

# 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,*

$$e^x \leq 1 + x + x^2 \quad \text{for } |x| \leq 1, \quad (\text{A.1})$$

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

$$|e^z - 1| \leq |z|e^{|z|} \quad \text{for } z \in \mathbb{C}, \quad (\text{A.3})$$

$$|e^z - 1 - z| \leq |z|^2 \quad \text{for } z \in \mathbb{C}, \quad |z| \leq 1/2, \quad (\text{A.4})$$

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

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

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

For the other half of (A.1),

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

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

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

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

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

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

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

$$\begin{aligned} |e^z - 1 - z| &\leq \sum_{k=2}^{\infty} \frac{|z|^k}{k!} \leq \frac{|z|^2}{2} \sum_{k=2}^{\infty} |z|^{k-2} = \frac{|z|^2}{2} \cdot \frac{1}{1-|z|} \leq |z|^2, \\ |\log(1-z) + z| &\leq \sum_{k=2}^{\infty} \frac{|z|^k}{k} \leq \frac{|z|^2}{2} \sum_{k=2}^{\infty} |z|^{k-2} \leq |z|^2. \quad \square \end{aligned}$$

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

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

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

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

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

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

$$\int_0^y e^{ix} dx = \begin{cases} y + i \int_0^y e^{ix} (y-x) dx, \\ \frac{e^{iy} - 1}{i}, \end{cases}$$

so that, by equating these expressions, we obtain

$$e^{iy} = 1 + iy + i^2 \int_0^y e^{ix} (y-x) dx. \quad (\text{A.7})$$

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

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

and, hence,

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

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

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

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

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

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

Another estimate concerns the integral  $\int_0^t \frac{\sin x}{x} dx$  as  $t \rightarrow \infty$ . 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  $n \rightarrow \infty$ , which proves that the limit as  $t \rightarrow \infty$  exists.

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

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

*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  $\sin x \leq x$ .

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

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

we may apply Fubini's theorem to obtain

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

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

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

so that

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

Inserting this into the double integral yields

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

so that, finally,

$$\int_0^\infty \frac{\sin x}{x} dx = \lim_{t \rightarrow \infty} \int_0^t \frac{\sin x}{x} dx = \frac{\pi}{2},$$

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

## 2 Mill's Ratio

The function  $e^{-x^2/2}$ , 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*

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

*In particular,*

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

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

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

Rearranging this proves the right-most inequality. Similarly,

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

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:

$$1 - \Phi(x) = \int_x^\infty \frac{y}{y} \phi(y) dy < \frac{1}{x} \int_x^\infty y \phi(y) dy = \frac{\phi(x)}{x}. \quad \square$$

### 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 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,

$$\lim_{n \rightarrow \infty} n^\alpha \sum_{k=n}^{\infty} \frac{1}{k^{\alpha+1}} = \frac{1}{\alpha} \quad \text{as } n \rightarrow \infty.$$

(ii) For  $\beta > 0$ ,

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

and

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

(iii)

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

and

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

*Proof.* (i): We have

$$\frac{1}{\alpha n^\alpha} = \int_n^\infty \frac{dx}{x^{\alpha+1}} = \sum_{k=n}^{\infty} \int_k^{k+1} \frac{dx}{x^{\alpha+1}} \left\{ \begin{array}{l} \leq \sum_{k=n}^{\infty} \frac{1}{k^{\alpha+1}}, \\ \geq \sum_{k=n}^{\infty} \frac{1}{(k+1)^{\alpha+1}}. \end{array} \right.$$

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

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

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

The point of departure for (iii) is

$$\log n = \int_1^n \frac{dx}{x}. \quad \square$$

**Exercise 3.1.** Finish the proof of the lemma. □

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

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

where  $\gamma = 0.5772\dots$  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

$$\prod(\dots) = \exp \left\{ \sum \log(\dots) \right\}.$$

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

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

*Convergence thus holds iff*

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

*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  $a_n \rightarrow 0$  as  $n \rightarrow \infty$ , 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

$$-\frac{3}{2}a_n \leq \log(1 - a_n) \leq -a_n. \quad \square$$

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

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

*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  $(1 - a_n)^n \rightarrow 1$  as  $n \rightarrow \infty$ . Recalling (A.2),

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

which establishes the preliminary fact that

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

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  $|x+y| \leq |x|+|y|$ . Following are some analogs for powers.

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

$$(x+y)^r \leq \begin{cases} 2^r(x^r + y^r), & \text{for } r > 0, \\ x^r + y^r, & \text{for } 0 < r \leq 1, \\ 2^{r-1}(x^r + y^r), & \text{for } r \geq 1. \end{cases}$$

*Proof.* For  $r > 0$ ,

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

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

$$\left(\frac{x^r}{x^r + y^r}\right)^{1/r} + \left(\frac{y^r}{x^r + y^r}\right)^{1/r} \leq \frac{x^r}{x^r + y^r} + \frac{y^r}{x^r + y^r} = 1,$$

and, hence, that

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

which is the same as the second inequality.

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

$$\left(\frac{x+y}{2}\right)^r \leq \frac{1}{2}x^r + \frac{1}{2}y^r,$$

which is easily reshuffled into the third inequality.  $\square$

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

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad \text{for } x, y > 0.$$

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

$$\begin{aligned} xy &= \exp\{\log xy\} = \exp\left\{\frac{1}{p} \log x^p + \frac{1}{q} \log x^q\right\} \\ &\leq \exp\left\{\log\left(\frac{x^p}{p} + \frac{y^q}{q}\right)\right\} = \frac{x^p}{p} + \frac{y^q}{q}. \end{aligned} \quad \square$$

*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  $(x - y)^2 \geq 0$ .  $\square$

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

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

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*

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

*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

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

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,

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

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

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

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

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

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

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

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  $g$ .  $\square$

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

## 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  $a_n \rightarrow a$  as  $n \rightarrow \infty$ , then*

$$\frac{1}{n} \sum_{k=1}^n a_k \rightarrow a \quad \text{as } n \rightarrow \infty.$$

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

$$\frac{1}{B_n} \sum_{k=1}^n w_k a_k \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*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  $n > n_0$ ,

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

so that

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

which does it, since  $\varepsilon$  can be made arbitrarily small.

This proves the first statement. The second one follows similarly.  $\square$

**Exercise 6.1.** Carry out the proof of the second half of the lemma.  $\square$

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

$$\begin{aligned} \frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} a_k &\rightarrow a, \\ \frac{1}{\log \log n} \sum_{k=1}^n \frac{1}{k \log k} a_k &\rightarrow a, \\ n^\alpha \sum_{k=1}^n \frac{1}{k^{\alpha+1}} a_k &\rightarrow \frac{a}{\alpha}, \quad \alpha > 0. \end{aligned} \quad \square$$

Next, an important further development in this context.

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

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

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

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

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

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,  $b_n \rightarrow b_\infty$  as  $n \rightarrow \infty$  by assumption, and

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

by the second half of Lemma 6.1, since we are faced with a weighted average of quantities tending to  $b_\infty$ .  $\square$

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

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

The following continuous version is proved similarly.

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

$$\limsup_{n \rightarrow \infty} \left| \frac{1}{2h} \int_{|x-b|<h} g_n(x) dx - g(b) \right| < \varepsilon \quad \text{for all } h \in (0, h_0).$$

**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  $\rho$ ,  $-\infty < \rho < \infty$ , denoted  $u \in \mathcal{RV}(\rho)$ , iff*

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

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

Typical examples of regularly varying functions are

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

Typical slowly varying functions are the above when  $\rho = 0$ . Moreover, every positive function with a finite limit as  $x \rightarrow \infty$  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

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

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  $u_2(x) \rightarrow \infty$  as  $x \rightarrow \infty$ , 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  $\frac{u(tx)}{u(t)} \rightarrow x^\rho$  as  $t \rightarrow \infty$  yields

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

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

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

If  $\rho_1 < \rho_2$  the limit equals  $x^{\rho_2}$ , and if the exponents are equal (to  $\rho$ ) the limit becomes  $x^\rho$ .

- (e): An application of Lemma 7.2 yields

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

*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*

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

Then

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

*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  $n \rightarrow \infty$ .

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

Suppose that  $u \in \mathcal{RV}(\rho)$ , where  $\rho > -1$ . Then, since the slowly varying component is “negligible” with respect to  $x^\rho$ , 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  $U(x) = \int_a^x u(y) dy \in \mathcal{RV}(\rho + 1)$ .*

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

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

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

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

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

and that

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

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  $\varepsilon$  and  $\delta$  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,

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

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  $g(x) = e^{cx}$  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

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

so that, by Lemma 8.1,  $g(e^x) = cx$ , which is the same as  $g(x) = c \log x$ .

(b): In this case a change of variables yields

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

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

(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,  $g(x) = e^{cx}$ . □

## 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  $C$  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  $n \geq 1$ ,

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

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  $x \rightarrow \pm\infty$  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,

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

Combining this with the previous inequality proves that

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

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  $\varepsilon$ , 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

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

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

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

Moreover,

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

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

$$\begin{aligned} G_n(x) &\rightarrow G(x) \quad \text{as } n \rightarrow \infty, \quad \text{for all } x \in D, \\ G_n(\pm\infty) &\rightarrow G(\pm\infty) \quad \text{as } n \rightarrow \infty, \end{aligned}$$

then

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

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

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

then

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

*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  $\varepsilon$  and the accompanying  $\delta$  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

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

so that

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

so that

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

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

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

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

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

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

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

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

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  $\{n_k, k \geq 1\}$ ,  $n_k \nearrow \infty$  as  $k \rightarrow \infty$ , such that

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

The first observation is that we cannot have  $x_k \rightarrow \pm\infty$ , because of the second assumption, which means that  $\{x_k, k \geq 1\}$  is bounded, which implies that there exists a convergent subsequence,  $x_{k_j} \rightarrow x$ , say, as  $j \rightarrow \infty$ . 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  $j \rightarrow \infty$ :

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

- $x_{k_j} \searrow x$ , 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

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

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,

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

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

In the third case,

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

after which we let  $r_1$  approach  $x-$ .

Finally,

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

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

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

and

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

Then, for  $i = 1, 2$ ,

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

as soon as  $n$  is large enough. In addition,  $f$  is sandwiched between the  $g$ -functions;  $g_1(x) \leq f(x) \leq g_2(x)$  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  $|x| \leq b$  (that is,  $a = -b$ ) and equal to 0 otherwise. By a slight modification the sandwiching effect can also be retained.  $\square$

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

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

*Proof.* Set, for  $n \geq 1$ ,

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

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

**Exercise 9.1.** Please pursue the checking.  $\square$

---

## 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  $r$ th 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. CSÖRGÖ, M., AND RYCHLIK, Z. Weak convergence of sequences of random elements with random indices. *Math. Proc. Camb. Phil. Soc.* **88**, 171–174 (1980).
58. CSÖRGÖ, 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'espace  $H^p$ . 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 Function 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 SPĂȚARU, 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 SPĂȚARU, 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. HÄHLIN, 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  $\phi$ -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 Exponentialgesetzes 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 logarithme itéré. *Fund. Math.* **4**, 82–105 (1937).
182. MARCINKIEWICZ, J., AND ZYGMUND, A. Quelques 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  $\phi$ -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, Oszk. 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. SLABY, 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. SPĂȚARU, 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

- $\aleph_0$ , 3
- $\aleph_1$ , 3
- $\alpha$ -mixing, 450
- $\alpha$ -quantile, 140
- $\lambda$ -system, 6
- $\phi$ -mixing, 450
- $\pi$ -system, 6, 10
- $\psi$ -mixing, 450
- $\rho$ -mixing, 450
- $\sigma$ -algebra, 6
  - approximation, 14
  - filtration, 477, 541
    - natural, 477
  - generator, 7, 10
  - intersection, 6
  - minimal, 7
  - of sets, 10
  - pre- $\tau$ - $\sigma$ -algebra, 493
  - tail- $\sigma$ -field, 20, 72, 520
  - union, 6, 22, 477
- absolute central moments, 62
- absolute continuity, 470
- absolute moments, 62
- absolutely continuous distribution, *see*
  - distribution, absolutely continuous
- algebra, 6, 10
- almost sure convergence, 48, *see*
  - convergence
- alternating renewal process, 379
- amart, 514
- Anscombe's theorem
  - central limit theorem, 346, 348
  - law of the iterated logarithm, 416
  - stable distributions, 440
- antrax cells, 87
- asymptotic martingales, 514
- availability, 379
- Bayes' formula, 18
- Bernoulli random walk, 88
- Berry-Esseen
  - Hartman-Wintner, 412
  - law of the iterated logarithm, 412
- Berry-Esseen theorem, 355, 356
  - counting process
    - records, 364
    - renewal theory, 364
- Bible code, 100
- Bonferroni inequalities, 23
- Borel
  - normal numbers, 305, 383
  - sets, 15
    - generator, 15
    - on  $\mathbb{R}^n$ , 16
  - space, 15
- Borel-Cantelli lemmas, 96–113
  - first, 96
  - pair-wise independence, 104
  - second, 97
  - subsequence, 98
- branching process, 86
  - martingales, 483
- Brownian motion, 375
- Burkholder inequalities, 506, 507
- $C$ , 570
- $C_B$ , 70, 203, 226, 335

- $C[0, 1]$ , 46, 203, 374, 375  
 Cantelli's inequality, 154  
 Cantor  
   distribution, 40, 73  
   mean, 73  
   variance, 73  
   set, 41  
 Cantor distribution  
   characteristic function, 159, 167  
 cardinal number, 3  
 Cauchy convergence  
   almost sure, 256  
   in  $r$ -mean, 257  
   in distribution, 258  
   in probability, 257  
 Cauchy's functional equation, 568  
 Cauchy-Schwarz inequality, 130  
 central limit theorem, 329–381  
   Anscombe, 346  
   generalized, 348  
   Berry-Esseen, 355, 356  
   convergence rates, 366  
   domain of attraction, 439  
   Edgeworth expansions, 363  
   empirical correlation coefficient, 380  
   Esseen's lemma, 357  
   first passage times, 351  
   functional, 376, 451  
   i.i.d. case, 330  
   Lindberg conditions, 330  
   Lindeberg-Lévy-Feller, 331  
   Lyapounov's condition, 339, 378  
   martingales, 515  
   moment convergence, 353  
   non-uniform remainder term  
     estimates, 363  
   pair-wise independence, 344  
   random indices, 346  
   records, 351  
   remainder term estimates, 355, 356  
   sample variance, 380  
   sums of dependent random variables,  
     450  
   triangular array, 345  
   triangular arrays, 377  
   uniformly asymptotically negligible,  
     340  
 central moments, 62  
 change of variables formula, 60  
 characteristic function, 157–184  
   absolutely continuous distributions,  
     159  
   inversion, 161, 252  
 Cantor distribution, 159, 167  
 continuity theorem, 238, 239, 251  
 degenerate distribution, 164  
 discrete distributions, 159  
   inversion, 163  
 infinite divisibility, 444  
 inversion, 160, 161, 163, 252  
 lattice distribution, 163, 164  
 moments, 176–178  
 multiplication theorem, 164  
 multivariate normal, 182  
 Pólya's theorem, 262  
 random vector, 180, 246  
 stable distribution, 171  
 stable distributions, 426, 427  
 sums of a random number of random  
   variables, 193  
 surprise, 173, 174  
 uniqueness, 160, 173, 250  
 Chebyshev's inequality, 121  
   truncated, 121  
 Clarkson's inequality, 138, 563  
 closed set, 15  
 closing random variable, 517, 518, 520,  
   544  
 cluster set, 416, 417  
   definition, 405  
   law of the iterated logarithm, 405,  
     411  
   subsequences, 405, 406, 411  
 complement, 3  
 complete convergence, 312  
   definition, 203  
   law of large numbers, 312, 315  
   partial maxima, 312  
   random indices, 322, 323  
 completeness, 12  
 concentration function, 155  
 conditional  
   density, 80  
   distribution function, 79  
   probability, 17  
   probability function, 79  
 conditional expectation  
   defining relation, 468

- definition, 468
- equivalence class, 470
- Fatou's lemma, 473
- monotone convergence, 473
- properties, 471
- smoothing, 474
- conditional inequalities, 476, 477
  - $c_r$ , 476
  - Hölder, 476
  - Jensen, 477
  - Lyapounov, 477
  - Minkowski, 477
- conditional probability, 468
- conditional square function, 506
- conditional variance, 475
- confidence interval, 249
- conjugate exponents, 132
- consistency, 51
- continuity point, 202, 203
- continuity set, 30, 202, 204
- continuity theorem, 238–243
  - characteristic function, 238, 239
  - cumulant generating function, 240
  - generating function, 241
  - moment generating function, 242
- continuous distribution, 36
- continuous mapping theorem, 246, 260, 264
- continuous singular distribution, 37, 40
- continuum hypothesis, 3
- convergence, 201–264
  - almost sure, 48, 218, 244, 292
  - definition, 202
  - Fatou's lemma, 218
  - martingales, 508, 517, 520
  - random indices, 302, 303
  - reversed martingales, 543, 544
  - reversed submartingales, 544
  - submartingales, 508, 518
  - upcrossings, 512
- Cauchy, *see* Cauchy convergence
- complete, *see* complete convergence
- conditional
  - dominated, 473
  - Fatou's lemma, 473
  - monotone, 473
- continuity theorem, *see* continuity theorem
- continuous mapping theorem, *see* continuous mapping theorem
- Cramér's theorem, 249
- definition, 202
  - almost sure (a.s.), 202
  - in  $r$ -mean, 202
  - in distribution, 202, 203
  - in probability, 202
  - in square mean, 202
  - in total variation, 227
  - vague, 231
- dependence, 450, 462
- dominated, 57
- equivalence, 266
  - distributional, 268
- expectation, 54
- Fatou's lemma, 56
- functions of random variables, 243–246
- in  $L^1$ 
  - martingales, 517, 520
  - reversed martingales, 543
  - reversed submartingales, 544
  - submartingales, 518
- in  $L^p$ 
  - martingales, 517
  - reversed martingales, 544
- in  $r$ -mean
  - definition, 202
  - in distribution, 222–224, 227, 232, 235, 246, 292
  - continuity point, 202, 203
  - continuity set, 202, 204
  - definition, 202, 203, 225
  - Fatou's lemma, 223
  - random indices, 345
  - uniqueness, 226
- in probability, 221, 244, 245, 292
  - definition, 202
  - Fatou's lemma, 220
  - random indices, 303
- in square mean
  - definition, 202
- in total variation, 227, 461
  - definition, 227
- infinite divisibility, 446, 447
- Kolmogorov criterion, 286
- martingales, 508–515, 517
  - Garsia's proof, 508–511

- reversed, 543
  - upcrossings proof, 511–513
- measurability, 205
- monotone, 55, 56, 213
- of moments, 218–225, 309–311, 353
- of series, 284–294
  - bounded random variables, 287
  - in  $L^2$ , 285
  - three-series theorem, 289
- Pratt's lemma, 221
- pseudo-distribution
  - function, 233, 234
- random indices, 302, 346, 348, 440
- relations
  - between concepts, 209–213, 227, 232, 235, 292, 377
  - converses, 212, 237, 253, 255
  - reversed martingales, 544
  - reversed submartingales, 544
  - Scheffé's lemma, 227, 228
- Skorohod's representation theorem, 258
- stable distributions
  - first passage times, 441
- Stein-Chen method, 459–464
- submartingales, 518
  - Garsia's proof, 508–511
  - upcrossings proof, 511–513
- subsequences, 212
- sums of sequences, 247, 248
  - independence, 247
- tight, 236
- uniform integrability, *see* uniform integrability
- uniqueness, 208
- vague, 231–237
  - definition, 231
- weak, 203
- convergence rates
  - central limit theorem, 366
  - law of large numbers, 312, 316, 321
  - last exit times, 321
  - law of the iterated logarithm, 418
  - partial maxima, 312
  - random indices, 322, 323
- convergence to types theorem, 428, 431, 452, 453
- convexity, 132, 487, 561–564
  - definition, 561
- convolution formula, 67
- convolution table, 113, 170
- correlation, 130
  - independence, 130
  - multivariate normal, 182, 183
- countable additivity, 11
- counting process
  - records, 93, 101, 308, 351, 364
  - precise asymptotics, 374
- renewal, 89, 90, 364
- counting variable, 418
  - law of large numbers, 315
  - moments, 321
  - law of the iterated logarithm
  - moments, 418
- coupling, 460, 464
- coupon collector's problem, 115
- covariance, 130
  - matrix, 181
- $c_r$ -inequality, 127
  - conditional, 476
- Cramér's theorem, 249
- Cramér-Wold device, 246
- cumulant generating function, 184
  - continuity theorem, 240
  - multiplication theorem, 185
  - uniqueness, 185
- cumulants, 185
  - moments, 185
- $D$ , 30, 46, 571
- $D^+$ , 30, 570–574
- $D[0, 1]$ , 46
- de Morgan formulas, 4
- defining relation
  - conditional expectation, 468
  - martingales, 479
  - submartingales, 479
  - supermartingales, 479
- degenerate
  - characteristic function, 164
  - distribution, 181
  - distribution function, 30
  - random variable, 27, 30
- Delta method, 349
- dense, 31, 233, 511
  - sets
    - definition, 570
    - properties, 570–574

- density, 36
  - conditional, 80
  - joint, 43
  - multivariate normal, 182
- dependence
  - convergence, 450
  - Poisson approximation, 462
- dependence concepts, 448–451
  - $\alpha$ -mixing, 450
  - $\phi$ -mixing, 450
  - $\psi$ -mixing, 450
  - $\rho$ -mixing, 450
  - $(m + 1)$ -block factor, 449
  - $m$ -dependence, 448
  - measures of dependence, 449
  - mixing, 450
- determining class, 16
- difference of sets, 3
- discrete distribution, *see* distribution, discrete
- disjoint sets, 4, 10
- distribution, 30
  - absolutely continuous, 36, 40
    - characteristic function, *see* characteristic function, absolutely continuous distributions
    - mean, *see* mean, absolutely continuous distributions
    - variance, *see* variance, absolutely continuous distributions
  - Cantor, 40, 73
  - continuous, 36
  - continuous singular, 37, 40
  - decomposition, 36
  - degenerate, 181
  - discrete, 36, 39
    - characteristic function, *see* characteristic function, discrete distributions
    - mean, *see* mean, discrete distributions
    - variance, *see* variance, discrete distributions
  - empirical, *see* empirical distribution
  - extremal, 451, 452
  - function
    - sub-probability, 31
  - lattice, *see* lattice distribution
  - marginal, 44, 181
    - max-stable, 451, 452
    - mean, *see* mean
    - perverse, 42, 63
    - pseudo-, 231
    - singular, 36
    - stable, *see* stable distribution
    - sub-probability, 231
    - variance, *see* variance
    - with random parameters, *see* random parameter
- distribution function, 30
  - conditional, 79
  - degenerate, 30
  - joint, 43
  - marginal, 44
  - non-degenerate, 30
- distributional equality, 27
- domain of attraction, 430–441
  - central limit theorem, 439
  - definition, 431
  - extremes, 456, 457
    - definition, 456
  - moments, 436
  - normal distribution, 432, 439
  - regular variation, 432
  - slow variation, 432
  - stable distributions, 431, 432
- dominated convergence, 57
  - conditional, 473
- Donsker's theorem, 375, 451
- Doob decomposition, 489, 551
- Doob's maximal inequality, 505, 543
- Doob's optional sampling theorem, 495, 497
- double or nothing, 482, 519
- Dynkin system, 6
  - generator, 7
- Dynkin's  $\pi$ - $\lambda$  theorem, 10
- Edgeworth expansions, 363
- elementary event, 10
- elementary random variable, 25
- empirical
  - correlation coefficient, 324, 380
  - distribution, 276, 306
- empty set, 4, 11
- entropy, 379
- equality
  - distributional, 27

- point-wise, 27
- equivalence
  - class
    - conditional expectations, 470
    - random variables, 26, 468, 470
  - convergence, 266
    - distributional, 268
  - moments/tail probabilities, 75
- Esscher transform, 541
- Esseen's lemma, 357
- estimator
  - unbiased, 116, 475, 476
- Etemadi's inequality, 144
- Euler's constant, 560
- events, 2, 10
  - elementary, 10
  - independence, 16
  - independence of collections, 18
  - pair-wise independent, 19
  - remote, 21
- excess, 186
- expectation, 32, 46
  - basic properties
    - general case, 53
    - non-negative random variables, 50
    - simple random variables, 48
  - conditional, *see* conditional expectation
  - convergence, 54
  - definition
    - general case, 48
    - non-negative random variable, 47
    - simple random variable, 47
  - indefinite
    - general case, 59
    - non-negative random variables, 58
  - martingale differences, 481
  - martingales, 480
  - submartingale differences, 481
  - submartingales, 480
  - supermartingale differences, 481
  - supermartingales, 480
- exponential
  - martingales, 483, 525, 527
- exponential bounds
  - lower, 386
  - upper, 386
- exponential tilting, 540
- extended random variable, 26
- extremal distribution, 452
  - definition, 451
  - Gnedenko, 452
  - record values, 457
  - types, 452
- extremes, 93, 107–109, 451–458
  - domain of attraction, 456, 457
    - definition, 456
  - extremal distribution, *see* extremal distribution
- Fatou's lemma, 56, 218, 220, 223, 473
  - conditional, 473
  - stopping times, 513
- filtration, 477, 541
  - natural, 477
- finite additivity, 11
- first entrance times, 492
- first passage times, 90, 535–541
  - central limit theorem, 351
  - complete convergence, 323
  - convergence
    - stable distributions, 441
  - convergence rates, 323
  - expectation, 536
  - law of the iterated logarithm, 417
  - local limit theorem, 365
  - martingales, 499
  - moments, 537
  - precise asymptotics, 373
  - regularity, 536
  - strong law, 306
- Fourier
  - series, 158
  - transform, 158
- Fréchet distribution, 452
- Fubini's theorem, 65
- functional central limit theorem, 376, 451
- functions of random variables, 28
- Galton-Watson process, 86
  - martingales, 483, 519
- gambler's ruin problem, 526, 541
  - expected duration, 528
- Garsia
  - (sub)martingale convergence, 508
  - reversed martingale convergence, 543
- generating function, 186

- continuity theorem, 241
- moments, 188
- multiplication theorem, 187
- random vector, 188
- sums of a random number of random variables, 194
- uniqueness, 186
- generator, 7, 31
- $\sigma$ -algebra, 7
  - Borel sets, 15
  - Dynkin system, 7
  - monotone class, 7
- Glivenko-Cantelli theorem, 306
- Gumbel distribution, 452
  
- Hájek-Rényi's inequality, 125
- Hamel equation, 425, 570
- harmonic series
  - Euler's constant, 560
  - random, 288
- Hartman-Wintner
  - Berry-Esseen, 412
  - law of the iterated logarithm, 384, 387
- Hausdorff dimension, 74
- Helly's selection principle, 232
- Hoeffding's inequality, 120
- Hölder's inequality, 129
  - conditional, 476
- Hsu-Robbins-Erdős strong law, 312
  
- Ibragimov conjecture, 450
- Ibragimov-Iosifescu conjecture, 451
- inclusion-exclusion formula, 24
- increasing process, 477
- independence
  - $\sigma$ -algebras, 71
  - collections of events, 18
  - correlation, 130
  - events, 16
  - functions of random variables, 71
  - pair-wise
    - Borel-Cantelli lemmas, 104
    - Chebyshev's inequality, 122
    - events, 19
    - random variables, 71
    - zero-one law, 105
  - random variables, 68, 70
- induced probability measure, 26
- induced probability space, 26
- inequalities, 119–155
  - Burkholder, 506, 507
  - Cantelli, 154
  - Cauchy-Schwarz, 130
  - Chebyshev, 121
    - truncated, 121
  - Clarkson, 138, 563
  - conditional, *see* conditional inequalities
  - $c_r$ , 127
  - Doob's maximal, 505, 543
  - Etemadi, 144
  - Hájek-Rényi, 125
    - martingales, 503
  - Hájek-Rényi-Chow, 503
  - Hoeffding, 120
  - Hölder, 129
  - Jensen, 132
  - Kahane-Hoffmann-Jørgensen, 141
  - Khintchine, 147, 149
  - KHJ, 141
  - Kolmogorov, 122
    - the other, 123
  - Kolmogorov-Doob, 501, 543
  - Lévy, 139, 140
    - extended, 140
  - Lyapounov, 129
  - Marcinkiewicz-Zygmund, 150, 151
    - martingales, 506
  - Markov, 120
  - median/moments, 133
  - Minkowski, 129
  - Ottaviani, 145
  - Rosenthal, 152, 507
  - Skorohod, 145
  - symmetrization
    - strong, 134
    - weak, 134
- infinite divisibility, 442–447
  - characteristic function, 444
  - finite variance, 444
  - characterization, 442, 444
    - finite variance, 444, 447
  - convergence, 447
    - finite variance, 446
  - definition, 442
  - properties, 443
  - random indices, 466

- triangular array, 442
  - finite variance, 447
- infinitely often (i.o.), 96
- insurance risk theory, 87
- integral
  - Lebesgue, 32, 34, 35
  - Riemann, 32, 33
  - Riemann-Stieltjes, 32, 33
- integration, 32
  - partial, 66
- intersection, 3
- invariance principle, 376, 451
  
- Jensen's inequality, 132
  - conditional, 477
- $\mathbb{J}_G$ , 571
- joining
  - martingales, 486, 497, 498
  - stopping times, 494
  - submartingales, 486
  - supermartingales, 486, 498
- joint
  - density, 43
  - distribution function, 43
  - probability function, 43
  
- Kahane-Hoffmann-Jørgensen inequality, 141
- Kertész, 480
  - Ad kudarc, 480
- Khintchine's inequality, 147, 149
- Kolmogorov, 2
  - axioms, 10, 26
  - convergence criterion, 286
  - inequality, 122, 502
    - the other, 123, 502
  - law of the iterated logarithm, 384, 396
  - strong law, 295
  - sufficient condition, 288
  - three-series theorem, 289
  - zero-one law, 20, 21, 72, 521
- Kolmogorov-Doob inequality, 501, 543
- Kolmogorov-Feller weak law, 279, 281
- Krickeberg decomposition, 490
- Kronecker's lemma, 565
  - random, 284
  
- large deviations, 366
  
- last exit times, 321, 418, 492
  - law of large numbers, 321
  - moments, 321
  - logarithmic moments, 418
  - moments, 418, 538
- lattice distribution, 163
  - characteristic function, 163, 164
  - span, 163
- law of large numbers, 265–327, 544–546
  - convergence rates, 321
  - counting variable, 321
  - last exit times, 321
  - precise asymptotics, 371
  - reversed martingales, 544–546
  - strong, *see* strong law of large numbers
  - weak, *see* weak law of large numbers
- law of the iterated logarithm, 383–421
  - Anscombe, 416
  - Chung, 419
  - cluster set, 405, 406, 411, 416, 417
  - convergence rates, 418
  - converse, 396
    - Strassen, 384
  - counting variable, 418
  - exponential bounds
    - lower, 386
    - upper, 386
  - first passage times, 417
  - Hartman-Wintner, 384, 387
    - Berry-Esseen, 412
  - Kolmogorov, 384, 396
  - last exit times, 418
  - precise asymptotics, 419
  - random index, 416
  - records, 417
  - subsequences, 398–413
    - cluster set, 405, 406, 411
    - dense, 399, 406, 411
    - sparse, 400, 402, 403, 405, 406, 414
  - the other LIL, 419
- law of total probability, 18
- Lebesgue
  - decomposition theorem, 37
  - dominated convergence theorem, 57
    - conditional, 473
  - integral, 32, 34, 35
  - integration, 46
  - measure, 35

- Lévy
  - convergence of series, 292
  - distance, 264
  - inequalities, 139, 140
- likelihood ratio test, 91
  - martingales, 483
  - regularity, 538
- limits
  - of probabilities, 11
  - of sets, 4
- Lindberg condition, 330
- Lindeberg-Lévy-Feller central limit theorem, 331
- local limit theorem, 365
  - first passage times, 365
- log-likelihood, 91, 154
- log-likelihood ratio, 551
- $L^p$ -spaces, 131
  - conjugate exponents, 132
  - norm, 131
- Lyapounov's condition, 339, 378
- Lyapounov's inequality, 129
  
- $m$ -dependence, 448
- Marcinkiewics-Zygmund strong law
  - precise asymptotics, 372
- Marcinkiewicz-Zygmund inequality, 150, 151
  - strong law, 298
  - $r$ -mean convergence, 311
  - weak law, 272
- marginal distribution, 44, 181
  - function, 44
- Markov's inequality, 120
- martingale
  - transform, 485
- martingale differences
  - definition, 478
  - expectation, 481
  - orthogonality, 488
- martingales, 467–553
  - almost sure convergence, 517, 520
  - asymptotic, 514
  - branching process, 483
  - Burkholder inequalities, 506, 507
  - central limit theorem, 515
  - closable, 517
  - closing random variable, 517, 520
  - complete, 517
  - convergence, 508–515
    - almost sure, 508
    - Garsia's proof, 508–511
    - in  $L^1$ , 517, 520
    - in  $L^p$ , 517
    - upcrossings proof, 511–513
  - convex functions, 487
  - defining relation, 479
  - definition, 478, 480, 496
  - Doob's maximal inequality, 505
  - double or nothing, 482, 519
  - exit times, 534
  - expectation, 480
  - exponential, 483, 525, 527
  - first passage times, 499
  - Galton-Watson process, 483, 519
  - Hájek-Rényi-Chow inequality, 503
  - inequalities, 501–508, 550
  - joined, 486, 497, 498
  - Kolmogorov-Doob inequality, 501
  - Krickeberg decomposition, 490
  - $L^1$ -convergence, 517, 520
  - $L^p$ -convergence, 517
  - likelihood ratio test, 483
  - main convergence theorem, 517
  - maximal function, 506
  - moment generating function, 483
  - non-convergent, 515
  - non-regular, *see* regular martingales
  - optional stopping, 495, 497
  - orthogonal increments, 488
  - powers, 487
  - predictable, 551
  - products of independent random variables, 482
  - regular, *see* regular martingales
  - regular stopping times, *see* regular stopping times
  - reversed, *see* reversed martingales
  - square function, 506, 552
    - conditional, 506, 552
  - stopped, *see* stopped martingales
  - sums of independent random variables, 482, 483, 507
  - uniform integrability, 484, 495, 497, 515, 517, 544
  - Wald, 539
    - non-regular, 541
- matching problem, 24

- max-stable distribution, 452
  - definition, 451
- maximal function, 506
- mean, 46, 62
  - absolutely continuous distributions, 62
  - Cantor distribution, 73
  - discrete distributions, 62
  - first passage times, 536
  - perverse distributions, 63
  - random vector, 181
  - stopped random walks, 523
  - sums of a random number of random variables, 84, 194
- measurability, 26–28, 205
- measurable function, 25
- measurable space, 14
- measures of dependence, 449
- median, 133, 140
- mesh, 33
- metatheorem, 9, 10, 13, 18
- method of moments, 237
- Mill's ratio, 109, 558
- Minkowski's inequality, 129
  - conditional, 477
- mixing coefficients, 450
- moment generating function, 189
  - continuity theorem, 242
  - moments, 190
  - multiplication theorem, 190
  - random vector, 191
  - sums of a random number of random variables, 194
  - uniqueness, 189
- moment problem, 194, 237
  - sums of a random number of random variables, 196
- moments, 62
  - absolute, 62
  - absolute central, 62
  - and tail probabilities, 74–79, 268
    - subsequences, 77, 78
  - central, 62
  - characteristic function, 176–178
  - convergence, *see* convergence
  - correlation, 130
  - covariance, 130
  - cumulants, 185
  - domain of attraction, 436
  - first passage times, 537
  - generating function, 188
  - last exit times, 538
  - $L^p$ -spaces, 131
  - mean, 62
  - moment generating function, 190
  - norm, 131
  - stable distributions, 426
  - stopped random walks, 521, 522, 530, 536
    - last summand, 536
    - stopping times, 530
  - variance, 62
- monkey and typewriter, 99
- monotone class, 6
  - generator, 7
  - theorem, 8
- monotone convergence, 55, 56, 213
  - conditional, 473
- multiplication theorem
  - characteristic function, 164
  - cumulant generating function, 185
  - generating function, 187
  - moment generating function, 190
- multivariate
  - random variable, 43
- multivariate normal distribution, 181, 375
  - characteristic function, 182
  - density, 182
  - independence
    - sample mean, variance, 183
    - uncorrelation, 182, 183
  - marginal, 181
- non-degenerate
  - distribution function, 30
  - random variable, 27
- norm, 131
- normal numbers, 305, 383
- null set, 12
- open set, 15
- Ottaviani's inequality, 145
- pair-wise independence
  - Borel-Cantelli lemmas, 104
  - central limit theorem, 344
  - Chebyshev's inequality, 122

- events, 19
- random variables, 71
- strong law, 297
- weak law, 276
- zero-one law, 105
- Parseval's relation, 172
- partial integration, 66
- partial maxima, 93, 94, 268
  - complete convergence, 312
  - convergence rates, 312
  - random indices, 466
  - weak law of large numbers, 270
- partition, 17
  - mesh, 33
- peak numbers, 448
- perverse distribution, 42, 63
- point-wise
  - equality, 27
- Poisson process, 89
- Poisson approximation, 364, 459
  - counting process
    - records, 364, 462
  - dependence, 462
  - Stein-Chen method, 459
- Pólya's theorem, 262
- potential
  - definition, 478
  - Doob decomposition, 551
- power set, 4, 6
- Pratt's lemma, 221
- pre- $\tau$ - $\sigma$ -algebra, 493
- precise asymptotics
  - first passage times, 373
  - law of large numbers, 371
  - law of the iterated logarithm, 419
  - Marcinkiewics-Zygmund strong law, 372
  - records, 374
- predictable, 477, 551
- probability function, 36
  - conditional, 79
  - joint, 43
- probability generating function, *see*
  - generating function
- probability measure, 10
- probability space, 2, 10
  - complete, 12
- probability triple, 2, 10
- probablity measure
  - induced, 26
- probablity space
  - induced, 26
- pseudo-distribution, 231, 235
  - function, 232–234
- quadratic variation, 148
- queueing models, 88
- Rademacher functions, 146
- Radon-Nikodym
  - derivative, 470
  - theorem, 470
- random element, 45
- random experiment, 1
- random index, 81–93, 192–194, 301–308
  - central limit theorem, 346
  - complete convergence, 322
  - convergence
    - almost sure, 302
    - in distribution, 345
    - stable distributions, 440
  - convergence rates, 322
  - infinite divisibility, 466
  - law of the iterated logarithm, 416
  - partial maxima, 466
- random parameter, 81–83
- random variables, 25, 27
  - closing, 517, 518, 520
  - degenerate, 27, 30
  - elementary, 25
  - equivalence class, 26, 468, 470
  - extended, 26
  - functions of, 28
    - independence, 71
  - independence, 68, 70
    - of functions, 71
    - pair-wise, 71, 276
  - multivariate, 43
  - non-degenerate, 27
  - non-negative, 29
  - simple, 25, 29
  - sums of a random number, *see* sums
    - of a random number of random variables
  - symmetrized, 133
  - truncation, 121
  - with random indices, *see* random index

- with random parameters, *see* random parameter
- random vector, 43
  - characteristic function, 180, 246
  - covariance matrix, 181
  - generating function, 188
  - mean, 181
  - moment generating function, 191
  - normal, *see* multivariate normal distribution
- random walks, 88
  - Bernoulli, 88
  - first passage times, *see* first passage times, 535–541
  - renewal theory, 90, 306, 351, 417, 441
  - simple, 88
    - recurrence, 103
    - symmetric, 88
    - transience, 103
  - stopped, *see* stopped random walks
- Rao-Blackwell theorem, 476, 548
- rapid variation, 568
- records
  - counting process, 93, 101, 364
    - central limit theorem, 351
    - law of the iterated logarithm, 417
    - precise asymptotics, 374
    - strong law of large numbers, 308
  - double records, 101
  - ranks, 94, 95
  - record times, 93, 101
    - central limit theorem, 351
    - law of the iterated logarithm, 417
    - precise asymptotics, 374
    - strong law of large numbers, 308
  - record values, 93, 94
    - types, 457
  - William's representation, 418
- recurrence, 103
- regular martingales, 517, 520, 532, 533
  - non-regular, 519, 551
- regular reversed martingales, 543
- regular stopping times, 531–541
  - exit times, 534
  - definition, 532
  - first passage times, 536
  - likelihood ratio test, 538
  - stopped random walks, 532, 539
- regular variation, 281, 566–568
- remote events, 21
- rencontre problem, 24
- renewal
  - alternating process, 379
  - counting process, 89, 364
  - process, 89
  - theory, 89
    - for random walks, *see* random walks
- replacement
  - based on age, 92
  - cost, 92
- replacement based on age, 552
- reversed martingales, 541–548
  - almost sure convergence, 543, 544
  - closing random variable, 543
  - convergence
    - Garsia's proof, 543
    - in  $L^1$ , 543
    - in  $L^p$ , 544
  - definition, 542
  - Doob's maximal inequality, 543
  - Kolmogorov-Doob inequality, 543
  - $L^1$ -convergence, 543
  - $L^p$ -convergence, 544
  - law of large numbers, 544–546
  - main convergence theorem, 543
  - $U$ -statistics, 547–548
  - uniform integrability, 542, 543
- reversed submartingales
  - almost sure convergence, 544
  - closing random variable, 544
  - convergence
    - in  $L^p$ , 544
  - definition, 542
  - $L^1$ -convergence, 544
  - main convergence theorem, 544
  - uniform integrability, 544
- reversed supermartingales
  - definition, 542
- Riemann integral, 32, 33
- Riemann-Lebesgue lemma, 162
- Riemann-Stieltjes integral, 32, 33
- Riesz decomposition, 491, 550
- $r$ -mean convergence, 310, 311
- Rosenthal's
  - inequality, 152
- runs (of zeroes), 117
- sample mean

- $U$ -statistics, 547
- sample space, 2, 10, 25, 43
- sample variance, 116, 380
- $U$ -statistics, 548
- Scheffé's lemma, 227, 228
- self-normalized sums, 380
- semi-invariants, 185
- sequential analysis, 91, 483, 538
- sets, 3
  - boundary, 15
  - Cantor, 41
  - cardinal number, 3
  - closed, 15
  - collection, 5
  - complement, 3
  - countable, 3
  - dense, 31, 570–574
  - difference, 3
  - disjoint, 4, 10
  - empty, 4, 11
  - i.o., 96
  - intersection, 3
  - liminf, 4, 96
  - limits, 4
  - limsup, 4, 96
  - non-decreasing, 4
  - non-increasing, 4
  - null set, 12
  - open, 15
  - operations, 3
  - power set, 4
  - subset, 4
  - symmetric difference, 3
  - union, 3
- Shakespeare, 99
- simple
  - function, 34
  - random variable, 25
  - random walks, 88
    - recurrence, 103
    - symmetric, 88
    - transience, 103
- singular distribution, 36
- skewness, 185
- Skorohod's inequality, 145
- Skorohod's representation theorem, 258
- slow variation, 281, 566–568
- smoothing, 474
- square function, 506, 552
- St. Petersburg
  - game, 283
  - paradox, 283
- stable distribution, 171, 424–427
  - Anscombe, 440
  - characteristic function, 426, 427
  - definition, 424
  - domain of attraction, 431
  - first passage times, 441
  - Hamel equation, 425
  - moments, 426
  - random indices, 440
  - strictly, 424
- Stein-Chen method, 459–464
  - coupling, 460, 464
  - Poisson approximation, 459
- Stirling's formula, 350
- stopped martingales, 485, 495, 497, 521–541
  - bounded stopping, 496, 497
- stopped random walks, 521–531
  - gambler's ruin problem, 526
    - expected duration, 528
  - last summand
    - moments, 536
  - mean, 523
  - moments, 521, 522, 536
    - converse, 530
  - regular stopping times, 532, 539
  - variance, 523
- stopping time, 145
- stopping times, 491–495
  - definition, 492
  - Fatou's lemma, 513
  - joined, 494
  - pre- $\tau$ - $\sigma$ -algebra, 493
  - properties, 493
  - regular, *see* regular stopping times
- strong law of large numbers, 288, 294–301
  - complete convergence, 315
  - convergence in  $L^1$ , 309
  - convergence rates, 316, 321
  - counting variable, 315
  - first passage times, 306
  - Hsu-Robbins-Erdős, 312
  - Kolmogorov, 295
  - Kolmogorov sufficient condition, 288

- Marcinkiewicz-Zygmund, *see*
  - Marcinkiewicz-Zygmund
  - pair-wise independence, 297
  - random indices, 302
  - records, 308
  - $r$ -mean convergence, 310, 311
- sub-probability distribution, 231
  - function, 31
- submartingale differences
  - definition, 478
  - expectation, 481
- submartingales
  - almost sure convergence, 518
  - closing random variable, 518
  - convergence
    - almost sure, 508
    - Garsia's proof, 508–511
    - in  $L^1$ , 518
    - upcrossings proof, 511–513
  - convex functions, 487
  - defining relation, 479
  - definition, 478, 480
  - Doob decomposition, 489
  - expectation, 480
  - inequalities, *see* martingales
  - joined, 486
  - Krickeberg decomposition, 490
  - $L^1$ -convergence, 518
  - main convergence theorem, 518
  - powers, 487
  - reversed, *see* reversed submartingales
  - uniform integrability, 518
- subsequence, 77, 78, 212, 235
  - Borel-Cantelli lemmas, 98
  - Helly's selection principle, 232
  - law of the iterated logarithm, 399,
    - 400, 402, 403, 405, 406, 411, 413, 414
  - principle, 229, 230
  - tight, 237
- subset, 4
- sufficient statistic, 475, 476, 548
- sums of a random number
  - mean, 84
  - variance, 84
- sums of a random number of random variables, 83–93, 192–194
  - characteristic function, 193
  - generating function, 194
  - mean, 194
  - moment generating function, 194
  - moment problem, 196
  - variance, 194
- sums of dependent random variables
  - central limit theorem, 450
  - invariance principle, 451
- supermartingale differences
  - definition, 478
  - expectation, 481
- supermartingales
  - defining relation, 479
  - definition, 478, 480
  - expectation, 480
  - joined, 486, 498
  - potential, *see* potential
  - Riesz decomposition, 491, 550
- symmetric difference, 3
- symmetrization, 133, 134, 255, 256, 291
- symmetrization inequalities, 134
- tail- $\sigma$ -field, 20, 72, 520, 521
- tight, 236, 375
  - subsequence, 237
  - uniform integrability, 236
- transience, 103
- triangular array, 344
  - central limit theorem, 345
  - infinite divisibility, 442, 447
- truncation, 121
- $U$ -statistics, 547–548
  - reversed martingales, 547–548
  - sample mean, 547
  - sample variance, 548
- unbiased estimator, 548
- uniform integrability, 214–218, 309–311, 353
  - convergence of moments, 218, 221, 224
  - definition, 214
  - martingales, 484, 495, 497, 515, 517
    - reversed, 542
  - products of sequences, 217
  - reversed martingales, 543, 544
  - reversed submartingales, 544
  - submartingales, 518
  - sums of sequences, 217
  - tight, 236

- uniformly asymptotically negligible, 340
- union, 3
- uniqueness
  - characteristic function, 160, 250
  - continuity theorem, 251
  - convergence, 208
  - cumulant generating function, 185
  - generating function, 186
  - moment generating function, 189
- upcrossings, 511
  - almost sure convergence, 512
  - expectation, 512
  - lemma, 511
    - random, 512
- Uppsala telephone directory, 99
- vague convergence, *see* convergence
- variance, 62
  - absolutely continuous distributions, 62
  - Cantor distribution, 73
  - conditional, 475
  - discrete distributions, 62
  - perverse distributions, 63
  - stopped random walks, 523
  - sums of a random number of random variables, 84, 194
- variational distance, 227, 459, 461
- Vieta's formula, 198
- Wald
  - equations, 523
  - fundamental identity, 539
  - martingale, 539
    - non-regular, 541
- weak convergence, *see* convergence
- weak law of large numbers, 270–284
  - general, 274
  - Kolmogorov-Feller, 279, 281
    - St Petersburg, 279
  - Marcinkiewicz-Zygmund, 272
  - pair-wise independence, 276
  - partial maxima, 270
- Weibull distribution, 452, 453
- Weierstrass approximation theorem, 277
- Wiener process, 375
- William's representation, 418
- zero-one law, 21, 97, 287, 290
  - Kolmogorov, 20, 21, 72, 521
  - pair-wise independence, 105