A

Some Useful Mathematics

In this appendix we collect a number of mathematical facts which may or may not be (or have been) familiar to the reader.

1 Taylor Expansion

A common method to estimate functions is to use Taylor expansion. Here are a few such estimates.

Lemma 1.1. *We have,*

$$\mathrm{e}^x \le 1 + x + x^2 \quad \textit{for } |x| \le 1, \tag{A.1}$$

$$-\frac{1}{1-\delta}x < \log(1-x) < -x \quad \textit{for } 0 < x < \delta < 1, \tag{A.2}$$

$$|\mathrm{e}^z - 1| \le |z|\mathrm{e}^{|z|} \quad \textit{for } z \in \mathbb{C}, \tag{A.3}$$

$$|\mathrm{e}^z - 1 - z| \le |z|^2 \quad \textit{for } z \in \mathbb{C}, \quad |z| \le 1/2, \tag{A.4}$$

$$|\log(1-z) + z| \le |z|^2 \quad \textit{for } z \in \mathbb{C}, \quad |z| \le 1/2. \tag{A.5}$$

Proof. Let $0 \le x \le 1$. By Taylor expansion, noticing that we have an alternating series that converges,

$$\mathrm{e}^{-x} \le 1 - x + \frac{x^2}{2}.$$

For the other half of (A.1),

$$\begin{aligned}
\mathrm{e}^x &= 1 + x + x^2 \sum_{k=2}^{\infty} \frac{x^{k-2}}{k!} \le 1 + x + x^2 \sum_{k=2}^{\infty} \frac{1}{k!} \\
&= 1 + x + x^2(\mathrm{e} - 2) \le 1 + x + x^2.
\end{aligned}$$

To prove (A.2), we use Taylor expansion to find that, for $0 < x < \delta < 1$,

$$\log(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k} \geq -x - x\sum_{k=2}^{\infty} \frac{\delta^{k-1}}{k}$$
$$\geq -x - x\sum_{k=1}^{\infty} \delta^k = -x - x\frac{\delta}{1-\delta} = -x\frac{1}{1-\delta},$$

which proves the lower inequality. The upper one is trivial.

Next, let $z \in \mathbb{C}$. Then

$$|\mathrm{e}^z - 1| \leq \sum_{k=1}^{\infty} \frac{|z|^k}{k!} = |z|\sum_{k=0}^{\infty} \frac{|z|^k}{(k+1)!} \leq |z|\sum_{k=0}^{\infty} \frac{|z|^k}{k!} = |z|\mathrm{e}^{|z|}.$$

If, in addition, $|z| \leq 1/2$, then

$$|\mathrm{e}^z - 1 - z| \leq \sum_{k=2}^{\infty} \frac{|z|^k}{k!} \leq \frac{|z|^2}{2}\sum_{k=2}^{\infty} |z|^{k-2} = \frac{|z|^2}{2} \cdot \frac{1}{1-|z|} \leq |z|^2,$$
$$|\log(1-z) + z| \leq \sum_{k=2}^{\infty} \frac{|z|^k}{k} \leq \frac{|z|^2}{2}\sum_{k=2}^{\infty} |z|^{k-2} \leq |z|^2.$$

□

We also need estimates for the tail of the Taylor expansion of the exponential function for imaginary arguments.

Lemma 1.2. *For any* $n \geq 0$,

$$\left|\mathrm{e}^{\mathrm{i}y} - \sum_{k=0}^{n} \frac{(\mathrm{i}y)^k}{k!}\right| \leq \min\left\{2\frac{|y|^n}{n!}, \frac{|y|^{n+1}}{(n+1)!}\right\}.$$

Proof. Let $y > 0$. By partial integration,

$$\int_0^y \mathrm{e}^{\mathrm{i}x}(y-x)^k \,\mathrm{d}x = \frac{y^{k+1}}{k+1} + \frac{\mathrm{i}}{k+1}\int_0^y \mathrm{e}^{\mathrm{i}x}(y-x)^{k+1}\,\mathrm{d}x, \quad k \geq 0\,. \tag{A.6}$$

For $k = 0$ the first formula and direct integration, respectively, yield

$$\int_0^y \mathrm{e}^{\mathrm{i}x}\,\mathrm{d}x = \begin{cases} y + \mathrm{i}\int_0^y \mathrm{e}^{\mathrm{i}x}(y-x)\,\mathrm{d}x, \\ \dfrac{\mathrm{e}^{\mathrm{i}y} - 1}{\mathrm{i}}, \end{cases}$$

so that, by equating these expressions, we obtain

$$\mathrm{e}^{\mathrm{i}y} = 1 + \mathrm{i}y + \mathrm{i}^2\int_0^y \mathrm{e}^{\mathrm{i}x}(y-x)\,\mathrm{d}x. \tag{A.7}$$

Inserting (A.6) into (A.7) iteratively for $k = 2, 3, \ldots, n-1$ (more formally, by induction), yields

$$e^{iy} = \sum_{k=0}^{n} \frac{(iy)^k}{k!} + \frac{i^{n+1}}{n!} \int_0^y e^{ix}(y-x)^n \, dx, \tag{A.8}$$

and, hence,

$$\left| e^{iy} - \sum_{k=0}^{n} \frac{(iy)^k}{k!} \right| \le \frac{1}{n!} \int_0^y (y-x)^n \, dx = \frac{y^{n+1}}{(n+1)!} \,.$$

Replacing n by $n-1$ in (A.8), and then adding and subtracting $\frac{(iy)^n}{n!}$, yields

$$\begin{aligned} e^{iy} &= \sum_{k=0}^{n-1} \frac{(iy)^k}{k!} + \frac{i^n}{(n-1)!} \int_0^y e^{ix}(y-x)^{n-1} \, dx \\ &= \sum_{k=0}^{n} \frac{(iy)^k}{k!} + \frac{i^n}{(n-1)!} \int_0^y \left(e^{ix} - 1\right)(y-x)^{n-1} \, dx, \end{aligned}$$

so that, in this case, noticing that $|e^{ix} - 1| \le 2$,

$$\left| e^{iy} - \sum_{k=0}^{n} \frac{(iy)^k}{k!} \right| \le \frac{2y^n}{n!} \,.$$

The proof is finished via the analogous estimates for $y < 0$. □

Another estimate concerns the integral $\int_0^t \frac{\sin x}{x} \, dx$ as $t \to \infty$. A slight delicacy is that the integral is not absolutely convergent. However, the successive slices $\int_{(n-1)\pi}^{n\pi} \frac{\sin x}{x} \, dx$ are alternating in sign and decreasing in absolute value to 0 as $n \to \infty$, which proves that the limit as $t \to \infty$ exists.

Lemma 1.3. *Let $\alpha > 0$. Then*

$$\int_0^t \frac{\sin \alpha x}{x} \, dx \begin{cases} \le \int_0^\pi \frac{\sin x}{x} \, dx \le \pi & \text{for all} \quad t > 0, \\ \to \frac{\pi}{2} & \text{as} \quad t \to \infty. \end{cases}$$

Proof. The change of variables $y = \alpha x$ shows that it suffices to check the case $\alpha = 1$.

The first inequality is a consequence of the behavior of the slices mentioned prior to the statement of the lemma, and the fact that $\sin x \le x$.

Since $\frac{1}{x} = \int_0^\infty e^{-yx} \, dy$, and since, for all t,

$$\int_0^t \int_0^\infty |\sin x e^{-yx}| dy \, dx \le \int_0^t \frac{|\sin x|}{x} \, dy \le \int_0^t dy = t,$$

we may apply Fubini's theorem to obtain

$$\int_0^t \frac{\sin x}{x} \, dx = \int_0^t \sin x \Big(\int_0^\infty e^{-yx} \, dy \Big) \, dx = \int_0^\infty \Big(\int_0^t \sin x e^{-yx} \, dx \Big) \, dy.$$

In order to evaluate the inner integral we use partial integration twice:

$$\begin{aligned} I_t(y) &= \int_0^t \sin x \mathrm{e}^{-yx}\,\mathrm{d}x = \big[-\cos x\mathrm{e}^{-yx}\big]_0^t - \int_0^t \cos x \cdot y\mathrm{e}^{-yx}\,\mathrm{d}x \\ &= 1 - \cos t\mathrm{e}^{-yt} - \big[\sin x \cdot y\mathrm{e}^{-yx}\big]_0^t - \int_0^t \sin x \cdot y^2\mathrm{e}^{-yx}\,\mathrm{d}x \\ &= 1 - \cos t\mathrm{e}^{-yt} - \sin t \cdot y\mathrm{e}^{-yt} - y^2 I_t(y), \end{aligned}$$

so that

$$I_t(y) = \frac{1}{1+y^2}\big(1 - \mathrm{e}^{-yt}(\cos t + y \sin t)\big).$$

Inserting this into the double integral yields

$$\int_0^t \frac{\sin x}{x}\,\mathrm{d}x = \int_0^\infty I_t(y)\,\mathrm{d}y = \frac{\pi}{2} - \int_0^\infty \frac{1}{1+y^2}\mathrm{e}^{-yt}(\cos t + y\sin t)\,\mathrm{d}y,$$

so that, finally,

$$\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \lim_{t\to\infty}\int_0^t \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2},$$

(since $\int_0^\infty \frac{1}{1+y^2}\mathrm{e}^{-yt}(\cos t + y\sin t)\,\mathrm{d}y \to 0$ as $t \to \infty$). □

2 Mill's Ratio

The function $\mathrm{e}^{-x^2/2}$, which is intimately related to the normal distribution, has no primitive function (expressable in terms of elementary functions), so that integrals must be computed numerically. In many situations, however, estimates or approximations are enough. Mill's ratio is one such result.

Lemma 2.1. *Let $\phi(x)$ be the standard normal density, and $\Phi(x)$ the corresponding distribution function. Then*

$$\Big(1 - \frac{1}{x^2}\Big)\frac{\phi(x)}{x} < 1 - \Phi(x) < \frac{\phi(x)}{x}, \quad x > 0.$$

In particular,

$$\lim_{x\to\infty}\frac{x(1-\Phi(x))}{\phi(x)} = 1.$$

Proof. Since $(\phi(x))' = -x\phi(x)$, partial integration yields

$$0 < \int_x^\infty \frac{1}{y^2}\phi(y)\,\mathrm{d}y = \frac{\phi(x)}{x} - (1-\Phi(x)).$$

Rearranging this proves the right-most inequality. Similarly,

$$0 < \int_x^\infty \frac{3}{y^4}\phi(y)\,\mathrm{d}y = \frac{\phi(x)}{x^3} - \int_x^\infty \frac{1}{y^2}\phi(y)\,\mathrm{d}y,$$

which, together with the previous estimate, proves the left-hand inequality.

The limit result follows immediately. □

Remark 2.1. If only at the upper estimate is of interest one can argue as follows:

$$1 - \Phi(x) = \int_x^\infty \frac{y}{y}\phi(y)\,\mathrm{d}y < \frac{1}{x}\int_x^\infty y\phi(y)\,\mathrm{d}y = \frac{\phi(x)}{x}. \qquad \square$$

3 Sums and Integrals

In general it is easier to integrate than to compute sums. One therefore often tries to switch from sums to integrals. Usually this is done by writing $\sum \sim \int$ or $\sum \leq C\int$, where C is some (uninteresting) constant. Following are some more precise estimates of this kind.

Lemma 3.1. (i) *For $\alpha > 0$, $n \geq 2$,*

$$\frac{1}{\alpha n^\alpha} \leq \sum_{k=n}^\infty \frac{1}{k^{\alpha+1}} \leq \frac{1}{\alpha(n-1)^\alpha} \leq \frac{2^\alpha}{\alpha n^\alpha}.$$

Moreover,

$$\lim_{n\to\infty} n^\alpha \sum_{k=n}^\infty \frac{1}{k^{\alpha+1}} = \frac{1}{\alpha} \qquad \textit{as} \quad n \to \infty.$$

(ii) *For $\beta > 0$,*

$$\frac{n^\beta}{\beta} \leq \sum_{k=1}^n k^{\beta-1} \leq \frac{n^\beta}{\beta} + n^{\beta-1} \leq \Big(\frac{1}{\beta} + 1\Big)n^\beta,$$

and

$$\lim_{n\to\infty} n^{-\beta} \sum_{k=1}^n k^{\beta-1} = \frac{1}{\beta}.$$

(iii)

$$\log n + \frac{1}{n} \leq \sum_{k=1}^n \frac{1}{k} \leq \log n + 1,$$

and

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} = 1.$$

Proof. (i): We have

$$\frac{1}{\alpha n^\alpha} = \int_n^\infty \frac{\mathrm{d}x}{x^{\alpha+1}} = \sum_{k=n}^\infty \int_k^{k+1} \frac{\mathrm{d}x}{x^{\alpha+1}} \begin{cases} \leq \sum_{k=n}^\infty \frac{1}{k^{\alpha+1}}, \\ \geq \sum_{k=n}^\infty \frac{1}{(k+1)^{\alpha+1}}. \end{cases}$$

The proof of (ii) follows the same pattern by departing from

$$\frac{n^\beta}{\beta} = \int_0^n x^{\beta-1}\,\mathrm{d}x,$$

however, the arguments for $\beta > 1$ and $0 < \beta < 1$ have to be worked out separately.

The point of departure for (iii) is

$$\log n = \int_1^n \frac{\mathrm{d}x}{x}. \qquad \square$$

Exercise 3.1. Finish the proof of the lemma. □

Remark 3.1. An estimate which is sharper than (iii) is

$$\sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + o(1) \quad \text{as} \quad n \to \infty,$$

where $\gamma = 0.5772\ldots$ is *Euler's constant*. However, the corresponding limit coincides with that of the lemma. □

4 Sums and Products

There is a strong connection between the convergence of sums and that of products, for example through the formula

$$\prod(\cdots) = \exp\Big\{\sum \log(\cdots)\Big\}.$$

One can transform criteria for convergence of sums into criteria for convergence of products, and vice versa, essentially via this connection. For example, if a sum converges, then the tails are small. For a product this means that the tails are close to 1. Here are some useful connections.

Lemma 4.1. *For $n \geq 1$, let $0 \leq a_n < 1$. Then*

$$\sum_{n=1}^{\infty} a_n \text{ converges} \quad \Longleftrightarrow \quad \prod_{n=1}^{\infty}(1-a_n) \text{ converges.}$$

Convergence thus holds iff

$$\sum_{k=m}^{n} a_k \to 0 \quad \Longleftrightarrow \quad \prod_{k=m}^{n}(1-a_n) \to 1 \quad \text{as} \quad m, n \to \infty.$$

Proof. Taking logarithms, shows that the product converges iff

$$\sum_{n=1}^{\infty} \log(1-a_n) < \infty.$$

No matter which of the sums we assume convergent, we must have $a_n \to 0$ as $n \to \infty$, so we may assume, without restriction, that $a_n < 1/3$ for all n. Formula (A.2) with $\delta = 1/3$ then tells us that

$$-\tfrac{3}{2}a_n \leq \log(1 - a_n) \leq -a_n. \qquad \square$$

Lemma 4.2. *For $n \geq 1$, let $0 \leq a_n < \delta < 1$. Then*

$$(1 - a_n)^n \to 1 \quad \textit{as} \quad n \to \infty \quad \Longleftrightarrow \quad na_n \to 0 \quad \textit{as} \quad n \to \infty.$$

Moreover, in either case, given $\delta \in (0, 1)$, we have $na_n < \delta(1 - \delta) < 1$ for n large enough, and

$$(1 - \delta)na_n \leq 1 - (1 - a_n)^n \leq na_n/(1 - \delta).$$

Proof. The sufficiency is well known. Therefore, suppose that $(1 - a_n)^n \to 1$ as $n \to \infty$. Recalling (A.2),

$$1 \leftarrow (1-a_n)^n = \exp\{n \log(1-a_n)\} \begin{cases} \leq \exp\{-na_n\}, \\ \geq \exp\{-na_n/(1 - \delta)\}, \end{cases} \quad \text{as} \quad n \to \infty,$$

which establishes the preliminary fact that

$$(1 - a_n)^n \to 1 \quad \text{as} \quad n \to \infty \quad \Longleftrightarrow \quad na_n \to 0 \quad \text{as} \quad n \to \infty.$$

With this in mind we return to the upper bound and apply (A.1). Choose n so large that $na_n < \delta(1 - \delta) < 1$. Then,

$$(1 - a_n)^n \leq \exp\{-na_n\} \leq 1 - na_n + (na_n)^2 \leq 1 - na_n(1 - \delta).$$

For the lower bound there is a simpler way out;

$$(1 - a_n)^n \geq \exp\{-na_n/(1 - \delta)\} \geq 1 - na_n/(1 - \delta).$$

The double inequality follows by joining the upper and lower bounds. □

5 Convexity; Clarkson's Inequality

Convexity plays an important role in many branches of mathematics. Our concern here is some inequalities, such as generalizations of the triangle inequality.

Definition 5.1. *A real valued function g is* convex *iff, for every $x, y \in \mathbb{R}$, and $\alpha \in [0, 1]$,*

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y).$$

The function is concave *if the inequality is reversed.* □

In words, g is convex if a chord joining two points lies on, or above, the function between those points.

Convex functions always possess derivatives from the left and from the right. The derivatives agree on almost all points. The typical example is $|x|$, which is convex but does not possess a derivative at 0, only left- and right-hand ones.

A twice differentiable function is convex if and only if the second derivative is non-negative (and concave if and only if it is non-positive).

For $x, y \in \mathbb{R}$ the standard triangular inequality states that $|x+y| \le |x|+|y|$. Following are some analogs for powers.

Lemma 5.1. *Let* $r > 0$, *and suppose that* $x, y > 0$. *Then*

$$(x+y)^r \le \begin{cases} 2^r(x^r + y^r), & \textit{for} \quad r > 0, \\ x^r + y^r, & \textit{for} \quad 0 < r \le 1, \\ 2^{r-1}(x^r + y^r), & \textit{for} \quad r \ge 1. \end{cases}$$

Proof. For $r > 0$,

$$(x+y)^r \le (2\max\{x,y\})^r = 2^r(\max\{x,y\})^r \le 2^r(x^r + y^r).$$

Next, suppose that $0 < r \le 1$. Then, since $x^{1/r} \le x$ for any $0 < x < 1$, it follows that

$$\Big(\frac{x^r}{x^r + y^r}\Big)^{1/r} + \Big(\frac{y^r}{x^r + y^r}\Big)^{1/r} \le \frac{x^r}{x^r + y^r} + \frac{y^r}{x^r + y^r} = 1,$$

and, hence, that

$$x + y \le (x^r + y^r)^{1/r},$$

which is the same as the second inequality.

For $r \ge 1$ we exploit the fact that the function $|x|^r$ is convex, so that, in particular,

$$\Big(\frac{x+y}{2}\Big)^r \le \frac{1}{2}x^r + \frac{1}{2}y^r,$$

which is easily reshuffled into the third inequality. □

Lemma 5.2. *Let* $p^{-1} + q^{-1} = 1$. *Then*

$$xy \le \frac{x^p}{p} + \frac{y^q}{q} \qquad \textit{for} \quad x, y > 0.$$

Proof. The concavity of the logarithm, and the fact that e^x is increasing, yield

$$\begin{aligned} xy &= \exp\{\log xy\} = \exp\Big\{\frac{1}{p}\log x^p + \frac{1}{q}\log x^q\Big\} \\ &\le \exp\Big\{\log\Big(\frac{x^p}{p} + \frac{y^q}{q}\Big)\Big\} = \frac{x^p}{p} + \frac{y^q}{q}\,. \end{aligned}$$ □

Remark 5.1. The numbers p and q are called *conjugate exponents.*

Remark 5.2. The case $p = q = 2$ is special, in the sense that the number 2 is the same as its conjugate. (This has a number of special consequences within the theory of functional analysis.) In this case the inequality above becomes

$$xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2,$$

which, on the other hand, is equivalent to the inequality $(x - y)^2 \geq 0$. □

Clarkson's inequality [50] generalizes the well-known parallelogram identity, which states that

$$|x + y|^2 + |x - y|^2 = 2(|x|^2 + |y|^2) \quad \text{for} \quad x, y \in \mathbb{R},$$

in the same vein as the Lemma 5.1 is a generalization of the triangle inequality. We shall need the following part of [50], Theorem 2.

Lemma 5.3. (Clarkson's inequality) *Let* $x, y \in \mathbb{R}$. *Then*

$$|x + y|^r + |x - y|^r \begin{cases} \leq 2(|x|^r + |y|^r), & \text{for} \quad 1 \leq r \leq 2, \\ \geq 2(|x|^r + |y|^r), & \text{for} \quad r \geq 2. \end{cases}$$

Proof. For $r = 1$ this is just a consequence of the triangular inequality. For $r = 2$ it is the parallelogram identity. We therefore assume that $r \neq 1, 2$ in the following.

First, let $1 < r < 2$. If $x = y$, or if one of x and y equals 0, the result is trivial. Moreover, if the inequality is true for x and y, then it is also true for $\pm x$ and $\pm y$. We therefore suppose, without restriction, that $0 < y < x$. Putting $a = y/x$ reduces our task to verifying that

$$(1 + a)^r + (1 - a)^r \leq 2(1 + a^r) \quad \text{for} \quad 0 < a < 1.$$

Toward that end, set $g(a) = 2(1 + a^r) - (1 + a)^r - (1 - a)^r$. We wish to prove that $g(a) \geq 0$ for $0 < a < 1$. Now,

$$\begin{aligned} g'(a) &= 2ra^{r-1} - r(1 + a)^{r-1} + r(1 - a)^{r-1} \\ &= r(2 - 2^{r-1})a^{r-1} + r\big((2a)^{r-1} + (1 - a)^{r-1} - (1 + a)^{r-1}\big) \geq 0. \end{aligned}$$

Here we have used the fact that $0 < r - 1 < 1$ to conclude that the first expression is non-negative, and Lemma 5.1 for the second one.

Next, let $r > 2$. The analogous argument leads to proving that

$$(1 + a)^r + (1 - a)^r \geq 2(1 + a^r) \quad \text{for} \quad 0 < a < 1,$$

so that with $g(a) = (1 + a)^r + (1 - a)^r - 2(1 + a^r)$,

$$\begin{aligned} g'(a) &= r\big((1+a)^{r-1} - (1-a)^{r-1} - 2a^{r-1}\big), \quad \text{and} \\ g''(a) &= r(r-1)\big((1+a)^{r-2} - (1-a)^{r-2} - 2a^{r-2}\big) \\ &= r(r-1)\big([(1+a)^{r-2} + (1-a)^{r-2} - 2] + 2[1 - a^{r-2}]\big) \geq 0, \end{aligned}$$

beacuse the first expression in brackets is non-negative by convexity;

$$\frac{1}{2}(1+a)^{r-2} + \frac{1}{2}(1-a)^{r-2} \geq \Big(\frac{(1+a)+(1-a)}{2}\Big)^{r-2} = 1,$$

and because the second expression is trivially non-negative (since $r - 2 > 0$).

Now, $g''(0) = 0$ and $g''(a) \geq 0$ implies that g' is non-decreasing. Since $g'(0) = 0$, it follows that g' is non-negative, so that g is non-decreasing, which, finally, since $g(0) = 0$, establishes the non-negativity of g. □

Remark 5.3. A functional analyst would probably say that (ii) trivially follows from (i) by a standard duality argument. □

6 Convergence of (Weighted) Averages

A fact that is frequently used, often without further ado, is that (weighted) averages of convergent sequences converge (too). After all, by abuse of language, this is pretty “obvious”. Namely, if the sequence is convergent, then, after a finite number of elements, the following ones are all close to the limit, so that the average is essentially equal to the average of the last group; the first couple of terms do not matter in the long run. But intuition and proof are not the same. However, a special feature here is that it is unusually transparent how the proof is, literally, a translation of intuition and common sense into formulas.

Lemma 6.1. *Suppose that $a_n \in \mathbb{R}$, $n \geq 1$. If $a_n \to a$ as $n \to \infty$, then*

$$\frac{1}{n}\sum_{k=1}^{n} a_k \to a \qquad \textit{as} \quad n \to \infty.$$

If, in addition, $w_k \in \mathbb{R}^+$, $k \geq 1$, and $B_n = \sum_{k=1}^{n} w_k$, $n \geq 1$, then

$$\frac{1}{B_n}\sum_{k=1}^{n} w_k a_k \to 0 \qquad \textit{as} \quad n \to \infty.$$

Proof. It is no restriction to assume that $a = 0$ (otherwise consider the sequence $a_n - a$, $n \geq 1$). Thus, given an arbitrary $\varepsilon > 0$, we know that $|a_n| < \varepsilon$ as soon as $n > n_0 = n_0(\varepsilon)$. It follows that, for $n > n_0$,

$$\Big|\frac{1}{n}\sum_{k=1}^{n} a_k\Big| \leq \Big|\frac{1}{n}\sum_{k=1}^{n_0} a_k\Big| + \frac{n-n_0}{n}\Big|\frac{1}{n-n_0}\sum_{k=n_0+1}^{n} a_k\Big| \leq \frac{1}{n}\sum_{k=1}^{n_0}|a_k| + \varepsilon,$$

so that

$$\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{k=1}^{n_0} a_k\right| \le \varepsilon,$$

which does it, since ε can be made arbitrarily small.

This proves the first statement. The second one follows similarly. □

Exercise 6.1. Carry out the proof of the second half of the lemma. □

Example 6.1. If $a_n \to a$ as $n \to \infty$, then, for example,

$$\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}\,a_k \to a,$$

$$\frac{1}{\log\log n}\sum_{k=1}^{n}\frac{1}{k\log k}\,a_k \to a,$$

$$n^{\alpha}\sum_{k=1}^{n}\frac{1}{k^{\alpha+1}}\,a_k \to \frac{a}{\alpha}, \quad \alpha > 0.$$

□

Next, an important further development in this context.

Lemma 6.2. (Kronecker's lemma) *Suppose that $x_n \in \mathbb{R}$, $n \ge 1$, set $a_0 = 0$, and let a_n, $n \ge 1$, be positive numbers increasing to $+\infty$. Then*

$$\sum_{n=1}^{\infty}\frac{x_n}{a_n} \text{ converges} \quad\Longrightarrow\quad \frac{1}{a_n}\sum_{k=1}^{n} x_k \to 0 \quad \text{as} \quad n\to\infty.$$

Proof. The essential tools are partial summation and Lemma 6.1.

Set, $b_0 = 0$, and, for $1 \le n \le \infty$,

$$b_n = \sum_{k=1}^{n}\frac{x_k}{a_k}.$$

Since $x_k = a_k(b_k - b_{k-1})$ for all k, it follows by partial summation that

$$\frac{1}{a_n}\sum_{k=1}^{n} x_k = b_n - \frac{1}{a_n}\sum_{k=0}^{n-1}(a_{k+1}-a_k)b_k.$$

Now, $b_n \to b_\infty$ as $n \to \infty$ by assumption, and

$$\frac{1}{a_n}\sum_{k=0}^{n-1}(a_{k+1}-a_k)b_k \to b_\infty \quad \text{as} \quad n\to\infty,$$

by the second half of Lemma 6.1, since we are faced with a weighted average of quantities tending to b_∞. □

Example 6.2. Let $x_n \in \mathbb{R}$, $n \geq 1$. Then

$$\sum_{n=1}^{\infty} \frac{x_n}{n} \text{ converges} \quad \Longrightarrow \quad \bar{x}_n = \frac{1}{n}\sum_{k=1}^{n} x_k \to 0 \quad \text{as} \quad n \to \infty. \qquad \square$$

The following continuous version is proved similarly.

Lemma 6.3. *Suppose that $\{g_n,\, n \geq 1\}$ are real valued continuous functions such that $g_n \to g$ as $n \to \infty$, where g is continuous in a neighborhood of $b \in \mathbb{R}$. Then, for every $\varepsilon > 0$, there exists $h_0 > 0$, such that*

$$\limsup_{n\to\infty} \Big| \frac{1}{2h} \int_{|x-b|<h} g_n(x)\,\mathrm{d}x - g(b) \Big| < \varepsilon \quad \textit{for all} \quad h \in (0, h_0).$$

Exercise 6.2. Prove the lemma.

Exercise 6.3. State and prove version for weighted averages. □

7 Regularly and Slowly Varying Functions

Regularly and slowly varying functions were introduced by Karamata [152]. Since then, the theory has become increasingly important in probability theory. For more on the topic we refer to [21, 88, 123, 217].

Definition 7.1. *Let $a > 0$. A positive measurable function u on $[a, \infty)$* varies regularly *at infinity with exponent ρ, $-\infty < \rho < \infty$, denoted $u \in \mathcal{RV}(\rho)$, iff*

$$\frac{u(tx)}{u(t)} \to x^{\rho} \quad \textit{as} \quad t \to \infty \quad \textit{for all} \quad x > 0.$$

If $\rho = 0$ the function is slowly varying *at infinity; $u \in \mathcal{SV}$.* □

Typical examples of regularly varying functions are

$$x^{\rho}, \quad x^{\rho}\log^+ x, \quad x^{\rho}\log^+\log^+ x, \quad x^{\rho}\frac{\log^+ x}{\log^+\log^+ x}, \qquad \text{and so on.}$$

Typical slowly varying functions are the above when $\rho = 0$. Moreover, every positive function with a finite limit as $x \to \infty$ is slowly varying. Regularly varying functions with a non-zero exponent are ultimately monotone.

Exercise 7.1. Check that the typical functions behave as claimed. □

The following lemma contains some elementary properties of regularly and slowly varying functions. The first two are a bit harder to verify, so we refer to the literature for them. The three others follow, essentially, from the definition and the previous lemma.

Lemma 7.1. *Let $u \in \mathcal{RV}(\rho)$ be positive on the positive half-axis.*

(a) *If* $-\infty < \rho < \infty$, *then* $u(x) = x^\rho \ell(x)$, *where* $\ell \in \mathcal{SV}$.
If, in addition, u *has a monotone derivative* u', *then*

$$\frac{xu'(x)}{u(x)} \to \rho \quad \text{as} \quad x \to \infty.$$

If, moreover, $\rho \neq 0$, *then* $\operatorname{sgn}(u) \cdot u' \in \mathcal{RV}(\rho - 1)$.

(b) *Let* $\rho > 0$, *and set* $u^{-1}(y) = \inf\{x : u(x) \geq y\}$, $y > 0$. *Then* $u^{-1} \in \mathcal{RV}(1/\rho)$.

(c) $\log u \in \mathcal{SV}$.

(d) *Suppose that* $u_i \in \mathcal{RV}(\rho_i)$, $i = 1, 2$. *Then* $u_1 + u_2 \in \mathcal{RV}(\max\{\rho_1, \rho_2\})$.

(e) *Suppose that* $u_i \in \mathcal{RV}(\rho_i)$, $i = 1, 2$, *that* $u_2(x) \to \infty$ *as* $x \to \infty$, *and set* $u(x) = u_1(u_2(x))$. *Then* $u \in \mathcal{RV}(\rho_1 \cdot \rho_2)$. *In particular, if one of* u_1 *and* u_2 *is slowly varying, then so is the composition.*

Proof. As just mentioned, we omit (a) and (b).
(c): The fact that $\frac{u(tx)}{u(t)} \to x^\rho$ as $t \to \infty$ yields

$$\frac{\log u(tx)}{\log u(t)} = \frac{\log \frac{u(tx)}{u(t)}}{\log u(t)} + 1 \to 0 + 1 \quad \text{as} \quad t \to \infty.$$

(d): Suppose that $\rho_1 > \rho_2$. Then

$$\begin{aligned}\frac{u_1(tx) + u_2(tx)}{u_1(t) + u_2(t)} &= \frac{u_1(tx)}{u_1(t)} \cdot \frac{u_1(t)}{u_1(t) + u_2(t)} + \frac{u_2(tx)}{u_2(t)} \cdot \frac{u_2(t)}{u_1(t) + u_2(t)} \\ &\to x^{\rho_1} \cdot 1 + x^{\rho_2} \cdot 0 = x^{\rho_1} \quad \text{as} \quad t \to \infty.\end{aligned}$$

If $\rho_1 < \rho_2$ the limit equals x^{ρ_2}, and if the exponents are equal (to ρ) the limit becomes x^ρ.
(e): An application of Lemma 7.2 yields

$$\frac{u(tx)}{u(t)} = \frac{u_1(u_2(tx))}{u_1(u_2(t))} = \frac{u_1(\frac{u_2(tx)}{u_2(t)} \cdot u_2(t))}{u_1(u_2(t))} \to \left(x^{\rho_2}\right)^{\rho_1} = x^{\rho_1 \cdot \rho_2} \quad \text{as} \quad t \to \infty. \; \square$$

Remark 7.1. Notice that (c) is contained in (e). □

In the definition of regular and slow variation the ratio between the arguments of the function is constant. However, the limits remain the same if the ratio converges to a constant.

Lemma 7.2. *Suppose that* $u \in \mathcal{RV}(\rho)$, $-\infty < \rho < \infty$, *and, in addition, that* u *is (ultimately) monotone if* $\rho = 0$. *Moreover, let, for* $n \geq 1$, $a_n, b_n \in \mathbb{R}^+$ *be such that*

$$a_n, b_n \to \infty \quad \text{and} \quad \frac{a_n}{b_n} \to c \quad \text{as} \quad n \to \infty \qquad (c \in (0, \infty)).$$

Then

$$\frac{u(a_n)}{u(b_n)} \to \begin{cases} 1, & \text{for} \quad \rho = 0, \\ c^\rho, & \text{for} \quad \rho \neq 0, \end{cases} \quad \text{as} \quad n \to \infty.$$

Proof. Suppose that $\rho > 0$, so that u is ultimately non-decreasing, let $\varepsilon > 0$, and choose n large enough to ensure that

$$b_n(c-\varepsilon) < a_n < b_n(c+\varepsilon).$$

Then

$$\frac{u((c-\varepsilon)b_n)}{u(b_n)} \leq \frac{u(a_n)}{u(b_n)} \leq \frac{u((c+\varepsilon)b_n)}{u(b_n)},$$

from which the conclusion follows from the fact that the extreme members converge to $(c \pm \varepsilon)^\rho$ as $n \to \infty$.

The case $\rho < 0$ is similar; the inequalities are reversed. In the slowly varying case ($\rho = 0$) the extreme limits are equal to 1 ($= c^\rho$). □

Suppose that $u \in \mathcal{RV}(\rho)$, where $\rho > -1$. Then, since the slowly varying component is "negligible" with respect to x^ρ, it is reasonable to believe that the integral of u is regularly varying with exponent $\rho + 1$. The truth of this fact, which is supported by Lemma 7.1(a) in conjunction with Lemma 3.1, is the first half of the next result.

Lemma 7.3. *Let* $\rho > -1$.

(i) *If* $u \in \mathcal{RV}(\rho)$, *then* $\quad U(x) = \int_a^x u(y)\,\mathrm{d}y \in \mathcal{RV}(\rho+1)$.

(ii) *If* $\ell \in \mathcal{SV}$, *then* $\quad \sum_{j \leq n} j^\rho \ell(j) \sim \frac{1}{\rho+1} n^{\rho+1}\ell(n)$ *as* $n \to \infty$.

There also exists something called *rapid variation*, corresponding to $\rho = +\infty$. A function u is *rapidly varying at infinity* iff

$$\frac{u(tx)}{u(t)} \to \begin{cases} 0, & \text{for } 0 < x < 1, \\ \infty, & \text{for } x > 1, \end{cases} \qquad \text{as} \quad t \to \infty.$$

This means that u increases faster than any power at infinity. The exponential function e^x is one example.

8 Cauchy's Functional Equation

This is a well known equation that enters various proofs. If g is a real valued *additive* function, that is,

$$g(x+y) = g(x) + g(y),$$

then it is immediate that $g(x) = cx$ is a solution for any $c \in \mathbb{R}$. The problem is: Are there any other solutions? Yes, there exist pathological ones if nothing more is assumed. However, under certain regularity conditions this is the only solution.

Lemma 8.1. *Suppose that g is real valued and additive on an arbitrary interval $I \subset \mathbb{R}$, and satisfies one of the following conditions:*

- *g is continuous;*
- *g is monotone;*
- *g is bounded.*

Then $g(x) = cx$ for some $c \in \mathbb{R}$.

Proof. For $x = y$ we find that $g(2x) = 2g(x)$, and, by induction, that

$$g(n) = ng(1) \quad \text{and} \quad g(1) = ng(1/n).$$

Combining these facts for $r = m/n \in \mathbb{Q}$ tells us that

$$g(r) = g(m/n) = mg(1/n) = m\big(g(1)/n\big) = rg(1),$$

and that

$$g(rx) = rg(x) \quad \text{for any} \quad x.$$

The remaining problem is to glue all x-values together.

Set $c = g(1)$. If g is continuous, the conclusion follows from the definition of continuity; for any $x \in \mathbb{R}$ there exists, for any given $\delta > 0$, $r \in \mathbb{Q}$, such that $|r - x| < \delta$, which implies that $|g(x) - g(r)| < \varepsilon$, so that

$$|g(x) - cx| \leq |g(x) - g(r)| + c|r - x| \leq \varepsilon + c\delta.$$

The arbitrariness of ε and δ completes the proof.

If g is monotone, say non-decreasing, then, for $r_1 < x < r_2$, where $r_2 - r_1 < \delta$,

$$cr_1 = g(r_1) \leq g(x) \leq g(r_2) = cr_2,$$

so that

$$|g(x) - cx| \leq c(r_2 - x) + c(x - r_1) = c(r_2 - r_1) < c\delta.$$

Finally, if g is bounded, it follows, in particular, that, for any given $\delta > 0$, there exists A, such that,

$$|g(x)| \leq A \quad \text{for} \quad |x| < \delta.$$

For $|x| < \delta/n$, this implies that

$$|g(x)| = |g(nx)/n| \leq \frac{A}{n}.$$

Next, let $x \in I$ be given, and choose $r \in \mathbb{Q}$, such that $|r - x| < \delta/n$. Then

$$\begin{aligned} |g(x) - cx| &= |g(x - r) + g(r) - cr - c(x - r)| = |g(x - r) - c(x - r)| \\ &\leq |g(x - r) - c(x - r)| \leq |g(x - r)| + c|x - r| \\ &\leq \frac{A}{n} + c\frac{\delta}{n} = \frac{C}{n}, \end{aligned}$$

which can be made arbitrarily small by choosing n sufficiently large. □

The following lemma contains variations of the previous one. For example, what happens if g is multiplicative?

Lemma 8.2. *Let g be a real valued function defined on some interval $I \subset \mathbb{R}^+$, and suppose that g is continuous, monotone or bounded.*

(a) *If $g(xy) = g(x) + g(y)$, then $g(x) = c \log x$ for some $c \in \mathbb{R}$.*

(b) *If $g(xy) = g(x)g(y)$, then $g(x) = x^c$ for some $c \in \mathbb{R}$.*

(c) *If $g(x + y) = g(x)g(y)$, then $g(x) = \mathrm{e}^{cx}$ for some $c \in \mathbb{R}$.*

Remark 8.1. The relation in (b) is called the *Hamel* equation. □

Proof. (a): A change of variable yields

$$g(\mathrm{e}^{x+y}) = g(\mathrm{e}^x \mathrm{e}^y) = g(\mathrm{e}^x) + g(\mathrm{e}^y),$$

so that, by Lemma 8.1, $g(\mathrm{e}^x) = cx$, which is the same as $g(x) = c \log x$.
(b): In this case a change of variables yields

$$\log g(\mathrm{e}^{x+y}) = \log g(\mathrm{e}^x \mathrm{e}^y) = \log \big(g(\mathrm{e}^x) \cdot g(\mathrm{e}^y)\big) = \log g(\mathrm{e}^x) + \log g(\mathrm{e}^y),$$

so that $\log g(\mathrm{e}^x) = cx$, and, hence, $g(x) = \mathrm{e}^{c \log x} = x^c$.
(c): We reduce to (b) via

$$g(\log xy) = g(\log x + \log y) = g(\log x)g(\log y),$$

so that $g(\log x) = x^c$, and, hence, $g(x) = \mathrm{e}^{cx}$. □

9 Functions and Dense Sets

Many proofs are based on the fact that it suffices to prove the desired result on a dense set. Others exploit the fact that the functions under consideration can be arbitrarily well approximated by other functions that are easier to handle; we mentioned this device in Chapter 1 in connection with Theorem 1.2.4 and the magic that "it suffices to check rectangles". In this section we collect some results which rectify some such arguments.

Definition 9.1. *Let A and B be sets. The set A is* dense *in B if the closure of A equals B; if $\bar{A} = B$.* □

The typical example one should have in mind is when $B = [0, 1]$ and $A = \mathbb{Q} \cap [0, 1]$:

$$\overline{\mathbb{Q} \cap [0, 1]} = [0, 1].$$

Definition 9.2. *Consider the following classes of real valued functions:*

- *C = the continuous functions;*
- *C_0 = the functions in $\mathcal{C}$ tending to 0 at $\pm\infty$;*

- *$C[a,b]$ = the functions in $\mathcal{C}$ with support on the interval $[a,b]$;*
- *D = the right-continuous, functions with left-hand limits;*
- *D^+ = the non-decreasing functions in D;*
- *$\mathbb{J}_G$ = the discontinuities of $G \in D$.* □

Lemma 9.1. (i) *If $G \in D^+$, then $\mathbb{J}_G$ is at most countable.*
(ii) *Suppose that $G_i \in D^+$ $i = 1, 2$, and that $G_1 = G_2$ on a dense subset of the reals. Then $G_1 = G_2$ for all reals.*

Proof. (i): Suppose, w.l.o.g. that $0 \leq G \leq 1$. Let, for $n \geq 1$,

$$\mathbb{J}_G^{(n)} = \Big\{x : G \text{ has a jump at } x \text{ of magnitude } \in \Big(\frac{1}{n+1}, \frac{1}{n}\Big]\Big\}.$$

The total number of points in $\mathbb{J}_G^{(n)}$ is at most equal to $n+1$, since G is non-decreasing and has total mass 1. The conclusion then follows from the fact that

$$\mathbb{J}_G = \bigcup_{n=1}^{\infty} \mathbb{J}_G^{(n)}.$$

(ii): We first show that a function in D^+ is determined by its values on a dense set. Thus, let D be dense in $\mathbb{R}$ (let $D = \mathbb{Q}$, for example), let $G_D \in D^+$ be defined on D, and set

$$G(x) = \inf_{\substack{y > x \\ y \in D}} G_D(y). \tag{A.9}$$

To prove that $G \in D^+$ we observe that the limits of G_D and G as $x \to \pm\infty$ coincide and that G is non-decreasing, so that the only problem is to prove right-continuity.

Let $x \in \mathbb{R}$ and $\varepsilon > 0$ be given, and pick $y \in D$ such that

$$G_D(y) \leq G(x) + \varepsilon.$$

Moreover, by definition, $G(y) \leq G_D(u)$ for all $u \geq y$, so that, in particular,

$$G(y) \leq G_D(u) \quad \text{for any} \quad u \in (x, y).$$

Combining this with the previous inequality proves that

$$G(u) \leq G(x) + \varepsilon \quad \text{for all} \quad u \in (x, y).$$

The monotonicity of G, and the fact that u may be chosen arbitrarily close to x, now together imply that

$$G(x) \leq G(x+) \leq G(x) + \varepsilon,$$

which, due to the arbitrariness of ε, proves that $G(x) = G(x+)$, so that $G \in D^+$.

Finally, if two functions in D^+ agree on a dense set, then the extensions to all of $\mathbb{R}$ via (A.9) does the same thing to both functions, so that they agree everywhere. □

Lemma 9.2. *Let G and $G_n \in D^+$, $n \geq 1$, and let $J(x) = G(x) - G(x-)$ and $J_n(x) = G_n(x) - G_n(x-)$ denote the jumps of G and G_n, respectively, at x.*
(i) *Suppose that $G \in D^+ \cap C[a,b]$, where $-\infty < a < b < \infty$. If*

$$G_n(x) \to G(x) \quad \text{as} \quad n \to \infty, \quad \text{for all} \quad x \in D,$$

then $G_n \to G$ uniformly on $[a,b]$;

$$\sup_{a \leq x \leq b} |G_n(x) - G(x)| \to 0 \quad \text{as} \quad n \to \infty.$$

Moreover,

$$\sup_{\substack{a \leq x \leq b \\ x \in \mathbb{J}_G}} |J_n(x)| \to 0 \quad \text{as} \quad n \to \infty.$$

(ii) *Suppose that $G \in D^+ \cap C$. If, for some dense subset $D \subset \mathbb{R}$,*

$$G_n(x) \to G(x) \quad \text{as} \quad n \to \infty, \quad \text{for all} \quad x \in D,$$
$$G_n(\pm\infty) \to G(\pm\infty) \quad \text{as} \quad n \to \infty,$$

then

$$\sup_{x \in \mathbb{R}} |G_n(x) - G(x)| \to 0 \quad \text{as} \quad n \to \infty,$$
$$\sup_{\substack{x \in \mathbb{R} \\ x \in \mathbb{J}_G}} |J_n(x)| \to 0 \quad \text{as} \quad n \to \infty.$$

(iii) *Suppose that $G \in D^+$. If, for some dense subset $D \subset \mathbb{R}$,*

$$G_n(x) \to G(x) \quad \text{as} \quad n \to \infty, \quad \text{for all} \quad x \in D,$$
$$J_n(x) \to J(x) \quad \text{as} \quad n \to \infty, \quad \text{for all} \quad x \in \mathbb{J}_G,$$
$$G_n(\pm\infty) \to G(\pm\infty) \quad \text{as} \quad n \to \infty,$$

then

$$\sup_{x \in \mathbb{R}} |G_n(x) - G(x)| \to 0 \quad \text{as} \quad n \to \infty,$$
$$\sup_{x \in \mathbb{J}_G} |J_n(x)| \to 0 \quad \text{as} \quad n \to \infty.$$

Proof. (i): Since G is continuous on a bounded interval it is uniformly continuous. Thus, for any $\varepsilon > 0$, there exists $\delta > 0$, such that

$$\omega_G(\delta) = \sup_{\substack{a \leq x,y \leq b \\ |x-y|<\delta}} |G(x) - G(y)| < \varepsilon.$$

Given the above ε and the accompanying δ we let $a = y_0 < y_1 < \cdots < y_m = b$, such that $y_k - y_{k-1} < \delta$ for all k. For any $x \in [y_{k-1} y_k]$, $1 \leq k \leq m$, it then follows that

$$G_n(y_{k-1}) - G(x) \le G_n(x) - G(x) \le G_n(y_k) - G(x),$$

so that

$$\begin{aligned}|G_n(x) - G(x)| &\le |G_n(y_{k-1}) - G(y_{k-1})| + |G_n(y_k) - G(y_k)| \\ &\qquad + |G(y_{k-1}) - G(x)| + |G(y_k) - G(x)| \\ &\le 2 \max_{1\le k\le m} |G_n(y_k) - G(y_k)| + 2\omega_G(\delta),\end{aligned}$$

so that

$$\limsup_{n\to\infty} \sup_{a\le x\le b} |G_n(x) - G(x)| \le 2\omega_G(\delta) \le 2\varepsilon.$$

As for the second statement, noticing that $J(x) = 0$, we obtain

$$\begin{aligned}\sup_{x\in\mathbb{J}} |G_n(x)| &\le \sup_{x\in\mathbb{J}} \big(|G_n(x) - G(x)| + |G(x-) - G_n(x)|\big) \\ &\le 2 \sup_{a\le x\le b} |G_n(x) - G(x)| \to 0 \quad \text{as} \quad n \to \infty.\end{aligned}$$

(ii): Since convergence at the infinities is assumed, we have, using (i),

$$\begin{aligned}\sup_{x\in\mathbb{R}} |G_n(x) - G(x)| &\le \sup_{|x|>A} |G_n(x) - G(x)| + \sup_{|x|\le A} |G_n(x) - G(x)| \\ &\le \big(G(\infty) - G_n(A)\big) + \big(G_n(-A) - G(-\infty)\big) \\ &\qquad + \big(G(\infty) - G(A)\big) + \big(G(-A) - G(-\infty)\big) \\ &\qquad + \sup_{|x|\le A} |G_n(x) - G(x)|.\end{aligned}$$

Thus, for $\pm A \in C(G)$, we obtain, recalling (i),

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}} |G_n(x) - G(x)| \le 2\big(G(\infty) - G(A)\big) + \big(G(-A) - G(-\infty)\big) + 0,$$

which can be made arbitrarily small by letting $A \to \infty$.

The second statement follows as in (i).

(iii): Assume that the conclusion does not hold, that is, suppose that there exist $\varepsilon > 0$ and a subsequence $\{n_k,\, k \ge 1\}$, $n_k \nearrow \infty$ as $k \to \infty$, such that

$$|G_{n_k}(x_k) - G(x_k)| > \varepsilon \quad \text{for all} \quad k.$$

The first observation is that we cannot have $x_k \to \pm\infty$, because of the second assumption, which means that $\{x_k,\, k \ge 1\}$ is bounded, which implies that there exists a convergent subsequence, $x_{k_j} \to x$, say, as $j \to \infty$. By diluting it further, if necessary, we can make it monotone. Since convergence can occur from above and below, and $G_{n_{k_j}}(x_{k_j})$ can be smaller as well as larger than $G(x_{k_j})$ we are faced with four different cases as $j \to \infty$:

- $x_{k_j} \searrow x, \quad \text{and} \quad G(x_{k_j}) - G_{n_{k_j}}(x_{k_j}) > \varepsilon;$

- $x_{k_j} \searrow x$, and $G_{n_{k_j}}(x_{k_j}) - G(x_{k_j}) > \varepsilon$;
- $x_{k_j} \nearrow x$, and $G(x_{k_j}) - G_{n_{k_j}}(x_{k_j}) > \varepsilon$;
- $x_{k_j} \nearrow x$, and $G_{n_{k_j}}(x_{k_j}) - G(x_{k_j}) > \varepsilon$.

Choose $r_1, r_2 \in D$, such that $r_1 < x < r_2$. In the first case this leads to

$$\begin{aligned}\varepsilon &< G(x_{k_j}) - G_{n_{k_j}}(x_{k_j}) \leq G(r_2) - G_{n_{k_j}}(x)\\ &\leq G(r_2) - G(r_1) + G(r_1) - G_{n_{k_j}}(r_1) + J_{n_{k_j}}(x)\\ &\to G(r_2) - G(r_1) + 0 - J(x) \quad \text{as} \quad j \to \infty.\end{aligned}$$

Since r_1, r_2 may be chosen arbitrarily close to x from below and above, respectively, the right-hand side can be made arbitrarily close to 0 if $x \in C(G)$, and arbitrarily close to $J(x) - J(x) = 0$ if $x \in \mathbb{J}_G$, which produces the desired contradiction.

The three other cases are treated similarly:

In the second case,

$$\varepsilon < G_{n_{k_j}}(x_{k_j}) - G(x_{k_j}) \leq G_{n_{k_j}}(r_2) - G(x) \to G(r_2) - G(x) \quad \text{as} \quad j \to \infty,$$

and the contradiction follows from the right-continuity of G by choosing r_2 close to x.

In the third case,

$$\varepsilon < G(x_{k_j}) - G_{n_{k_j}}(x_{k_j}) \leq G(x-) - G_{n_{k_j}}(r_1) \to G(x-) - G(r_1) \quad \text{as} \quad j \to \infty,$$

after which we let r_1 approach $x-$.

Finally,

$$\begin{aligned}\varepsilon &< G_{n_{k_j}}(x_{k_j}) - G(x_{k_j}) \leq G_{n_{k_j}}(x-) - G(r_1)\\ &\leq -J_{n_{k_j}}(x) + G_{n_{k_j}}(r_2) - G(r_2) + G(r_2) - G(r_1)\\ &\to -J(x) + G(r_2) - G(r_1) \quad \text{as} \quad j \to \infty,\end{aligned}$$

from which the contradiction follows as in the first variant. □

We close with two approximation lemmas.

Lemma 9.3. (Approximation lemma) *Let f be a real valued function such that either*

- $f \in C[a, b]$, *or*
- $f \in C_0$.

Then, for every $\varepsilon > 0$, there exists a simple function g, such that

$$\sup_{x \in \mathbb{R}} |f(x) - g(x)| < \varepsilon.$$

Proof. Suppose that $f \in C[a,b]$, and set

$$g_1(x) = \begin{cases} \frac{k-1}{2^n}, & \text{for} \quad \frac{k-1}{2^n} \le f(x) < \frac{k}{2^n}, \quad a < x < b, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$g_2(x) = \begin{cases} \frac{k}{2^n}, & \text{for} \quad \frac{k-1}{2^n} \le f(x) < \frac{k}{2^n}, \quad a < x < b, \\ 0, & \text{otherwise}. \end{cases}$$

Then, for $i = 1, 2$,

$$|f(x) - g_i(x)| \le g_2(x) - g_1(x) = \frac{1}{2^n} < \varepsilon,$$

as soon as n is large enough. In addition, f is sandwiched between the g-functions; $g_1(x) \le f(x) \le g_2(x)$ for all x.

If $f \in C_0$, then $|f(x)| < \varepsilon$ for $|x| > b$, so that g_1 and g_2 may be defined as above for $|x| \le b$ (that is, $a = -b$) and equal to 0 otherwise. By a slight modification the sandwiching effect can also be retained. □

Lemma 9.4. *Let* $-\infty < a \le b < \infty$. *Any indicator function* $I_{(a,b]}(x)$ *can be arbitrarily well approximated by a bounded, continuous function; there exists* f_n, $n \ge 1$, $0 \le f_n \le 1$, *such that*

$$f_n(x) \to I_{(a,b]}(x) \quad \textit{for all} \quad x \in \mathbb{R}.$$

Proof. Set, for $n \ge 1$,

$$f_n(x) = \begin{cases} 0, & \text{for} \quad x \le a, \\ n(x-a), & \text{for} \quad a < x \le a + \frac{1}{n}, \\ 1, & \text{for} \quad a + \frac{1}{n} < x \le b, \\ 1 - n(x-b), & \text{for} \quad b < x \le b + \frac{1}{n}, \\ 0, & \text{for} \quad x > b. \end{cases}$$

One readily checks that $\{f_n, n \ge 1\}$ does the job. □

Exercise 9.1. Please pursue the checking. □

References

In addition to the books and papers that have been cited in the text we also refer to a number of related items that have not been explicitly referred to, but which, nevertheless, may be of interest for further studies.

1. Aaronson, J., Gilat, D., Keane, M., and de Valk, V. An algebraic construction of a class of one-dependent processes. *Ann. Probab.* **17**, 128–143 (1989).
2. de Acosta, A. A new proof of the Hartman-Wintner law of the iterated logarithm. *Ann. Probab.* **11**, 270–276 (1983).
3. de Acosta, A., and Kuelbs, J. Some new results on the cluster set $C(\{S_n/a_n\})$ and the LIL. *Ann. Probab.* **11**, 102–122 (1983).
4. Aleškevičiene, A. On the local limit theorem for the first passage time across a barrier. *Litovsk. Mat. Sb.* **XV**, 23–66 (1975).
5. Anscombe, F.J. Large sample-theory of sequential estimation. *Proc. Cambridge Philos. Soc.* **48**, 600–607 (1952).
6. Austin, D.G., Edgar, G.A., and Ionescu Tulcea, A. Pointwise convergence in terms of expectations. *Z. Wahrsch. verw. Gebiete* **30**, 17–26 (1974).
7. von Bahr, B., and Esseen, C.-G. Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.* **36**, 299–302 (1965).
8. Barbour, A.D., and Hall, P. On the rate of Poisson convergence. *Math. Proc. Camb. Phil. Soc.* **95**, 473–480 (1984).
9. Barbour, A.D., Holst, L., and Janson, S. *Poisson Approximation.* Oxford Science Publications, Clarendon Press, Oxford (1992).
10. Barndorff-Nielsen O. On the rate of growth of the partial maxima of a sequence of independent identically distributed random variables. *Math. Scand.* **9**, 383–394 (1961).
11. Bauer, H. *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie.* De Gruyter, Berlin (1968).
12. Baum, L.E., and Katz, M. On the influence of moments on the asymptotic distribution of sums of random variables. *Ann. Math. Statist.* **34**, 1042–1044 (1963).

13. Baum, L.E., and Katz, M. Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120**, 108–123 (1965).
14. van Beek, P. An application of Fourier methods to the problem of sharpening of the Berry-Esseen inequality. *Z. Wahrsch. verw. Gebiete* **23**, 183–196 (1972).
15. Berbee, H.C.P. *Random Walks with Stationary Increments and Renewal Theory.* Mathematical Centre Tracts **112**, Amsterdam (1979).
16. Berkes, I., and Phillip, W. Limit theorems for mixing sequences without rate assumptions. *Ann. Probab.* **26**, 805–831 (1978).
17. Berry, A.C. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49**, 122–136 (1941).
18. Billingsley, P. The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.* **12**, 788–792 (1961).
19. Billingsley, P. *Probability and Measure*, 3rd ed. Wiley, New York (1995).
20. Billingsley, P. *Convergence of Probability Measures*, 2nd ed. Wiley, New York (1999).
21. Bingham, N.H., Goldie, C.M., and Teugels, J.L. *Regular Variation.* Cambridge University Press, Cambridge (1987).
22. Blackwell, D. On an equation of Wald. *Ann. Math. Statist.* **17**, 84–87 (1946).
23. Blackwell, D. Extension of a renewal theorem. *Pacific J. Math.* **3**, 315–320 (1953).
24. Blum, J.R., Hanson,D.L., and Rosenblatt, J.I. On the central limit theorem for the sum of a random number of independent random variables. *Z. Wahrsch. verw. Gebiete* **1**, 389–393 (1963).
25. Bondesson, L. *Generalized Gamma Convolutions and Related Classes of Distributions and Densities.* Lecture Notes in Statistics **76**. Springer-Verlag (1992).
26. Borel, E. Sur les probabilités et leurs applications arithmétiques. *Rend. Circ. Mat. Palermo* **26**, 247–271 (1909).
27. Bradley, R.C. Basic properties of strong mixing conditions. In: E. Eberlein and M.S. Taqqu, eds., *Dependence in Probability and Statistics,* 165–192. Birkhäuser, Boston, MA (1986).
28. Bradley, R.C. *Introduction to Strong Mixing Conditions, Vol. 1.* Technical Report (March 2003 printing), Department of Mathematics, Indiana University, Bloomington. Custom Publishing of I.U., Bloomington, March 2003.
29. Bradley, R.C. *Introduction to Strong Mixing Conditions, Vol. 2.* Technical Report, Department of Mathematics, Indiana University, Bloomington. Custom Publishing of I.U., Bloomington, March 2003.
30. Bradley, R.C. *Introduction to Strong Mixing Conditions, Vol. 3.* Technical Report, Department of Mathematics, Indiana University, Bloomington. Custom Publishing of I.U., Bloomington (in preparation).
31. Breiman, L. *Probability.* Addison-Wesley, Reading, MA (1968).
32. Burkholder, D.L. Successive conditional expectations of an integrable function. *Ann. Math. Statist.* **33**, 887–893 (1962).
33. Burkholder, D.L. Martingale transforms. *Ann. Math. Statist.* **37**, 1494–1504 (1966).
34. Burkholder, D.L. Distribution function inequalities for martingales. *Ann. Probab.* **1**, 19–42 (1973).
35. Burkholder, D.L., Davis, B.J., and Gundy, R.F. Integral inequalities for convex functions of operators on martingales. In: *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2**, 223–240 (1972).

36. Burkholder, D.L., and Gundy, R.F. Extrapolation and interpolation of quasi-linear operators on martingales. *Acta Math.* **124**, 249–304 (1970).
37. Carleman, T. Les fonctions quasi-analythiques. *Collection Borel*, Gauthiers-Villars, Paris (1926).
38. Chacon, R.V. A "stopped" proof of convergence. *Adv. in Math.* **14**, 365–368 (1974).
39. Chen, L.H.Y. Poisson approximation for dependent trials. *Ann. Probab.* **3**, 534–545 (1975)
40. Chen, P.-N. Asymptotic refinement of the Berry-Esseen constant. Preprint, National Chiao Tung University, Taiwan (2002).
41. Chen, R. A remark on the tail probability of a distribution. *J. Multivariate Analysis* **8**, 328–333 (1978).
42. Chern, H.-H., Hwang, H.-K., and Yeh, Y.-N. Distribution of the number of consecutive records. *Proceedings of the Ninth International Conference "Random Structures and Algorithms" (Poznań, 1999). Random Struct. Alg.* **17**, 169–196 (2000).
43. Chow, Y.S. A martingale inequality and the law of large numbers. *Proc. Amer. Math. Soc.* **11**, 107–111 (1960).
44. Chow, Y.S. Delayed sums and Borel summability of independent, identically distributed random variables. *Bull. Inst. Math. Acad. Sinica* **1**, 207–220 (1973).
45. Chow, Y.S., and Hsiung, C.A. Limiting behaviour of $\max_{j\leq n} S_j j^{-\alpha}$ and the first passage times in a random walk with positive drift. *Bull. Inst. Math. Acad. Sinica* **4**, 35–44 (1976).
46. Chow, Y.S., Hsiung, C.A., and Lai, T.L. Extended renewal theory and moment convergence in Anscombe's theorem. *Ann. Probab.* **7**, 304–318 (1979).
47. Chung, K.L. On the maximum partial sums of sequences of independent random variables. *Trans. Amer. Math. Soc.* **64**, 205–233 (1948).
48. Chung, K.L. *A Course in Probability Theory*, 2nd ed. Academic Press, New York (1974).
49. Chung, K.L., and Erdős, P. On the application of the Borel-Cantelli lemma. *Trans. Amer. Math. Soc.* **72**, 179–186 (1952).
50. Clarkson, J.A. Uniformly convex spaces. *Trans. Amer. Math. Soc.* **40**, 396–414 (1936).
51. Cohen, P.J. *Set Theory and the Continuum Hypothesis.* W. A. Benjamin, Inc., New York (1966).
52. Cohn, D.L. *Measure Theory.* Birkhäuser, Boston, MA (1997).
53. Cox, D. *Renewal Theory.* Methuen, London (1962).
54. Cramér, H. Über eine Eigenschaft der normalen Verteilungsfunktion. *Math. Z.* **41**, 405–414 (1936).
55. Cramér, H. Sur un nouveau théorème-limite de la théorie des probabilités. *Actual. Sci. Indust.* **736**, 5–23 (1938).
56. Cramér, H. *Mathematical Methods of Statistics.* Princeton University Press, Princeton, NJ (1946).
57. Csörgő, M., and Rychlik, Z. Weak convergence of sequences of random elements with random indices. *Math. Proc. Camb. Phil. Soc.* **88**, 171–174 (1980).
58. Csörgő, M., and Rychlik, Z. Asymptotic properties of randomly indexed sequences of random variables. *Canad. J. Statist.* **9**, 101–107 (1981).
59. Davis, B. On the integrability of the martingale square function. *Israel J. Math.* **8**, 187–190 (1970).

60. Davis, B. Stopping rules for S_n/n and the class $L \log L$. *Z. Wahrsch. verw. Gebiete* **17**, 147–150 (1971).
61. Davis, J.A. Convergence rates for the law of the iterated logarithm. *Ann. Math. Statist.* **39**, 1479–1485 (1968).
62. Davis, J.A. Convergence rates for probabilities of moderate deviations (1968). *Ann. Math. Statist.* **39**, 2016–2028.
63. Donsker, M. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6** (1951).
64. Doob, J.L. Regularity properties of certain families of chance variables. *Trans. Amer. Math. Soc.* **47**, 455–486 (1940).
65. Doob, J.L. Renewal theory from the point of view of probability. *Trans. Amer. Math. Soc.* **63**, 422–438 (1948).
66. Doob, J.L. *Stochastic Processes.* Wiley, New York (1953).
67. Drosin, M. *The Bible Code.* Simon and Schuster, New York (1997).
68. Dudley, R. *Real Analysis and Probability.* Wadsworth & Brooks/Cole, Belmont, CA (1989).
69. Durrett, R. *Probability: Theory and Examples.* Wadsworth & Brooks/Cole, Belmont, CA (1991).
70. Edgar, G.A., and Sucheston, L. Amarts: A class of asymptotic martingales. *J. Multivariate Anal.* **6**, 193-221 (1976).
71. Embrechts, P., Klüppelberg, C., and Mikosch, T. *Modelling Extremal Events for Insurance and Finance.* Springer-Verlag, Berlin (1997).
72. Englund, G. Remainder term estimates for the asymptotic normality of the number of renewals. *J. Appl. Probab.* **17**, 1108–1113 (1980).
73. Englund, G. On the coupon collector's remainder term. *Z. Wahrsch. verw. Gebiete* **60**, 381–391 (1982).
74. Erdős, P. On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20**, 286–291 (1949).
75. Erdős, P. Remark on my paper "On a theorem of Hsu and Robbins". *Ann. Math. Statist.* **21**, 138 (1950).
76. Esseen, C.-G. On the Liapounoff limit of error in the theory of probability. *Ark. Mat., Astr. o. Fysik* **28A** 9, 1-19 (1942).
77. Esseen, C.-G. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.* **77**, 1–125 (1945).
78. Esseen, C.-G. A moment inequality with an application to the central limit theorem. *Skand. Aktuar. Tidskr.* **XXXIX**, 160–170 (1956).
79. Esseen, C.-G. On the Kolmogorov-Rogozin inequality for the concentration function. *Z. Wahrsch. verw. Gebiete* **5**, 210–216 (1966).
80. Esseen, C.-G. On the concentration function of a sum of independent random variables. *Z. Wahrsch. verw. Gebiete* **9**, 290–308 (1968).
81. Esseen, C.-G. On the remainder term in the central limit theorem. *Arkiv Matematik* **8**, 7–15 (1968).
82. Etemadi, N. An elementary proof of the strong law of large numbers. *Z. Wahrsch. verw. Gebiete* **55**, 119–122 (1981).
83. Feller, W. Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung. *Math. Z.* **40**, 521–559 (1935).
84. Feller, W. Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung II. *Math. Z.* **42**, 301–312 (1937).
85. Feller, W. Fluctuation theory of recurrent events. *Trans. Amer. Math. Soc.* **67**, 98–119 (1949).

86. Feller, W. An extension of the law of the iterated logarithm to variables without variance. *J. Math. Mech.* **18**, 343–355 (1968).
87. Feller, W. *An Introduction to Probability Theory and Its Applications, Vol 1*, 3rd ed. Wiley, New York (1968).
88. Feller, W. *An Introduction to Probability Theory and Its Applications, Vol 2*, 2nd ed. Wiley, New York (1971).
89. Fisher, R.A., and Tippett, L.H.C. Limiting forms of the frequency of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.* **24**, 180–190 (1928).
90. Friedman, N., Katz, M., and Koopmans, L.H. Convergence rates for the central limit theorem. *Proc. Nat. Acad. Sci. USA* **56**, 1062–1065 (1966).
91. Galambos, J. *The Asymptotic Theory of Extreme Order Statistics,* 2nd ed. Krieger, Malabar, FL (1987).
92. Garsia, A.M. *Topics in Almost Everywhere Convergence.* Markham, Chicago (1970).
93. Garsia, A.M. *Martingale Inequalities. Seminar Notes on Recent Progress.* W.A. Benjamin, Inc., Reading MA (1973).
94. Gilat, D. Every nonnegative submartingale is the absolute value of a martingale. *Ann. Probab.* **5**, 475–481 (1977).
95. Gnedenko, B.V. Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**, 423–453 (1943).
96. Gnedenko, B.V. On the local limit theorem in the theory of probability. *Uspekhi Mat. Nauk.* **3** (1948) (in Russian).
97. Gnedenko, B.V. *Theory of Probability*, 4th ed. Chelsea, New York (1967).
98. Gnedenko, B.V., and Kolmogorov, A.N. *Limit Distributions for Sums of Independent Random Variables*, 2nd ed. Addison-Wesley, Cambridge, MA (1968).
99. Grimmett, G.R., and Stirzaker, D.R. *Probability Theory and Random Processes*, 2nd ed. Oxford University Press, Oxford (1992).
100. Gumbel, E.J. *Statistics of Extremes.* Columbia University Press, New York (1958).
101. Gundy, R.F. On the class $L \log L$, martingales, and singular integrals. *Studia Math.* **XXXIII**, 109–118 (1969).
102. Gundy, R.F. Inégalités pour martingales à un et deux indices: L'éspace H^p. In: *École d'Été de Probabilités de Saint Fluor VIII-1978.* Lecture Notes in Mathematics **774**, 251–334. Springer-Verlag, New York (1980).
103. Gut, A. On the moments and limit distributions of some first passage times. *Ann. Probab.* **2**, 277–308 (1974).
104. Gut, A. Marcinkiewicz laws and convergence rates in the law of large numbers for random variables with multidimensional indices. *Ann. Probab.* **6**, 469–482 (1978).
105. Gut, A. Convergence rates for probabilities of moderate deviations for sums of random variables with multidimensional indices. *Ann. Probab.* **8**, 298–313 (1980).
106. Gut, A. An introduction to the theory of asymptotic martingales. In: Gut, A., and Schmidt, K.D. *Amarts and Set Funtion Processes.* Lecture Notes in Mathematics **284**, 4–49. Springer-Verlag (1983).
107. Gut, A. Complete convergence and convergence rates for randomly indexed partial sums with an application to some first passage times. *Acta Math. Acad. Sci. Hungar.* **42**, 225-232 (1983). Correction *ibid.* **45**, 235–236 (1985).

108. Gut, A. On the law of the iterated logarithm for randomly indexed partial sums with two applications. *Studia Sci. Math. Hungar.* **20**, 63–69 (1985).
109. Gut, A. Law of the iterated logarithm for subsequences. *Probab. Math. Statist.* **7**, 27-58 (1986).
110. Gut, A. *Stopped Random Walks.* Springer-Verlag, New York (1988).
111. Gut, A. Limit theorems for record times. In: *Proc. 5th Vilnius Conf. on Prob. Theory and Math. Stat., Vol. 1*, 490-503. Utrecht, The Netherlands (1990).
112. Gut, A. Convergence rates for record times and the associated counting process. *Stochastic process. Appl.* **36**, 135–151 (1990).
113. Gut, A. *An Intermediate Course in Probability.* Springer-Verlag, New York (1995).
114. Gut, A. Convergence rates in the central limit theorem for multidimensionally indexed random variables. *Studia Math. Hung.* **37**, 401–418 (2001).
115. Gut, A Precise asymptotics for record times and the associated counting process. *Stochastic Process. Appl.* **101**, 233–239 (2002).
116. Gut, A. The moment problem for random sums. *J. Appl. Probab.* **40**, 797–802 (2003).
117. Gut, A. An extension of the Kolmogorov-Feller weak law of large numbers with an application to the St Petersburg game. *J. Theoret. Probab.* **17**, 769–779 (2004).
118. Gut, A., and Janson, S. The limiting behaviour of certain stopped sums and some applications. *Scand. J. Statist.* **10**, 281–292 (1983).
119. Gut, A., and Janson, S. Converse results for existence of moments and uniform integrability for stopped random walks. *Ann. Probab.* **14**, 1296–1317 (1986).
120. Gut, A., and Spătaru, A. Precise asymptotics in the Baum-Katz and Davis law of large numbers. *J. Math. Anal. Appl.* **248**, 233–246 (2000).
121. Gut, A., and Spătaru, A. Precise asymptotics in the law of the iterated logarithm. *Ann. Probab.* **28**, 1870–1883 (2000).
122. Gut, A., and Steinebach, J. Convergence rates and precise asymptotics for renewal counting processes and some first passage times. In: *Proceedings of the International Conference on Asymptotic Methods in Stochastics, May 23–25, 2002, Carleton University Ottawa. Fields Institute Communications* **44**, 205–227 (2004).
123. de Haan, L. *On Regular Variation and Its Application to the Weak Convergence of Extremes.* Mathematical Centre Tracts **32**, Amsterdam (1970).
124. de Haan, L., and Hordijk, A. The rate of growth of sample maxima. *Ann. Math. Statist.* **43**, 1185–1196 (1972).
125. Hahlin, L.-O. Double records. *Report U.U.D.M.* 1995:12, Uppsala University (1995).
126. Hájek, J., and Rényi, A. Generalization of an inequality of Kolmogorov. *Acta Math. Acad. Sci. Hung.* **6**, 281–283 (1955).
127. Hall, P. Characterizing the rate of convergence in the central limit theorem. *Ann. Probab.* **8**, 1037–1048 (1980).
128. Hall, P., and Heyde, C.C. *Martingale Limit Theory and Its Applications.* Academic Press, Cambridge, MA (1980).
129. Halmos, P.R. *Measure Theory.* Van Nostrand, Princeton, NJ (1950).
130. Hartman, P., and Wintner, A. On the law of the iterated logarithm. *Amer. J. Math.* **63**, 169–176 (1941).

131. Herrndorff, N. Stationary strongly mixing sequences not satisfying the central limit theorem. *Ann. Probab.* **11**, 809–813 (1983).
132. Hewitt, E., and Savage, L.J. Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80**, 470–501 (1955).
133. Heyde, C.C. On a property of the lognormal distribution. *J. Roy. Statist. Soc. Ser. B* **29**, 392–393 (1963).
134. Heyde, C.C. On the influence of moments on the rate of convergence to the normal distribution. *Z. Wahrsch. verw. Gebiete* **8**, 12–18 (1967).
135. Heyde, C.C. On the converse to the iterated logarithm law. *J. Appl. Probab.* **5**, 210–215 (1968).
136. Heyde, C.C. A supplement to the strong law of large numbers. *J. Appl. Probab.* **12**, 173–175 (1975).
137. Hodges, J.L., and Le Cam, L. The Poisson approximation to the Poisson binomial distribution. *Ann. Math. Statist.* **31**, 737–740 (1960).
138. Hoeffding, W. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30 (1963).
139. Hoffmann-Jørgensen, J. Sums of independent Banach space valued random variables. *Studia Math.* **LII**, 159–186 (1974).
140. Hsu, P.L., and Robbins, H. Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA* **33**, 25–31 (1947).
141. Ibragimov, I.A. Some limit theorems for stationary sequences. *Theor. Prob. Appl.* **7**, 349–382 (1962).
142. Ibragimov, I.A., and Linnik, Yu.V. *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen (1971).
143. Iosifescu, M. Limit theorems for ϕ-mixing sequences. A survey. In: *Proc. 5th Conf. on Prob. Theory, Braşov, 1974*, 51-57. Editura Acad. R.S.R. Bucureşti, Romania (1977).
144. Iosifescu, M., and Theodorescu, R. *Random Processes and Learning.* Springer-Verlag, New York (1969).
145. Jacod, J., and Protter, P. *Probability Essentials*, 2nd ed. Springer-Verlag, Berlin (2003).
146. Jain, N. Tail probabilities for sums of independent Banach space valued random variables. *Z. Wahrsch. verw. Gebiete* **33**, 155–166 (1975).
147. Janson, S. Moments for first passage and last exit times, the minimum and related quantities for random walks with positive drift. *Adv. in Appl. Probab.* **18**, 865–879 (1986).
148. Janson, S. Renewal theory for m-dependent variables. *Ann. Probab.* **11**, 558–568 (1983).
149. Janson, S. Some pairwise independent sequences for which the central limit theorem fails. *Stochastics* **23**, 439–448 (1988).
150. Kahane, J.-P. *Some Random Series of Functions*, 2nd ed. Cambridge University Press, Cambridge (1985).
151. Kallenberg, O. *Foundations of Modern Probability,* 2nd ed. Springer-Verlag, New York (2001).
152. Karamata, J. Sur une mode de croissance régulière des fonctions. *Mathematica* (Cluj) **4**, 38–53 (1930).
153. Katz, M. The probability in the tail of a distribution. *Ann. Math. Statist.* **34**, 312–318 (1963).
154. Katz, M. Note on the Berry-Esseen theorem. *Ann. Math. Statist.* **34**, 1107–1108 (1963).

155. Kertész, I. *A Kudarc.* (1988). Swedish transl: *Fiasko.* Norstedts Förlag, Stockholm (2000).
156. Kesten, H. A sharper form of the Doeblin-Lévy-Kolmogorov-Rogozin inequality for concentration functions. *Math. Scand.* **25**, 133–144 (1969).
157. Khan, R.A. A probabilistic proof of Stirling's formula. *Amer. Math. Monthly* **81**, 366–369 (1974).
158. Khinthine, A. Über die diadischen Brüche. *Math. Z.* **18**, 109–116 (1923).
159. Khinthine, A. Über einen Satz der Wahrscheinlichkeitsrechnung. *Fund. Math.* **6**, 9–20 (1924).
160. Kolmogorov, A.N. Über die Summen durch den Zufall bestimmter unabhängiger Größen. *Math. Ann.* **99**, 309–319 (1928).
161. Kolmogorov, A.N. Über das Gesetz des iterierten Logarithmus. *Math. Ann.* **101**, 126–135 (1929).
162. Kolmogorov, A.N. Bemerkungen zu meiner Arbeit "Über die Summen zufälliger Größen". *Math. Ann.* **102**, 484–488 (1930).
163. Kolmogorov, A.N. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* (1933). English transl: *Foundations of the Theory of Probability.* Chelsea, New York (1956).
164. Krein, M.G. On one extrapolation problem of A.N. Kolmogorov. *Doklady Akad. Nauk SSSR* **46**, 339–342 (1944) (in Russian).
165. Lalley, S.P. Limit theorems for first-passage times in linear and nonlinear renewal theory. *Adv. in Appl. Probab.* **16**, 766–803 (1984).
166. Lamb, C.W. A short proof of the martingale convergence theorem. *Proc. Amer. Math. Soc.* **38**, 215–217 (1973).
167. Lamperti, J. Wiener's test and Markov chains. *J. Math. Anal. Appl.* **6**, 58–66 (1963).
168. Leadbetter, M.R., Lindgren, G., and Rootzén, H. *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New York (1983).
169. Le Cam, L. An approximation theorem for the Poisson binomial distribution. *Pacific. J. Math.* **10**, 1181–1197 (1960).
170. Lévy, P. *Calcul des Probabilités.* Gauthier-Villars, Paris (1925).
171. Lévy, P. Propriétés asymptotiques de sommes de variables aléatoires indépendantes ou enchaînées. *J. Math. Pures Appl.* **14**, 37–402 (1935).
172. Lévy, P. *Théorie de l'Addition des Variables Aléatoires.* Gauthier-Villars, Paris (1937). 2nd ed. (1954).
173. Li, D., Wang, X., and Rao, M.B. Some results on convergence rates for probabilities of moderate deviations for sums of random variables. *Internat. J. Math. & Math. Sci.* **15**, 481–498 (1992).
174. Lin, G.D. On the moment problems. *Statist. Probab. Lett.* **35**, 85-90 (1997). Erratum *ibid.* **50**, 205 (2000).
175. Lin G.D., and Stoyanov, J.M. On the moment determinacy of the distributions of compound geometric sums. *J. Appl. Probab.* **39**, 545–554 (2002).
176. Lindeberg, J.W. Eine neue Herleitung des Exponentialgezetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.* **15**, 211–225 (1922).
177. Loève, M. *Probability Theory*, 3rd ed. Van Nostrand, Princeton, NJ (1963).
178. Lyapounov, A.M. Sur une proposition de la théorie des probabilités. *Bull. Acad. Sci. St.-Pétersbourg* (5) **13**, 359–386 (1900).
179. Lyapounov, A.M. Nouvelle forme du théorème sur la limite de probabilités. *Mem. Acad. Sci. St.-Pétersbourg* (8) **12** No 5 (1901).

180. MARCINKIEWICZ, J., AND ZYGMUND, A. Sur les fonctions indépendantes. *Fund. Math.* **29**, 60–90 (1937).
181. MARCINKIEWICZ, J., AND ZYGMUND, A. Remarque sur la loi du logarithm itéré. *Fund. Math.* **4**, 82–105 (1937).
182. MARCINKIEWICZ, J., AND ZYGMUND, A. Quelque théorèmes sur les fonctions indépendantes. *Studia Math.* **VII**, 104–120 (1938).
183. MCCABE, B.J., AND SHEPP, L.A. On the supremum of S_n/n. *Ann. Math. Statist.* **41**, 2166–2168 (1970).
184. MEYER, P.A. *Martingales and Stochastic Integrals I.* Lecture Notes in Mathematics **284**, Springer-Verlag (1972).
185. VON MISES, R. La distribution de la plus grande de n valeurs. *Rev. Math. Interbalkanique* **1**, 1–20 (1936). *Selected papers* **II**, 271–294. Amer. Math. Soc., Providence, R.I. (1964)
186. NEVEU, J. *Discrete-parameter Martingales.* North-Holland, Amsterdam (1975).
187. NEVZOROV, V.B. *Records: Mathematical Theory.* Translation of Mathematical Monographs. Amer. Math. Soc. **194**, Providence, RI. (2001).
188. PARTHASARATHY, K.R. *Probability Measures on Metric Spaces.* Academic Press, Cambridge, MA (1967).
189. PELIGRAD, M. Invariance principles for mixing sequences of random variables. *Ann. Probab.* **10**, 968–981 (1982).
190. PELIGRAD, M. Recent advances in the central limit theorem and its invariance principle for mixing sequences of random variables (a survey). In: E. Eberlein and M.S. Taqqu, eds., *Dependence in Probability and Statistics,* 193–223. Birkhäuser, Boston, MA (1986).
191. PELIGRAD, M. On Ibragimov-Iosifescu conjecture for ϕ-mixing sequences. *Stoch. Proc. Appl.* **35**, 293–308 (1990).
192. PELIGRAD, M. On the asymptotic normality of sequences of weak dependent random variables. *J. Theoret. Probab.* **9**, 703–715 (1996).
193. PETROV, V.V. An estimate of the deviation of the distribution of a sum of independent random variables from the normal law. *Dokl. Akad. Nauk* **160**, 1013–1015 (1965). (In Russian). Engl. transl.: *Sov. Math. Dokl.* **6**, 242–244.
194. PETROV, V.V. *Sums of Independent Random Variables.* Springer-Verlag, New York (1975).
195. PETROV, V.V. *Limit Theorems of Probability Theory.* Oxford Science Publications, Clarendon Press, Oxford (1995).
196. PETROV, V.V. A note on the Borel-Cantelli lemma. *Statist. Probab. Lett.* **58**, 283-286 (2002).
197. PETROV, V.V. A generalization of the Borel-Cantelli lemma. *Statist. Probab. Lett.* **67**, 233-239 (2004).
198. PHILLIP, W. Invariance principles for independent and weakly dependent random variables. In: E. Eberlein and M.S. Taqqu, eds., *Dependence in Probability and Statistics,* 225–268. Birkhäuser, Boston, MA (1986).
199. PHILLIP, W., AND STOUT, W.F. Almost sure invariance principles for sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* **161** (1975).
200. POLLARD, D. *Convergence of Stochastic Processes.* Springer-Verlag, New York (1984).
201. PRABHU, N.U. *Stochastic Processes.* Macmillan, New York (1965).
202. PRATT, J.W. On interchanging limits and integrals. *Ann. Math. Statist.* **31**, 74–77 (1960).

203. Pyke, R., and Root D. On convergence in r-mean for normalized partial sums. *Ann. Math. Statist.* **39**, 379–381 (1968).
204. Rényi, A. On the asymptotic distribution of the sum of a random number of independent random variables. *Acta Math. Acad. Sci. Hungar.* **8**, 193–199 (1957).
205. Rényi, A. On the extreme elements of observations. *MTA III, Oszt. Közl.* **12**, 105–121 (1962). *Collected Works* **III**, 50–66. Akadémiai Kiadó, Budapest (1976).
206. Resnick, S.I. Limit laws for record values. *Stochastic Process. Appl.* **1**, 67–82 (1973).
207. Resnick, S.I. *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag (1987).
208. Resnick, S.I. *A Probability Path.* Birkhäuser, Boston, MA (1999).
209. Richter, W. Limit theorems for sequences of random variables with sequences of random indices. *Theor. Prob. Appl.* **X**, 74–84 (1965).
210. Rootzén, H. On the functional central limit theorem for martingales. *Z. Wahrsch. verw. Gebiete* **38**, 199–210 (1977).
211. Rootzén, H. On the functional central limit theorem for martingales, II. *Z. Wahrsch. verw. Gebiete* **51**, 79–93 (1980).
212. Rootzén, H. Central limit theory for martingales via random change of time. In: A. Gut and L. Holst, eds., *Essays in Honour of Carl Gustav Esseen*, 154–190. Uppsala University (1983).
213. Rosén, B. On the asymptotic distribution of sums of independent identically distributed random variables. *Arkiv Matematik* **4**, 323–332 (1962).
214. Rosenthal, H.P. On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8**, 273–303 (1970).
215. Russo, R.R. On the last exit time and the number of exits of partial sums over a moving boundary. *Statist. Probab. Lett.* **6**, 371–377 (1988).
216. Samorodnitsky, G., and Taqqu, M.S. *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York (1994).
217. Seneta, E. *Regularly Varying Functions.* Lecture Notes in Mathematics **508**, Springer-Verlag, Berlin (1976).
218. Serfling, R.J. A general Poisson approximation theorem. *Ann. Probab.* **3**, 726–731 (1975).
219. Serfling, R.J. *Approximation Theorems of Mathematical Statistics.* Wiley, New York (1980).
220. Shiganov, I.S. Refinement of the upper bound of the constant in the central limit theorem. *J. Sov. Math.* **35**, 2545–2550 (1986).
221. Shorack, G.A., and Wellner, J.A. *Empirical Processes with Applications to Statistics.* Wiley, New York (1986).
222. Shohat, J.A., and Tamarkin, J.D. *The Problem of Moments.* Amer. Math. Soc., Providence, R.I. (1943).
223. Siegmund, D. *Sequential Analysis. Tests and Confidence Intervals.* Springer-Verlag, New York (1985).
224. Skorohod, A.V. Limit theorems for stochastic processes. *Theor. Prob. Appl.* **I**, 261–290 (1956).
225. Słaby, M. The law of the iterated logarithm for subsequences and characterizations of the cluster set of $S_{n_k}/(2n_k \log\log n_k)^{1/2}$ in Banach spaces. *J. Theoret. Probab.* **2**, 343–376 (1989).

226. Slivka, J. On the law of the iterated logarithm. *Proc. Nat. Acad. Sci. USA* **63**, 289–291 (1969).
227. Slivka, J., and Severo, N.C. On the strong law of large numbers. *Proc. Amer. Math. Soc.* **24**, 729–734 (1970).
228. Slud, E.V. The moment problem for polynomial forms in normal random variables. *Ann. Probab.* **21**, 2200–2214 (1993).
229. Smith, W.L. Asymptotic renewal theorems. *Proc. Roy. Soc. Edinburgh Sect A.* **64**, 9–48 (1954).
230. Smith, W.L. Renewal theory and its ramifications. *J. Roy. Stat. Soc. Ser. B* **20**, 243–302 (1958).
231. Spătaru, A. Precise asymptotics in Spitzer's law of large numbers. *J. Theoret. Probab.* **12**, 811–819 (1999).
232. Spitzer, F. A combinatorial lemma and its applications to probability theory. *Trans. Amer. Math. Soc.* **82**, 323–339 (1956).
233. Spitzer, F. A Tauberian theorem and its probability interpretation. *Trans. Amer. Math. Soc.* **94**, 150–169 (1960).
234. Spitzer, F. *Principles of Random Walk,* 2nd ed. Springer-Verlag, New York (1976).
235. Steele, J.M. *Stochastic Calculus and Financial Applications.* Springer-Verlag, New York (2000).
236. Steiger, W.L., and Zaremba, S.K. The converse of the Hartman-Wintner theorem. *Z. Wahrsch. verw. Gebiete* **22**, 193–194 (1972).
237. Stein, C. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In: *Proc. Sixth Berkeley Symp. Math. Statist. Prob. 1970* **2**, 583–603 (1972).
238. Strassen, V. A converse to the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **4**, 265–268 (1966).
239. Stout, W.F. *Almost Sure Convergence.* Academic Press, Cambridge, MA (1974).
240. Sudderth, W.D. A "Fatou equation" for randomly stopped variables. *Ann. Math. Statist.* **42**, 2143–2146 (1971).
241. Takács, L. On a probability problem arising in the theory of counters. *Proc. Cambridge Philos. Soc.* **52**, 488–498 (1956).
242. Thiele, T.N. *Almindelig Iakttagelselære: Sandsynlighetsregning og Mindste Kvadraters Methode.* C.A. Reitzel, Copenhagen (1889).
243. Thiele, T.N. *Theory of Observations.* Layton, London (1903). Reprinted in *Ann. Math. Statist.* **2**, 165–307 (1931).
244. Thomas, D. Hidden messages and the Bible code. *Sceptical Inquirer.* November/December 1997.
245. Torrång, I. Law of the iterated logarithm – cluster points of deterministic and random subsequences. *Probab. Math. Statist.* **8**, 133–141 (1987).
246. Trotter, H.F. An elementary proof of the central limit theorem. *Arch. Math.* **10**, 226–234 (1959).
247. Utev, S., and Peligrad, M. Maximal inequalities and an invariance principle for a class of weakly dependent random variables. *J. Theoret. Probab.* **16**, 101–115 (2003).
248. van der Vaart, A.W., and Wellner, J.A. *Weak Convergence and Empirical Processes.* Springer-Verlag, New York (1996).
249. Ville, J. *Étude Critique de la Notion de Collectif.* Monographie des Probabilités. Gauthier-Villars, Paris (1939).

250. WALD, A. *Sequential Analysis.* Wiley, New York (1947).
251. WILLIAMS, D. On Rényi's "record" problem and Engel's series. *Bull. London Math. Soc.* **5**, 235–237 (1973).
252. WILLIAMS, D. *Probability with Martingales.* Cambridge University Press, Cambridge (1991).
253. WITTENBERG, H. Limiting distributions of random sums of independent random variables. *Z. Wahrsch. verw. Gebiete* **3**, 7–18 (1964).
254. WITZTUM, D., RIPS, E., AND ROSENBERG, Y. Equidistant letter sequences in the book of Genesis. *Statistical Science* **9**, 429–438 (1994).

Index