated convergence theorem,

$$\sqrt{n}\,I_1 \to \int \exp\left\{-iuv - \frac{u^2}{2}\right\}du \qquad (8.7.9)$$

uniformly in v, since the integral on the right-hand side of (8.7.8) is uniformly continuous in v. But the integral on the right-hand side of (8.7.9) is simply (up to the factor $1/(2\pi)$) the result of applying the inversion formula to the ch.f. of the normal distribution, so that

$$\lim_{n\to\infty} \sqrt{n}\,I_1 = \sqrt{2\pi}\,e^{-v^2/2}. \qquad (8.7.10)$$

It remains to consider the integral

$$I_2 := \int_{|t|\ge\gamma} e^{-itv\sqrt{n}}\varphi^n(t)\widehat{\varphi}(t)\,dt.$$

By virtue of (8.7.2) and non-latticeness of the distribution of ξ,

$$q := \sup_{|t| \geq \gamma} |\varphi(t)| < 1 \qquad (8.7.11)$$

and therefore

$$|I_2| \leq q^n \int_{|t| \geq \gamma} |\widehat{\varphi}(t)| \, dt \leq q^n c(\Delta, \delta), \qquad \lim_{n \to \infty} \sqrt{n} I_2 = 0 \qquad (8.7.12)$$

uniformly in v, where $c(\Delta, \delta)$ depends on Δ and δ only. We have established that, for $x = v\sqrt{n}$, as $n \to \infty$, the relations

$$
\begin{aligned}
I_1 + I_2 &= \sqrt{\frac{2\pi}{n}} \, e^{-v^2/2} + o\left(\frac{1}{\sqrt{n}}\right), \\
\mathbf{P}\big(Z_n \in \Delta[x)\big) &= \frac{\Delta}{\sqrt{2\pi n}} e^{-x^2/(2n)} + o\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}
\qquad (8.7.13)
$$

hold uniformly in v (see (8.7.5)). This means that representation (8.7.13) holds uniformly for all x.

Further, by (8.7.3),

$$\big\{Z_n \in [x, x + \Delta - \delta)\big\} \subset \big\{S_n \in \Delta[x)\big\} \subset \big\{Z_n \in [x - \delta, x + \Delta)\big\} \qquad (8.7.14)$$

and, so, in particular,

$$\mathbf{P}\big(S_n \in \Delta[x)\big) \leq \frac{\Delta + \delta}{\sqrt{2\pi n}} e^{-(x-\delta)^2/(2n)} + o\left(\frac{1}{\sqrt{n}}\right) = \frac{\Delta + \delta}{\sqrt{2\pi n}} e^{-x^2/(2n)} + o\left(\frac{1}{\sqrt{n}}\right).$$

By (8.7.14) an analogous converse inequality also holds. Since δ is arbitrary, this is possible only if

$$\mathbf{P}\big(S_n \in \Delta[x)\big) = \frac{\Delta}{\sqrt{2\pi n}} e^{-x^2/(2n)} + o\left(\frac{1}{\sqrt{n}}\right). \qquad (8.7.15)$$

The theorem is proved.   □

8.7.2 Local Theorems

If the distribution of S_n has a density than we can obtain local theorems on the asymptotics of this density.

Theorem 8.7.2 *Let $\mathbf{E}\xi = 0$, $\mathbf{E}\xi^2 = \sigma^2 < \infty$ and suppose there exists an $m \geq 1$ such that at least one of the following three conditions is met*:

(a) *the distribution of S_m has a bounded density*;

(b) *the distribution of S_m has a density from L_2;*
(c) *the ch.f. $\varphi^m(t)$ of the sum S_m is integrable.*

Then, for $n \geq m$, the distribution of the sum S_n has density $f_{S_n}(x)$ for which the representation

$$f_{S_n}(x) = \frac{1}{\sqrt{2\pi n}\sigma} \exp\left\{-\frac{x^2}{2n\sigma^2}\right\} + o\left(\frac{1}{\sqrt{n}}\right) \tag{8.7.16}$$

holds uniformly in x as $n \to \infty$.

Conditions (a)–(c) *are equivalent to each other* (*possibly with different values of m*).

Proof We first establish the equivalence of (a)–(c). The fact that a bounded density belongs to L_2 was proved in Sect. 7.2.3. Conversely, if $f \in L_2$ then

$$\left|f^{(2)*}(t)\right| = \left|\int f(u)f(t-u)\,du\right|$$

$$\leq \left[\int f^2(u)\,du \times \int f^2(t-u)\,du\right]^{1/2} = \int f^2(u)\,du < \infty.$$

Hence the relationship $f_{S_m} \in L_2$ implies the boundedness of $f_{S_{2m}}$, and thus (a) and (b) are equivalent.

If φ^m is integrable then by Theorem 7.2.2 the density f_{S_m} exists and is bounded. Conversely, if f_{S_m} is bounded then $f_{S_m} \in L_2$, $\varphi_{S_m} \in L_2$ and $\varphi_{S_{2m}} \in L_1$ (see Sect. 8.7.2). This proves the equivalence of (a) and (c).

We will now prove (8.7.16). By the inversion formula (7.2.1),

$$f_{S_n}(x) = \frac{1}{2\pi}\int e^{-itx}\varphi^n(t)\,dt.$$

Here the integral on the right-hand side does not "qualitatively" differ from the integral on the right-hand side of (8.7.5), we only have to put $\widehat{\varphi}(t) \equiv 1$ in the part I_1 of the integral (8.7.5) (the integral over the set $|t| < \gamma$), and, in the part I_2 (over the set $|t| \geq \gamma$), to replace the integrable function $\widehat{\varphi}(t)$ with the integrable function $\varphi^m(t)$ and to replace the function $\varphi^n(t)$ with $\varphi^{n-m}(t)$. After these changes the whole argument in the proof of relation (8.7.13) remains valid, and therefore the same relation (up to the factor Δ) will hold for

$$f_{S_n}(x) = \frac{1}{\sqrt{2\pi n}\,\sigma} \exp\left\{-\frac{x^2}{2n\sigma^2}\right\} + o\left(\frac{1}{\sqrt{n}}\right).$$

The theorem is proved.   □

Theorem 8.7.2 implies that the density f_{ζ_n} of the random variable $\zeta_n = \frac{S_n}{\sigma\sqrt{n}}$ converges to the density ϕ of the standard normal law:

$$f_{\zeta_n}(v) \to \phi(v)$$

uniformly in v as $n \to \infty$.

For instance, the density of the uniform distribution over $[-1, 1]$ satisfies the conditions of this theorem, and hence the density of S_n at the point $x = v\sigma\sqrt{n}$ ($\sigma^2 = 1/3$) will behave as $\frac{1}{\sigma\sqrt{2\pi n}} e^{-v^2/(2\sigma^2)}$ (cf. the remark to Example 3.6.1).

In the *arithmetic case*, where the random variable ξ is integer-valued and the greatest common divisor of all possible values of ξ equals 1 (see Sect. 7.1), it is the asymptotics of the probabilities $\mathbf{P}(S_n = x)$ for integer x that become the subject of interest for local theorems. In this case we cannot assume without loss of generality that $\mathbf{E}\xi = 0$.

Theorem 8.7.3 (Gnedenko)   *Let* $\mathbf{E}\xi = a$, $\mathbf{E}\xi^2 = \sigma^2 < \infty$ *and* ξ *have an arithmetic distribution. Then, uniformly over all integers* x, *as* $n \to \infty$,

$$\mathbf{P}(S_n = x) = \frac{1}{\sqrt{2\pi n}\sigma} \exp\left\{\frac{(x - an)^2}{2n\sigma^2}\right\} + o\left(\frac{1}{\sqrt{n}}\right). \tag{8.7.17}$$

Proof When proving limit theorems for arithmetic ξ, it is more convenient to use the generating functions (see Sects. 7.1, 7.7)

$$p(z) \equiv p_\xi(z) := \mathbf{E}\, z^\xi, \quad |z| = 1,$$

so that $p(e^{it}) = \varphi(t)$, where φ is the ch.f. of ξ.

In this case the inversion formulas take the following form (see (7.2.10)): for integer x,

$$\mathbf{P}(\xi = x) = \frac{1}{2\pi i} \int_{|z|=1} z^{-x-1} p(z)\, dz,$$

$$\mathbf{P}(S_n = x) = \frac{1}{2\pi i} \int_{|z|=1} z^{-x-1} p^n(z)\, dz = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itx} \varphi^n(t)\, dt.$$

As in the proof of Theorem 8.7.1, here we split the integral on the right-hand side into two subintegrals: over the domain $|t| < \gamma$ and over the complementary set. The treatment of the first subintegral

$$I_1 := \int_{|t|<\gamma} e^{-itx} \varphi^n(t)\, dt = \int_{|t|<\gamma} e^{-ity} \left[e^{-ita}\varphi(t)\right]^n dt$$

for $y = x - an$ differs from the considerations for I_1 in Theorem 8.7.1 only in that it is simpler and yields (see (8.7.10))

$$I_1 = \frac{\sqrt{2\pi}}{\sigma\sqrt{n}} \exp\left\{-\frac{y^2}{2\pi\sigma^2}\right\} + o\left(\frac{1}{\sqrt{n}}\right).$$

Similarly, the treatment of the second subintegral differs from that of I_2 in Theorem 8.7.1 in that it becomes simpler, since the range of integration here is compact

and on that one has

$$\left|\varphi(t)\right| \leq q(\gamma) < 1. \tag{8.7.18}$$

Therefore, as in Theorem 8.7.1,

$$I_2 = o\left(\frac{1}{\sqrt{n}}\right), \qquad \mathbf{P}(S_n = x) = \frac{1}{\sqrt{2\pi n}\sigma} \exp\left\{-\frac{y^2}{2n\sigma^2}\right\} + o\left(\frac{1}{\sqrt{n}}\right).$$

The theorem is proved.   □

Evidently, for the values of y of order $\sqrt{n}$ Theorem 8.7.3 is a generalisation of the local limit theorem for the Bernoulli scheme (see Corollary 5.2.1).

8.7.3 The Proof of Theorem 8.7.1 in the General Case

To prove Theorem 8.7.1 in the general case we will use the same approach as in Sect. 7.1. We will again employ the smoothing method, but now, when specifying the random variable Z_n in (8.7.3), we will take $\theta\eta$ instead of η_δ, where $\theta = \mathrm{const}$, η is a random variable with the ch.f. from Example 7.2.1 (see the end of Sect. 7.2) equal to

$$\varphi_\eta(t) = \begin{cases} 1 - |t|, & |t| \leq 1; \\ 0, & |t| > 1, \end{cases}$$

so that for $Z_n = S_n + \theta\eta$, similarly to (8.7.5), we have

$$\mathbf{P}\big(Z_n \in \Delta[x)\big) = \frac{\Delta}{2\pi} \int_{|t| \leq \frac{1}{\theta}} e^{-itx} \varphi^n(t) \varphi_{\eta_\Delta}(t) \varphi_{\theta\eta}(t)\, dt, \tag{8.7.19}$$

where $\varphi_{\theta\eta}(t) = \max(0, 1 - \theta|t|)$. As in Sect. 8.7.1, split the integral on the right-hand side of (8.7.19) into two subintegrals: I_1 over the domain $|t| < \gamma$ and I_2 over the domain $\gamma \leq |t| \leq 1/\theta$. The asymptotic behaviour of these integrals is investigated in almost the same way as in Sect. 8.7.1, but is somewhat simpler, since the domain of integration in I_2 is compact, and so, by the non-latticeness of ξ, one has on it the upper bound

$$q := \sup_{\gamma \leq |t| \leq 1/\theta} \left|\varphi(t)\right| < 1. \tag{8.7.20}$$

Therefore, to bound I_2 we no longer need condition (8.7.2).

Thus we have established, as above, relation (8.7.13).

To derive from this fact the required relation (8.7.15) we will need the following.

Lemma 8.7.1 *Let $f(y)$ be a bounded uniformly continuous function, η an arbitrary proper random variable independent of S_n and $b(n) \to \infty$ as $n \to \infty$. If, for any*

fixed $\Delta > 0$ and $\theta > 0$, as $n \to \infty$, we have

$$\mathbf{P}\big(S_n + \theta\eta \in \Delta[x]\big) = \frac{\Delta}{b(n)}\left[f\left(\frac{x}{b(n)}\right) + o(1)\right], \tag{8.7.21}$$

then

$$\mathbf{P}\big(S_n \in \Delta[x]\big) = \frac{\Delta}{b(n)}\left[f\left(\frac{x}{b(n)}\right) + o(1)\right]. \tag{8.7.22}$$

In this assertion we can take S_n to be any sequence of random variables satisfying (8.7.21). In this section we will set $b(n)$ to be equal to $\sqrt{n}$, but later (see the proof of Theorem A7.2.1 in Appendix 7) we will need some other sequences as well.

Proof Put $\theta := \delta^2\Delta$, where $\delta > 0$ will be chosen later, $\Delta_\pm := (1\pm2\delta)\Delta$, $\Delta_\pm[x] := [x, x + \Delta_\pm)$ and $f_0 := \max f(y)$. We first obtain an upper bound for $\mathbf{P}(S_n \in \Delta[x])$. We have

$$\mathbf{P}\big(Z_n \in \Delta_+[x - \Delta\delta)\big) \geq \mathbf{P}\big(Z_n \in \Delta_+[x - \Delta\delta); |\eta| < 1/\delta\big).$$

On the event $|\eta| < 1/\delta$ one has $-\delta\Delta < \theta\eta < \delta\Delta$, and hence on this event

$$\big\{Z_n \in \Delta_+[x - \Delta\delta)\big\} \supset \big\{S_n \in \Delta[x]\big\}.$$

Thus, by independence of η and S_n,

$$\mathbf{P}\big(Z_n \in \Delta_+[x - \Delta\delta)\big) \geq \mathbf{P}\big(S_n \in \Delta[x]; |\eta| < 1/\delta\big) = \mathbf{P}\big(S_n \in \Delta[x]\big)\big(1 - h(\delta)\big),$$

where $h(\delta) := \mathbf{P}(|\eta| \geq 1/\delta) \to 0$ as $\delta \to 0$. By condition (8.7.21) and the uniform integrability of f we obtain

$$\mathbf{P}\big(S_n \in \Delta[x]\big) \leq \mathbf{P}\big(Z_n \in \Delta_+[x - \Delta\delta)\big)\big(1 - h(\delta)\big)^{-1}$$

$$\leq \left[\frac{\Delta}{b(n)}f\left(\frac{x}{b(n)}\right) + \frac{2\delta\Delta f_0}{b(n)} + o\left(\frac{1}{b(n)}\right)\right]\big(1 - h(\delta)\big)^{-1}.$$

$$\tag{8.7.23}$$

If, for a given $\varepsilon > 0$, we choose $\delta > 0$ such that

$$\big(1 - h(\delta)\big)^{-1} \leq 1 + \frac{\varepsilon\Delta}{3}, \qquad 2\delta f_0 \leq \frac{\varepsilon}{3},$$

then we derive from (8.7.23) that, for all n large enough and ε small enough,

$$\mathbf{P}\big(S_n \in \Delta[x]\big) \leq \frac{\Delta}{b(n)}\left(f\left(\frac{x}{b(n)}\right) + \varepsilon\right). \tag{8.7.24}$$

This implies, in particular, that for all x,

$$\mathbf{P}\big(S_n \in \Delta[x)\big) \leq \frac{\Delta}{b(n)}(f_0 + \varepsilon). \tag{8.7.25}$$

Now we will obtain a lower bound for $\mathbf{P}(S_n \in \Delta[x))$. For the event

$$A := \big\{Z_n \in \Delta_-[x + \Delta\delta)\big\}$$

we have

$$\mathbf{P}(A) = \mathbf{P}\big(A; |\eta| < 1/\delta\big) + \mathbf{P}\big(A; |\eta| \geq 1/\delta\big). \tag{8.7.26}$$

On the event $|\eta| < 1/\delta$ we have

$$\big\{Z_n \in \Delta_-[x + \Delta\delta)\big\} \subset \big\{S_n \in \Delta[x)\big\},$$

and hence

$$\mathbf{P}\big(A; |\eta| < 1/\delta\big) \leq \mathbf{P}\big(S_n \in \Delta[x)\big). \tag{8.7.27}$$

Further, by independence of η and S_n and inequality (8.7.25),

$$
\begin{aligned}
\mathbf{P}\big(A; |\eta| \geq 1/\delta\big) &= \mathbf{E}\big[\mathbf{P}(A \mid \eta); |\eta| \geq 1/\delta\big] \\
&= \mathbf{E}\big[\mathbf{P}\big(S_n \in \Delta_-[x + \theta\eta + \Delta\delta) \mid \eta\big); |\eta| \geq 1/\delta\big] \\
&\leq \frac{\Delta}{b(n)}(f_0 + \varepsilon)h(\delta).
\end{aligned}
$$

Therefore, combining (8.7.26), (8.7.27) and (8.7.21), we get

$$\mathbf{P}\big(S_n \in \Delta[x)\big) \geq \frac{\Delta}{b(n)} f\left(\frac{x}{b(n)}\right) - \frac{2\delta\Delta f_0}{b(n)} + o\left(\frac{1}{b(n)}\right) - \frac{\Delta}{b(n)}(f_0 + \varepsilon)h(\delta).$$

In addition, choosing δ such that

$$f_0 h(\delta) < \frac{\varepsilon}{3}, \qquad 2\delta f_0 < \frac{\varepsilon}{3},$$

we obtain that, for all n large enough and ε small enough,

$$\mathbf{P}\big(S_n \in \Delta[x)\big) \geq \frac{\Delta}{b(n)}\left(f\left(\frac{x}{b(n)}\right) - \varepsilon\right). \tag{8.7.28}$$

Since ε is arbitrarily small, inequalities (8.7.24) and (8.7.28) prove the required relation (8.7.22). The lemma is proved. $\square$

To prove the theorem it remains to apply Lemma 8.7.1 in the case (see (8.7.13)) where $f = \phi$ and $b(n) = \sqrt{n}$. Theorem 8.7.1 is proved. $\square$

8.7.4 Uniform Versions of Theorems 8.7.1–8.7.3 for Random Variables Depending on a Parameter

In the next chapter, we will need uniform versions of Theorems 8.7.1–8.7.3, where the summands ξ_k depend on a parameter λ. Denote such summands by $\xi_{(\lambda)k}$, the corresponding distributions by $\mathbf{F}_{(\lambda)}$, and put

$$S_{(\lambda)n} := \sum_{k=1}^{n} \xi_{(\lambda)k},$$

where $\xi_{(\lambda)k}$ are independent copies of $\xi_{(\lambda)} \Subset \mathbf{F}_{(\lambda)}$. If λ is only determined by the number of summands n then we will be dealing with the triangular array scheme considered in Sects. 8.3–8.6 (the summands there were denoted by $\xi_{k,n}$). In the general case we will take the segment $[0, \lambda_1]$ for some $\lambda_1 > 0$ as the parametric set, keeping in mind that $\lambda \in [0, \lambda_1]$ may depend on n (in the triangular array scheme one can put $\lambda = 1/n$).

We will be interested in what conditions must be imposed on a family of distributions $\mathbf{F}_{(\lambda)}$ for the assertions of Theorems 8.7.1–8.7.3 to hold uniformly in $\lambda \in [0, \lambda_1]$. We introduce the following notation:

$$a(\lambda) = \mathbf{E}\xi_{(\lambda)}, \qquad \sigma^2(\lambda) = \mathrm{Var}(\xi_{(\lambda)}), \qquad \varphi_{(\lambda)}(t) = \mathbf{E}e^{it\xi_{(\lambda)}}.$$

The next assertion is an analogue of Theorem 8.7.1.

Theorem 8.7.1A *Let the distributions* $\mathbf{F}_{(\lambda)}$ *satisfy the following properties*: $0 < \sigma_1 < \sigma(\lambda) < \sigma_2 < \infty$, *where* σ_1 *and* σ_2 *do not depend on* λ:

(a) *the relation*

$$\varphi_{(\lambda)}(t) - 1 - ia(\lambda)t + \frac{t^2 m_2(\lambda)}{2} = o(t^2), \quad m_2(\lambda) := \mathbf{E}\,\xi_{(\lambda)}^2, \qquad (8.7.29)$$

holds uniformly in $\lambda \in [0, \lambda_1]$ *as* $t \to 0$, *i.e. there exist a* $t_0 > 0$ *and a function* $\varepsilon(t) \to 0$ *as* $t \to 0$, *independent of* λ, *such that, for all* $|t| \le t_0$, *the absolute value of the left-hand side of* (8.7.29) *does not exceed* $\varepsilon(t)t^2$;

(b) *for any fixed* $0 < \theta_1 < \theta_2 < \infty$,

$$q_{(\lambda)} := \sup_{\theta_1 \le |t| \le \theta_2} \left|\varphi_{(\lambda)}(t)\right| \le q < 1, \qquad (8.7.30)$$

where q *does not depend on* λ.

Then, for each fixed $\Delta > 0$,

$$\mathbf{P}\big(S_{(\lambda)n} - na(\lambda) \in \Delta[x)\big) = \frac{\Delta}{\sigma(\lambda)\sqrt{n}} \phi\left(\frac{x}{\sigma(\lambda)\sqrt{n}}\right) + o\left(\frac{1}{\sqrt{n}}\right), \qquad (8.7.31)$$

where the remainder term $o(1/\sqrt{n})$ *is uniform in* x *and* $\lambda \in [0, \lambda_1]$.

Proof Going through the proof of Theorem 8.7.1 in its general form (see Sect. 7.3), we see that, to ensure the validity of all the proofs of the intermediate assertions in their uniform forms, it suffices to have uniformity in the following two places:

(a) the uniformity in λ of the estimate $o(t^2)$ as $t \to 0$ in relation (8.7.6) for the expansion of the ch.f. of the random variable $\xi = \frac{\xi_{(\lambda)} - a(\lambda)}{\sigma(\lambda)}$;

(b) the uniformity in relation (8.7.20) for the same ch.f.

We verify the uniformity in (8.7.6). For $\varphi(t) = \mathbf{E}\, e^{it\xi}$, we have by (8.7.29)

$$\ln \varphi(t) = -\frac{ita(\lambda)}{\sigma(\lambda)} + \ln \varphi_{(\lambda)}\left(\frac{t}{\sigma(\lambda)}\right)$$

$$= -\frac{t^2(m_2(\lambda) - a^2(\lambda))}{2\sigma^2(\lambda)} + o(t^2) = -\frac{t^2}{2} + o(t^2),$$

where the remainder term is uniform in λ.

The uniformity in relation (8.7.20) clearly follows from condition b), since $\sigma(\lambda)$ is uniformly separated from both 0 and ∞. The theorem is proved. $\qquad\square$

Remark 8.7.3 Conditions (a) and (b) of Theorem 8.7.1A are essential for (8.7.31) to hold. To see this, consider random variables ξ and η with fixed distributions, $\mathbf{E}\xi = \mathbf{E}\eta = 0$ and $\mathbf{E}\xi^2 = \mathbf{E}\eta^2 = 1$. Let $\lambda \in [0, 1]$ and the random variable $\xi_{(\lambda)}$ be defined by

$$\xi_{(\lambda)} := \begin{cases} \xi & \text{with probability } 1 - \lambda, \\ \frac{\eta}{\sqrt{\lambda}} & \text{with probability } \lambda, \end{cases} \qquad (8.7.32)$$

so that $\mathbf{E}\xi_{(\lambda)} = 0$ and $\mathrm{Var}(\xi_{(\lambda)}) = 2 - \lambda$ (in the case of the triangular array scheme one can put $\lambda = 1/n$). Then, under the obvious notational conventions, for $\lambda = t^2$, $t \to 0$, we have

$$\varphi_{(\lambda)}(t) = (1 - \lambda)\varphi_\xi(t) + \lambda\varphi_\eta\left(\frac{t}{\sqrt{\lambda}}\right) = 1 - \frac{3t^2}{2} + o(t^2) + t^2\varphi_\eta(1).$$

This implies that (8.7.29) does not hold and hence condition a) is not met for the values of λ in the vicinity of zero. At the same time, the uniform versions of relation (8.7.31) and the central limit theorem will fail to hold. Indeed, putting $\lambda = 1/n$, we obtain the triangular array scheme, in which the number ν_n of the summands of the form $\eta_i/\sqrt{\lambda}$ in the sum $S_{(\lambda)n} = \sum_{i=1}^n \xi_{(\lambda)i}$ converges in distribution to $\nu \Subset \mathbf{\Pi}_1$ and

$$\frac{1}{\sqrt{n(2 - \lambda)}} S_{(\lambda)n} \overset{d}{=} \frac{S_{n - \nu_n}}{\sqrt{2n - 1}} + \frac{H_{\nu_n}}{\sqrt{2 - 1/n}}, \qquad \text{where } H_k = \sum_{i=1}^k \eta_i.$$

The first term on the right-hand side weakly converges in distribution to $\zeta \Subset \mathbf{\Phi}_{0, 1/2}$, while the second term converges to $H_\nu/\sqrt{2}$. Clearly, the sum of these independent summands is, generally speaking, not distributed normally with parameters $(0, 1)$.

To see that condition (b) is also essential, consider an arithmetic random variable ξ with $\mathbf{E}\xi = 0$ and $\mathrm{Var}(\xi) = 1$, take η to be a random variable with the uniform distribution $\mathbf{U}_{-1,1}$, and put

$$\xi_{(\lambda)} := \begin{cases} \xi & \text{with probability } 1 - \lambda, \\ \eta & \text{with probability } \lambda. \end{cases}$$

Here the random variable $\xi_{(\lambda)}$ is non-lattice (its distribution has an absolutely continuous component), but

$$\varphi_{(\lambda)}(2\pi) = (1 - \lambda) + \lambda\varphi_\eta(2\pi), \qquad q_{(\lambda)} \geq 1 - 2\lambda.$$

Again putting $\lambda = 1/n$, we get the triangular array scheme for which condition (b) is not met. Relation (8.7.31) does not hold either, since, in the previous notation, the sum $S_{(\lambda)n}$ is integer-valued with probability $\mathbf{P}(\nu_n = 0) = e^{-1}$, so that its distribution will have atoms at integer points with probabilities comparable, by Theorem 8.7.3, with the right-hand side of (8.7.31). This clearly contradicts (8.7.31).

If we put $\lambda = 1/n^2$ then the sum $S_{(\lambda)n}$ will be integer-valued with probability $(1 - 1/n^2)^n \to 1$, and the failure of relation (8.7.31) becomes even more evident.

Uniform versions of the local Theorems 8.7.2 and 8.7.3 are established in a completely analogous way.

Theorem 8.7.2A *Let the distributions* $\mathbf{F}_{(\lambda)}$ *satisfy the conditions of Theorem* 8.7.1A *with* $\theta_2 = \infty$ *and the conditions of Theorem* 8.7.2, *in which conditions* (a)–(c) *are understood in the uniform sense (i.e.,* $\max_x f_{S_{(\lambda)m}}(x)$ *or the norm of* $f_{S_{(\lambda)m}}$ *in* L_2 *or* $\int |\varphi_{(\lambda)}^m(t)|\, dt$ *are bounded uniformly in* $\lambda \in [0, \lambda_1]$*).*

Then representation (8.7.16) *holds for* $f_{S_{(\lambda)n}}(x)$ *uniformly in* x *and* λ*, provided that on its right-hand side we replace* σ *by* $\sigma(\lambda)$*.*

Proof The conditions of Theorem 8.7.2A are such that they enable one to obtain the proof of the uniform version without any noticeable changes in the arguments proving Theorems 8.7.1A and 8.7.2. $\square$

The following assertion is established in the same way.

Theorem 8.7.3A *Let the arithmetic distributions* $\mathbf{F}_{(\lambda)}$ *satisfy the conditions of Theorem* 8.7.1A *for* $\theta_2 = \pi$*. Then representation* (8.7.17) *holds uniformly in* x *and* λ*, provided that* a *and* σ *on its right-hand side are replaced with* $a(\lambda)$ *and* $\sigma(\lambda)$*, respectively.*

Remark 8.7.3 applies to Theorems 8.7.2A and 8.7.3A as well.

8.8 Convergence to Other Limiting Laws

As we saw in previous sections, the normal law occupies a special place among all distributions—it is the limiting law for normed sums of arbitrary distributed random

variables. There arises the natural question of whether there exist any other limiting laws for sums of independent random variables.

It is clear from the proof of Theorem 8.2.1 for identically distributed random variables that the character of the limiting law is determined by the behaviour of the ch.f. of the summands in the vicinity of 0. If $\mathbf{E}\xi = 0$ and $\mathbf{E}\xi^2 = \sigma^2 = -\varphi''(0)$ exist, then

$$\varphi\left(\frac{1}{\sqrt{n}}\right) = 1 + \frac{\varphi''(0)t^2}{2n} + o\left(\frac{1}{n}\right),$$

and this determines the asymptotic behaviour of the ch.f. of $S_n/\sqrt{n}$, equal to $\varphi^n(t\sqrt{n})$, which leads to the normal limiting law. Therefore, if one is looking for different limiting laws for the sums $S_n = \xi_1 + \cdots + \xi_n$, it is necessary to renounce the condition that the variance is finite or, which is the same, that $\varphi''(0)$ exists. In this case, however, we will have to impose some conditions on the regular variation of the functions $F_+(x) = \mathbf{P}(\xi \geq x)$ and/or $F_-(x) = \mathbf{P}(\xi < -x)$ as $x \to \infty$, which we will call the *right* and the *left tail* of the distribution of ξ, respectively. We will need the following concepts.

Definition 8.8.1 A positive (Lebesgue) measurable function $L(t)$ is called a *slowly varying function* (s.v.f.) as $t \to \infty$, if, for any fixed $v > 0$,

$$\frac{L(vt)}{L(t)} \to 1 \quad \text{as } t \to \infty. \tag{8.8.1}$$

A function $V(t)$ is called a *regularly varying function* (r.v.f.) (of index $-\beta$) as $t \to \infty$ if it can be represented as

$$V(t) = t^{-\beta}L(t), \tag{8.8.2}$$

where $L(t)$ is an s.v.f. as $t \to \infty$.

One can easily see that, similarly to (8.8.1), the characteristic property of regularly varying functions is the convergence

$$\frac{V(vt)}{V(t)} \to v^{-\beta} \quad \text{as } t \to \infty \tag{8.8.3}$$

for any fixed $v > 0$. Thus an s.v.f. is an r.v.f. of index zero.

Among typical representatives the class of s.v.f.s are the logarithmic function and its powers $\ln^\gamma t$, $\gamma \in \mathbb{R}$, linear combinations thereof, multiple logarithms, functions with the property that $L(t) \to L = \text{const} \neq 0$ as $t \to \infty$ etc. As an example of *a bounded oscillating* s.v.f. we mention

$$L_0(t) = 2 + \sin(\ln \ln t), \quad t > 1.$$

The main properties of r.v.f.s are given in Appendix 6.

As has already been noted, for $S_n/b(n)$ to converge to a "nondegenerate" limiting law under a suitable normalisation $b(n)$, we will have to impose conditions on the regular variation of the distribution tails of ξ. More precisely, we will need a regular variation of the "two-sided tail"

$$F_0(t) = F_-(t) + F_+(t) = \mathbf{P}\big(\xi \notin [-t, t)\big).$$

We will assume that the following condition is satisfied for some $\beta \in (0, 2]$, $\rho \in [-1, 1]$:

$[\mathbf{R}_{\beta,\rho}]$ The two-sided tail $F_0(x) = F_-(x) + F_+(x)$ is an r.v.f. as $x \to \infty$, i.e. it can be represented as

$$F_0(x) = t^{-\beta} L_{F_0}(x), \quad \beta \in (0, 2], \tag{8.8.4}$$

where $L_{F_0}(x)$ is an s.v.f., and the following limit exists

$$\rho_+ := \lim_{x \to \infty} \frac{F_+(x)}{F_0(x)} \in [0, 1], \qquad \rho := 2\rho_+ - 1. \tag{8.8.5}$$

If $\rho_+ > 0$, then clearly the right tail $F_+(x)$ is an r.v.f. like $F_0(x)$, i.e. it can be represented as

$$F_+(x) = V(x) := x^{-\beta} L(x), \quad \beta \in (0, 2], \ L(x) \sim \rho_+ L_{F_0}(x).$$

(Here, and likewise in Appendix 6, we use the symbol V to denote an r.v.f.) If $\rho_+ = 0$, then the right tail $F_+(x) = o(F_0(x))$ is not assumed to be regularly varying.

Relation (8.8.5) implies that the following limit also exists

$$\rho_- := \lim_{x \to \infty} \frac{F_-(x)}{F_0(x)} = 1 - \rho_+.$$

If $\rho_- > 0$, then, similarly to the case of the right tail, the left tail $F_-(x)$ can be represented as

$$F_-(x) = W(x) := x^{-\beta} L_W(x), \quad \beta \in (0, 2], \ L_W(x) \sim \rho_- L_{F_0}(x).$$

If $\rho_- = 0$, then the left tail $F_-(x) = o(F_0(x))$ is not assumed to be regularly varying.

The parameters $\rho_\pm$ are related to the parameter ρ in the notation $[\mathbf{R}_{\beta,\rho}]$ through the equalities

$$\rho = \rho_+ - \rho_- = 2\rho_+ - 1 \in [-1, 1].$$

Clearly, in the case $\beta < 2$ we have $\mathbf{E}\xi^2 = \infty$, so that the representation

$$\varphi(t) = 1 - \frac{t^2\sigma^2}{2} + o(t^2) \quad \text{as } t \to 0$$

no longer holds, and the central limit theorem is not applicable. If $\mathbf{E}\xi$ exists and is finite then everywhere in what follows it will be assumed without loss of generality

that

$$\mathbf{E}\xi = 0.$$

Since $F_0(x)$ is non-increasing, there always exists the "generalised" inverse function $F_0^{(-1)}(u)$ understood as

$$F_0^{(-1)}(u) := \inf\{x : F_0(x) < u\}.$$

If the function F_0 is strictly monotone and continuous then $b = F_0^{(-1)}(u)$ is the unique solution to the equation

$$F_0(b) = u, \quad u \in (0, 1).$$

Set

$$\zeta_n := \frac{S_n}{b(n)},$$

wherein the case $\beta > 2$ we define the normalising factor $b(n)$ by

$$b(n) := F_0^{(-1)}(1/n). \tag{8.8.6}$$

For $\beta = 2$ put

$$b(n) := Y^{(-1)}(1/n), \tag{8.8.7}$$

where

$$Y(x) := 2x^{-2} \int_0^x y F_0(y)\, dy = 2x^{-2}\left[\int_0^x y F_+(y)\, dy + \int_0^x y F_-(y)\, dy\right]$$
$$= x^{-2}\mathbf{E}\big(\xi^2; -x \le \xi < x\big) = x^{-2} L_Y(x), \tag{8.8.8}$$

L_Y is an s.v.f. (see Theorem A6.2.1(iv) in Appendix 6). It follows from Theorem A6.2.1(v) in Appendix 6 that, under condition (8.8.4), we have

$$b(n) = n^{1/\beta} L_b(n), \quad \beta \le 2,$$

where L_b is an s.v.f.

We introduce the functions

$$V_I(x) = \int_0^x V(y)\, dy, \qquad V^I(x) = \int_x^\infty V(y)\, dy.$$

8.8.1 The Integral Theorem

Theorem 8.8.1 *Let condition* $[\mathbf{R}_{\beta,\rho}]$ *be satisfied. Then the following assertions hold true.*

(i) *For $\beta \in (0, 2)$, $\beta \neq 1$ and the normalising factor* (8.8.6), *as $n \to \infty$,*

$$\zeta_n \Rightarrow \zeta^{(\beta,\rho)}.\tag{8.8.9}$$

The distribution $\mathbf{F}_{\beta,\rho}$ of the random variable $\zeta^{(\beta,\rho)}$ depends on parameters β and ρ only and has a ch.f. $\varphi^{(\beta,\rho)}(t)$, given by

$$\varphi^{(\beta,\rho)}(t) := \mathbf{E}e^{it\zeta^{(\beta,\rho)}} = \exp\{|t|^\beta B(\beta, \rho, \vartheta)\},\tag{8.8.10}$$

where $\vartheta = \operatorname{sign} t$,

$$B(\beta, \rho, \vartheta) = \Gamma(1 - \beta)\left[i\rho\vartheta \sin\frac{\beta\pi}{2} - \cos\frac{\beta\pi}{2}\right]\tag{8.8.11}$$

and, for $\beta \in (1, 2)$, we put $\Gamma(1 - \beta) = \Gamma(2 - \beta)/(1 - \beta)$.

(ii) *When $\beta = 1$, for the sequence ζ_n with the normalising factor* (8.8.6) *to converge to a limiting law, the former, generally speaking, needs to be centred. More precisely, as $n \to \infty$, the following convergence takes place:*

$$\zeta_n - A_n \Rightarrow \zeta^{(1,\rho)},\tag{8.8.12}$$

where

$$A_n = \frac{n}{b(n)}\left[V_I\big(b(n)\big) - W_I\big(b(n)\big)\right] - \rho C,\tag{8.8.13}$$

$C \approx 0.5772$ is the Euler constant, and

$$\varphi^{(1,\rho)}(t) = \mathbf{E}e^{it\zeta^{(1,\rho)}} = \exp\left\{-\frac{\pi|t|}{2} - i\rho t \ln|t|\right\}.\tag{8.8.14}$$

If $n[V_I(b(n)) - W_I(b(n))] = o(b(n))$, then $\rho = 0$ and we can put $A_n = 0$. If $\mathbf{E}\xi$ exists and equals zero then

$$A_n = \frac{n}{b(n)}\left[W^I\big(b(n)\big) - V^I\big(b(n)\big)\right] - \rho C.$$

If $\mathbf{E}\xi = 0$ and $\rho \neq 0$ then $\rho A_n \to -\infty$ as $n \to \infty$.

(iii) *For $\beta = 2$ and the normalising factor* (8.8.7), *as $n \to \infty$,*

$$\zeta_n \Rightarrow \zeta^{(2,\rho)}, \qquad \varphi^{(2,\rho)}(t) := \mathbf{E}e^{it\zeta^{(2,\rho)}} = e^{-t^2/2},$$

so that $\zeta^{(2,\rho)}$ has the standard normal distribution that is independent of ρ.

The Proof of Theorem 8.8.1 is based on the same considerations as the proof of Theorem 8.2.1, i.e. on using the asymptotic behaviour of the ch.f. $\varphi(t)$ in the vicinity of zero. But here it will be somewhat more difficult from the technical viewpoint. This is why the proof of Theorem 8.8.1 appears in Appendix 7. □

Remark 8.8.1 The last assertion of the theorem (for $\beta = 2$) shows that the limiting distribution may be normal even in the case of infinite variance of ξ.

Besides with the normal distribution, we also note "extreme" limit distributions, corresponding to the $\rho = \pm 1$ where the ch.f. $\varphi^{(\beta,\rho)}$ (or the respective Laplace transform) takes a very simple form. Let, for example, $\rho = -1$. Since $e^{i\pi\vartheta/2} = \vartheta i$, then, for $\beta \neq 1, 2$,

$$B(\beta, -1, \vartheta) = -\Gamma(1 - \beta)\left[i \sin \frac{\beta\pi\vartheta}{2} + \cos \frac{\beta\pi\vartheta}{2}\right]$$

$$= -\Gamma(1 - \beta)e^{i\beta\pi\vartheta/2} = -\Gamma(1 - \beta)(i\vartheta)^{\beta},$$

$$\varphi^{(\beta,-1)}(t) = \exp\{-\Gamma(1 - \beta)(it)^{\beta}\},$$

$$\mathbf{E}\,e^{\lambda\zeta^{(\beta,-1)}} = \exp\{-\Gamma(1 - \beta)\lambda^{\beta}\}, \quad \mathrm{Re}\,\lambda \geq 0.$$

Similarly, for $\beta = 1$, by (8.8.14) and the equalities $-\frac{\pi\vartheta}{2} = i\,\frac{i\pi\vartheta}{2} = i\ln i\vartheta$ we have

$$\ln\varphi^{(1,-1)}(t) = -\frac{\pi\vartheta t}{2} + it\ln|t| = it\ln i\vartheta + it\ln|t| = it\ln it,$$

$$\mathbf{E}\,e^{\lambda\zeta^{(1,-1)}} = \exp\{\lambda\ln\lambda\}, \quad \mathrm{Re}\,\lambda \geq 0.$$

A similar formula is valid for $\rho = 1$.

Remark 8.8.2 If $\beta < 2$, then by virtue of the properties of s.v.f.s (see Theorem A6.2.1(iv) in Appendix 6), as $x \to \infty$,

$$\int_0^x y F_0(y)\,dy = \int_0^x y^{1-\beta} L_{F_0}(y)\,dy \sim \frac{1}{2 - \beta} x^{2-\beta} L_{F_0}(x) = \frac{1}{2 - \beta} x^2 F_0(x).$$

Therefore, for $\beta < 2$, we have $Y(x) \sim 2(2 - \beta)^{-1} F_0(x)$,

$$Y^{(-1)}(1/n) \sim F_0^{(-1)}\left(\frac{2 - \beta}{2n}\right) \sim \left(\frac{2}{2 - \beta}\right)^{1/\beta} F_0^{(-1)}(1/n)$$

(cf. (8.8.6)). On the other hand, for $\beta = 2$ and $\sigma^2 := \mathbf{E}\xi^2 < \infty$ one has

$$Y(x) \sim x^{-2}\sigma^2, \qquad b(n) = Y^{(-1)}(1/n) \sim \sqrt{\sigma}n.$$

Thus normalisation (8.8.7) is "transitional" from normalisation (8.8.6) (up to the constant factor $(2/(2 - \beta))^{1/\beta}$) to the standard normalisation $\sigma\sqrt{n}$ in the central limit theorem in the case where $\mathbf{E}\xi^2 < \infty$. This also means that normalisation (8.8.7) is "universal" and can be used for all $\beta \leq 2$ (as it is done in many textbooks on probability theory). However, as we will see below, in the case $\beta < 2$ normalisation (8.8.6) is easier and simpler to deal with, and therefore we will use that scaling.

Recall that $\mathbf{F}_{\beta,\rho}$ denotes the distribution of the random variable $\zeta^{(\beta,\rho)}$. The parameter β takes values in the interval $(0, 2]$, the parameter $\rho = \rho_+ - \rho_-$ can assume any value from $[-1, 1]$. The role of the parameters β and ρ will be clarified below.

Theorem 8.8.1 implies that each of the laws $\mathbf{F}_{\beta,\rho}$, $0 < \beta \le 2$ and $-1 \le \rho \le 1$ is *limiting* for the distributions of suitably normalised sums of independent identically distributed random variables. It follows from the law of large numbers that the degenerate distribution $\mathbf{I}_a$ concentrated at the point a is also a limiting one. Denote the set of all such distributions by $\mathfrak{S}_0$. Furthermore, it is not hard to see that if $\mathbf{F}$ is a distribution from the class $\mathfrak{S}_0$ then the law that differs from $\mathbf{F}$ by scaling and shifting, i.e. the distribution $\mathbf{F}_{\{a,b\}}$ defined, for some fixed $b > 0$ and a, by the relation

$$\mathbf{F}_{\{a,b\}}(B) := \mathbf{F}\left(\frac{B-a}{b}\right), \quad \text{where } \frac{B-a}{b} = \{u \in \mathbb{R} : ub + a \in B\},$$

is also limiting for the distributions of sums of random variables $(S_n - a_n)/b_n$ as $n \to \infty$ for appropriate $\{a_n\}$ and $\{b_n\}$.

It turns out that the *class of distributions $\mathfrak{S}$ obtained by the above extension from $\mathfrak{S}_0$ exhausts all the limiting laws for sums of identically distributed independent random variables*.

Another characterisation of the class of limiting laws $\mathfrak{S}$ is also possible.

Definition 8.8.2 We call a distribution $\mathbf{F}$ *stable* if, for any a_1, a_2, $b_1 > 0$, $b_2 > 0$, there exist a and $b > 0$ such that

$$\mathbf{F}_{\{a_1,b_1\}} * \mathbf{F}_{\{a_2,b_2\}} = \mathbf{F}_{\{a,b\}}.$$

This definition means that the convolution of a stable distribution $\mathbf{F}$ with itself again yields the same distribution $\mathbf{F}$, up to a scaling and shift (or, which is the same, for independent random variables $\xi_i \Subset \mathbf{F}$ we have $(\xi_1 + \xi_2 - a)/b \Subset \mathbf{F}$ for appropriate a and b).

In terms of the ch.f. φ, the stability property has the following form: for any $b_1 > 0$ and $b_2 > 0$, there exist a and $b > 0$ such that

$$\varphi(tb_1)\varphi(tb_2) = e^{ita}\varphi(tb), \quad t \in \mathbb{R}. \tag{8.8.15}$$

Denote the class of all stable laws by $\mathfrak{S}^S$. The remarkable fact is that *the class of all limiting laws $\mathfrak{S}$ (for $(S_n - a_n)/b_n$ for some a_n and b_n) and the class of all stable laws $\mathfrak{S}^S$ coincide*.

If, under a suitable normalisation, as $n \to \infty$,

$$\zeta_n \Rightarrow \zeta^{(\beta,\rho)},$$

then one says that *the distribution $\mathbf{F}$ of the summands ξ belongs to the domain of attraction of the stable law $\mathbf{F}_{\beta,\rho}$*.

Theorem 8.8.1 means that, if $\mathbf{F}$ satisfies condition $[\mathbf{R}_{\beta,\rho}]$, then $\mathbf{F}$ belongs to the domain of attraction of the stable law $\mathbf{F}_{\beta,\rho}$.

One can prove the converse assertion (see e.g. Chap. XVII, § 5 in [30]): if $\mathbf{F}$ belongs to the domain of attraction of a stable law $\mathbf{F}_{\beta,\rho}$ for $\beta < 2$, then $[\mathbf{R}_{\beta,\rho}]$ is satisfied.

As for the role of the parameters β and ρ, note the following. The parameter β *characterises the rate of convergence to zero as $x \to \infty$* for the functions

$$F_{\beta,\rho,-}(x) := \mathbf{F}_{\beta,\rho}\big((-\infty, -x)\big) \quad \text{and} \quad F_{\beta,\rho,+}(x) := \mathbf{F}_{\beta,\rho}\big([x, \infty)\big).$$

One can prove that, for $\rho_+ > 0$, as $t \to \infty$,

$$F_{\beta,\rho,+}(t) \sim \rho_+ t^{-\beta}, \tag{8.8.16}$$

and, for $\rho_- > 0$, as $t \to \infty$,

$$F_{\beta,\rho,-}(t) \sim \rho_- t^{-\beta}. \tag{8.8.17}$$

Note that, for $\xi \in \mathbf{F}_{\beta,\rho}$, the asymptotic relations in Theorem 8.8.1 turn into precise equalities provided that we replace in them $b(n)$ with $b_n := n^{1/\beta}$. In particular,

$$\mathbf{P}\left(\frac{S_n}{b_n} \geq t\right) = F_{\beta,\rho,+}(t). \tag{8.8.18}$$

This follows from the fact that $[\varphi^{(\beta,\rho)}(t/b_n)]^n$ coincides with $\varphi^{(\beta,\rho)}(t)$ (see (8.8.10)) and hence the distribution of the normalised sum S_n/b_n coincides with the distribution of the random variable ξ.

The parameter ρ taking values in $[-1, 1]$ is the *measure of asymmetry* of the distribution $\mathbf{F}_{\beta,\rho}$. If, for instance, $\rho = 1$ ($\rho_- = 0$), then, for $\beta < 1$, the distribution $\mathbf{F}_{\beta,1}$ is concentrated entirely on the positive half-line. This is evident from the fact that in this case $\mathbf{F}_{\beta,1}$ can be considered as the limiting distribution for the normalised sums of independent identically distributed random variables $\xi_k \geq 0$ (with $F_-(0) = 0$). Since all the prelimit distributions are concentrated on the positive half-line, so is the limiting distribution.

Similarly, for $\rho = -1$ and $\beta < 1$, the distribution $\mathbf{F}_{\beta,-1}$ is entirely concentrated on the negative half-line. For $\rho = 0$ ($\rho_+ = \rho_- = 1/2$) the ch.f. of the distribution $\mathbf{F}_{\beta,0}$ will be real, and the distribution $\mathbf{F}_{\beta,0}$ itself is symmetric.

As we saw above, the ch.f.s $\varphi^{(\beta,\rho)}(t)$ of stable laws $\mathbf{F}_{\beta,\rho}$ admit closed-form representations. They are clearly integrable over $\mathbb{R}$, and the same is true for the functions $t^k \varphi^{(\beta,\rho)}(t)$ for any $k \geq 1$. Therefore all the stable distributions have densities that are differentiable arbitrarily many times (see e.g. the inversion formula (7.2.1)). As for explicit forms of these densities, they are only known for a few laws. Among them are:

1. The normal law $\mathbf{F}_{2,\rho}$ (which does not depend on ρ).

2. The Cauchy distribution $\mathbf{F}_{1,0}$ with density $2/(\pi^2 + 4x^2)$, $-\infty < x < \infty$. Scaling the x-axis with a factor of $\pi/2$ transforms this density into the form $1/\pi(1+x^2)$ corresponding to $\mathbf{K}_{0,1}$.

3. The Lévy distribution. This law can be obtained from the explicit form for the distribution of the maximum of the Wiener process. This will be the distribution $\mathbf{F}_{1/2,1}$ with parameters $1/2$, 1 and density (up to scaling; cf. (8.8.16))

$$f^{(1/2,1)}(x) = \frac{1}{\sqrt{2\pi}\,x^{3/2}}\, e^{-1/(2x)}, \quad x > 0$$

(this density has a first hitting time of level 1 by the standard Wiener process, see Theorem 19.2.2).

8.8.2 The Integro-Local and Local Theorems

Under the conditions of this section we can also obtain integro-local and local theorems in the same way as in Sect. 8.7 in the case of convergence to the normal law. As in Sect. 8.7, integro-local theorems deal here with the asymptotics of

$$\mathbf{P}\big(S_n \in \Delta[x)\big), \qquad \Delta[x) = [x, x + \Delta)$$

as $n \to \infty$ for a fixed $\Delta > 0$.

As we can see from Theorem 8.8.1, the ch.f. $\varphi^{(\beta,\rho)}(t)$ of the stable law $\mathbf{F}_{\beta,\rho}$ is integrable, and hence, by the inversion formula, there exists a uniformly continuous density $f^{(\beta,\rho)}$ of the distribution $\mathbf{F}_{\beta,\rho}$. (As has already been noted, it is not difficult to show that $f^{(\beta,\rho)}$ is differentiable arbitrarily many times, see Sect. 7.2.)

Theorem 8.8.2 (The Stone integro-local theorem) *Let ξ be a non-lattice random variable and the conditions of Theorem 8.8.1 be met. Then, for any fixed $\Delta > 0$, as $n \to \infty$,*

$$\mathbf{P}\big(S_n \in \Delta[x)\big) = \frac{\Delta}{b(n)}\, f^{(\beta,\rho)}\left(\frac{x}{b(n)}\right) + o\left(\frac{1}{b(n)}\right), \tag{8.8.19}$$

where the remainder term $o(\frac{1}{b(n)})$ is uniform over x.

If $\beta = 1$ and $\mathbf{E}|\xi|$ does not exist then, on the right-hand side of (8.8.20), we must replace $f^{(\beta,\rho)}(\frac{x}{b(n)})$ with $f^{(\beta,\rho)}(\frac{x}{b(n)} - A_n)$, where A_n is defined in (8.8.13).

All the remarks to the integro-local Theorem 8.7.1 hold true here as well, with evident changes.

If the distribution of S_n has a density then we can find the asymptotics of that density.

Theorem 8.8.3 *Let there exist an $m \geq 1$ such that at least one of conditions (a)–(c) of Theorem 8.7.2 is satisfied. Moreover, let the conditions of Theorem 8.8.1 be met. Then for the density $f_{S_n}(x)$ of the distribution of S_n one has the representation*

$$f_{S_n}(x) = \frac{1}{b(n)} f^{(\beta,\rho)}\left(\frac{x}{b(n)}\right) + o\left(\frac{1}{b(n)}\right) \tag{8.8.20}$$

which holds uniformly in x as $n \to \infty$.

If $\beta = 1$ and $\mathbf{E}|\xi|$ does not exist then, on the right-hand side of (8.8.20), we must replace $f^{(\beta,\rho)}(\frac{x}{b(n)})$ with $f^{(\beta,\rho)}(\frac{x}{b(n)} - A_n)$, where A_n is defined in (8.8.13).

The assertion of Theorem 8.8.3 can be rewritten for $\zeta_n = \frac{S_n}{b(n)} - A_n$ as

$$f_{\zeta_n}(v) \to f^{(\beta,\rho)}(v)$$

for any v as $n \to \infty$.

For integer-valued ξ_k the following theorem holds true.

Theorem 8.8.4 *Let the distribution of ξ be arithmetic and the conditions of Theorem 8.8.1 be met. Then, uniformly for all integers x, as $n \to \infty$,*

$$\mathbf{P}(S_n = x) = \frac{1}{b(n)} f^{(\beta,\rho)}\left(\frac{x - an}{b(n)}\right) + o\left(\frac{1}{\sqrt{n}}\right), \tag{8.8.21}$$

where $a = \mathbf{E}\xi$ if $\mathbf{E}|\xi|$ exists and $a = 0$ if $\mathbf{E}|\xi|$ does not exist, $\beta \neq 1$. If $\beta = 1$ and $\mathbf{E}|\xi|$ does not exist then, on the right-hand side of (8.8.21), we must replace $f^{(\beta,\rho)}(\frac{x-an}{b(n)})$ with $f^{(\beta,\rho)}(\frac{x}{b(n)} - A_n)$.

The proofs of Theorems 8.8.2–8.8.4 mostly repeat those of Theorems 8.7.1–8.7.3 and can be found in Appendix 7.

8.8.3 An Example

In conclusion we will consider an example.

In Sect. 12.8 we will see that in the fair game considered in Example 4.2.3 the ruin time $\eta(z)$ of a gambler with an initial capital of z units satisfies the relation $\mathbf{P}(\eta(z) \geq n) \sim z\sqrt{2/\pi n}$ as $n \to \infty$. In particular, for $z = 1$,

$$\mathbf{P}\big(\eta(1) \geq n\big) \sim \sqrt{2/\pi n}. \tag{8.8.22}$$

It is not hard to see (for more detail, see also Chap. 12) that $\eta(z)$ has the same distribution as $\eta_1 + \eta_2 + \cdots + \eta_z$, where η_j are independent and distributed as $\eta(1)$.

Thus for studying the distribution of $\eta(z)$ when z is large, by virtue of (8.8.22), one can make use of Theorem 8.8.4 (with $\beta = 1/2$, $b(n) = 2n^2/\pi$), by which

$$\lim_{z \to \infty} \mathbf{P}\left(\frac{2\pi \eta(x)}{z^2} < x\right) = F_{1/2,1}(x) \tag{8.8.23}$$

is the Lévy stable law with parameters $\beta = 1/2$ and $\rho = 1$. Moreover, for integer x and $z \to \infty$,

$$\mathbf{P}\big(\eta(z) = x\big) = \frac{\pi}{2z^2} f^{(1/2,1)}\left(\frac{x\pi}{2z^2}\right) + o\left(\frac{1}{z^2}\right).$$

These assertions enable one to obtain the limiting distribution for the number of crossings of an arbitrary strip $[u, v]$ by the trajectory $S_1, \ldots, S_n$ in the case where

$$\mathbf{P}(\xi_k = -1) = \mathbf{P}(\xi_k = -1) = 1/2.$$

Indeed, let for simplicity $u = 0$. By the first *positive crossing* of the strip $[0, v]$ we will mean the Markov time

$$\eta_+ := \min\{k : S_k = v\}.$$

The first *negative crossing* of the strip is then defined as the time $\eta_+ + \eta_-$, where

$$\eta_- := \min\{k : S_{\eta_+ + k} = 0\}.$$

The time $\eta_1 = \eta_+ + \eta_-$ will also be the time of the "double crossing" of $[0, v]$. The variables $\eta_\pm$ are distributed as $\eta(v)$ and are independent, so that η_1 has the same distribution as $\eta(2v)$. The variable $H_k = \eta_1(2v) + \cdots + \eta_k(2v)$, where $\eta_i(2v)$ have the same distribution as $\eta(2v)$ and are independent, is the time of the k-th double crossing. Therefore

$$\nu(n) := \max\{k : H_k \le n\} = \min\{k : H_k > n\} - 1$$

is the number of double crossings of the strip $[0, v]$ by time n. Now we can prove the following assertion:

$$\lim_{n \to \infty} \mathbf{P}\left(\frac{\nu(n)}{\sqrt{n}} \ge x\right) = F_{1/2,1}\left(\frac{\pi}{2v^2 x^2}\right). \tag{8.8.24}$$

To prove it, we will make use of the following relation (which will play, in its more general form, an important role in Chap. 10):

$$\big\{\nu(n) \ge k\big\} = \{H_k \le n\},$$

where H_k is distributed as $\eta(2vk)$. If $n/k^2 \to s^2$ as $n \to \infty$, then by virtue of (8.8.23)

$$\mathbf{P}(H_k \le n) = \mathbf{P}\left(\frac{2\pi H_k}{(2vk)^2} \le \frac{2\pi n}{(2vk)^2}\right) \to F_{1/2,1}\left(\frac{\pi s^2}{2v^2}\right),$$

and therefore

$$\mathbf{P}\left(\frac{\nu(n)}{\sqrt{n}} \geq x\right) = \mathbf{P}\left(\nu(n) \geq x\sqrt{n}\right) = \mathbf{P}(H_{\lfloor x\sqrt{n}\rfloor} \leq n) \to F_{1/2,1}\left(\frac{\pi}{2\nu^2 x^2}\right).$$

(Here for $k = \lfloor x\sqrt{n}\rfloor$ one has $n/k^2 \to s^2 = 1/x^2$.) Relation (8.8.24) is proved. $\square$

Assertion (8.8.24) will clearly remain true for the number of crossings of the strip $[u, v]$, $u \neq 0$; one just has to replace v with $v - u$ on the right-hand side of (8.8.24). It is also clear that (8.8.24) enables one to find the limiting distribution of the number of "simple" (not double) crossings of $[u, v]$ since the latter is equal to $2\nu(n)$ or $2\nu(n)+1$.