Chapter 10

Characteristic functions

This is based on [Billingsley, Section 26]

1. Complex numbers, Taylor polynomials, etc

**Theorem 10.1** (Taylor polynomials). For any "smooth enough" function \( f \) we have the following identity

\[
(10.1) \quad f(x) = f(0) + f'(0)x + \cdots + \frac{f^{(n)}(0)}{n!} x^n + \frac{1}{n!} \int_0^x f^{(n+1)}(s)(x-s)^n ds
\]

**Proof.** This is integration by parts formula: Case \( n = 0 \) is

\[
f(x) = f(0) + \int_0^x f'(s) ds
\]

Suppose the formula holds for some \( n \geq 0 \). Then

\[
\int_0^x f^{(n+1)}(t)(x-s)^n ds = \frac{1}{n+1} \left[ \int_0^x f^{(n+1)}(s)(-(x-s)^{n+1})' ds \right] - \frac{1}{n+1} f^{(n+1)}(x-s)^{n+1} \bigg|_{s=0}^{s=x} + \frac{1}{n+1} \int f^{(n+2)}(s)(x-s)^{n+1} ds
\]

Putting this back into (10.1), we get the same formula for \( n + 1 \).

When \( \frac{1}{n!} \int_0^x f^{(n+1)}(s)(x-t)^n ds \to 0 \) as \( n \to \infty \) we get the series expansion for \( f(x) \). Special cases of interest in this course are:

\[
(10.2) \quad e^x = \sum_{n=0}^\infty \frac{x^n}{n!}
\]
10. Characteristic functions

\[
\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}
\]

\[
\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}
\]

We will also need sharp error estimates!!

Exercise 10.1. Use Theorem 10.1 to prove (10.2).

1.1. Complex numbers. A complex number is an expression \( z = x + iy \) where \( i^2 = -1 \). Note: \( x \) is called the real part of \( z \) and \( y \) is called the imaginary part of \( z \).

The modulus of a complex number is \( |z| = \sqrt{x^2 + y^2} \). Noting that \( |z|^2 = z\bar{z} \) with complex conjugate \( \bar{z} = a - iy \), we get

\[
|z_1z_2| = |z_1||z_2|
\]

Since \( |z| \) is a distance in \( \mathbb{R}^2 \), we get the triangle inequality \( |z_1 + z_2| \leq |z_1| + |z_2| \).

Arithmetic works as usual. For example, \((1 + i)^2 = 2i\). Some powers of \( i \): \( i^3 = -i, i^4 = 1, i^5 = i, i^6 = -1 \).

(Omitted in 2018)
One simple explanation why arithmetics works comes from matrix interpretation: To a complex number \( a + ib \) associate the matrix \[
\begin{bmatrix}
a & b \\
b & a
\end{bmatrix}
\]. Then 1 corresponds to \( I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \) and \( i \) corresponds to \( J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \). Under this model, multiplication of complex numbers \((a+ib)(c+id)\) corresponds to multiplication of matrices. For example, \( J^2 = -I \).

It is clear that

\[
i^n = \begin{cases} 
  i & n \text{ odd} \\
  -1 & n \text{ even}
\end{cases}
\]

More specifically:

\[
i^n = \begin{cases} 
  (-1)^k & n = 2k \\
  (-1)^k i & n = 2k + 1
\end{cases}
\]

The following is called de Moivre formula

\[
e^{ix} = \cos x + i \sin x
\]
Proof. We use (10.2):
\[
e^{ix} = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}
\]
\[
= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} + \frac{ix}{2} \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} = \cos x + i \sin x
\]

In particular, \(e^{i(x+y)} = e^{ix}e^{iy}\) which is a just a version of the formula for \(\cos(x+y)\) and \(\sin(x+y)\).

Exercise 10.2. Find all complex numbers \(z\) with the property that \(z^2 = i\).

1.2. Complex version of Taylor’s formula. Integration by parts works also for complex functions, so (10.1) holds also for \(f(x) = e^{ix}\). This gives
\[
e^{ix} = \sum_{n=0}^{n} \frac{(ix)^n}{k!} + \frac{i^{n+1}}{n!} \int_{0}^{x} (x-s)^n e^{is}ds
\]

Note that for \(x > 0\) we have \(|\int_{0}^{x} (x-s)^n e^{is}ds| \leq \int_{0}^{x} (x-s)^n ds = x^{n+1}/(n+1)\). Similarly, for \(x < 0\) we have \(|\int_{0}^{x} (x-s)^n e^{is}ds| \leq \int_{x}^{0} (s-x)^n = (-x)^{n+1}/(n+1)\). This gives
\[
\left|e^{ix} - \sum_{n=0}^{n} \frac{(ix)^n}{k!}\right| \leq \frac{|x|^{n+1}}{(n+1)!}
\]

However, this bound is not as good as we need for large \(|x|\).

Lemma 10.2.
\[
\left|e^{ix} - \sum_{n=0}^{n} \frac{(ix)^n}{k!}\right| \leq \min \left\{ \frac{2|x|^n}{n!}, \frac{|x|^{n+1}}{(n+1)!} \right\}
\]

Proof. These improved bound is based on the identity
\[
e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + \frac{i^n}{(n-1)!} \int_{0}^{x} (x-s)^{n-1}(e^{is} - 1)ds
\]

which comes from integration by parts backwards:
\[
\int_{0}^{x} (x-s)^n e^{is}ds = \frac{1}{i}(x-s)^n e^{is}|_{s=0}^{x} + \frac{n}{i} \int_{0}^{x} (x-s)^{n-1} e^{is}ds = \frac{1}{i}x^n + \frac{n}{i} \int_{0}^{x} (x-s)^{n-1}e^{is}ds
\]
\[
= -\frac{n}{i} \int_{0}^{x} (x-s)^{n-1}ds + \frac{n}{i} \int_{0}^{x} (x-s)^{n-1}e^{is}ds
\]

The error estimate is
\[
\left|\frac{n}{(n-1)!} \int_{0}^{x} (x-s)^{n-1}(e^{is} - 1)ds\right| \leq 2|x|^n/n.\] This gives
\[
\left|e^{ix} - \sum_{n=0}^{n} \frac{(ix)^k}{k!}\right| \leq \frac{2|x|^n}{n!}
\]

Combining this with the previous estimate (10.5) we get (10.6).
We will only need (10.6) for $n = 1, 2, 3$.

\[|e^{ix} - 1| \leq \min\{|x|, 2\} \quad \text{(10.7)}\]
\[|e^{ix} - (1 + ix)| \leq \min\{\frac{1}{2}x^2, 2|x|\} \quad \text{(10.8)}\]
\[|e^{ix} - (1 + ix - \frac{1}{2}x^2)| \leq \min\{\frac{1}{6}|x|^3, x^2\} \quad \text{(10.9)}\]

1.3. Integrating complex-valued random variables. If $Z : \Omega \to \mathbb{C}$ is a random variable, then $Z(\omega) = X(\omega) + iY(\omega)$. The property we will need is integration of products of complex-valued expressions in independent random variables:

**Proposition 10.3.** If $Z_1 = X_1 + iY_1$ and $Z_2 = X_2 + iY_2$ are independent then $E(Z_1Z_2) = E(Z_1)E(Z_2)$.

**Proof.** Write $Z_1Z_2 = X_1X_2 - Y_1Y_2 + i(X_1Y_2 + Y_1X_2)$ and integrate each term. \(\square\)

Complex version of Theorem 6.17(ii) is a bit more difficult.

**Proposition 10.4.** If $|Z|$ is integrable, then $Z$ is integrable and $|E(Z)| \leq E(|Z|)$

**Proof.** Writing $Z = X + iY$, we have $|X| \leq |Z|$ and $|Y| \leq |Z|$ so $X, Y$ are integrable. The inequality says that $\sqrt{(EX)^2 + (EY)^2} \leq E\sqrt{X^2 + Y^2}$. This is Jensen’s inequality for the convex function $d(x,y) = \sqrt{x^2 + y^2}$ (the distance in $\mathbb{R}^2$ is a convex function!). \(\square\)

2. Characteristic functions

**Definition 10.1.** The characteristic function of a real-valued random variable $X$ is
\[\varphi(t) = Ee^{itX} \quad \text{(10.10)}\]

In principle, $\varphi(t)$ contains the same amount of "information" as a pair of functions $E\cos(tX)$ and $E\sin(tX)$. But it is convenient to use the standard properties of the exponential function.

**Remark 10.5.** Of course, $\varphi(0) = 1$ and $|\varphi(t)| \leq 1$ for all $t$. In fact, $\varphi(t)$ is uniformly continuous:
\[|\varphi(t + h) - \varphi(t)| \leq \int_{\mathbb{R}} |e^{ihx} - 1|F(dx) \to 0 \text{ as } h \to 0\]

This is a consequence of the Lebesgue dominated convergence theorem, but it is a good exercise to deduce it directly from (10.7):
\[\int_{\mathbb{R}} |e^{ihx} - 1|F(dx) = \int_{|x| > 1/\sqrt{h}} |e^{ihx} - 1|F(dx) + \int_{|x| \leq 1/\sqrt{h}} |e^{ihx} - 1|F(dx)\]
\[\leq \int_{|x| > 1/\sqrt{h}} 2F(dx) + \int_{|x| \leq 1/\sqrt{h}} \sqrt{h}F(dx) \leq 2P(|X| > 1/\sqrt{h}) + \sqrt{h} \to 0 \text{ as } h \to 0\]

The characteristic function $\varphi(t)$ can be thought as the moment generating function $M(it)$ applied to complex argument $it$.

**Example 10.1.** The moment generating function of the exponential random variable with density $f(x) = e^{-x}1_{x>0}$ is defined only for $t < 1$ and is given by $M(t) = \frac{1}{1-t}$. The characteristic function $\varphi(t) = \frac{1}{1-it}$ is defined for all real $t$. 
Example 10.2. The moment generating function of the standard normal random variable with density \( f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \) is given by \( M(t) = e^{t^2/2} \). The characteristic function is \( \varphi(t) = e^{-t^2/2} \).

The basic properties to establish are:

- The characteristic function uniquely determined the distribution.
- \( X_n \xrightarrow{D} X \text{ iff } \varphi_n(t) \to \varphi(t) \text{ for all } t \)
- If \( X, Y \) are independent then \( \varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) \).

The last property is the easiest, so we can prove it now:

**Proposition 10.6.** If \( X, Y \) are independent and \( U = aX + bY + c \) then \( \varphi_U(t) = \varphi_X(at)\varphi_Y(bt) e^{ict} \).

**Proof.** Use algebra \( e^{itU} = e^{iatX}e^{ibtY}e^{ict} \) and Proposition 10.3. \( \square \)

### 2.1. Useful estimates.

We conclude with some useful estimates that follow directly from (10.6)

If \( \varphi(t) = E e^{itX} \) and \( X \) is square-integrable, then

\[
\begin{align*}
(10.11) & \quad |\varphi(t) - 1| \leq E(\min\{(|tX|, 2)\}) \\
(10.12) & \quad |\varphi(t) - (1 + itE(X))| \leq E(\min\{\frac{1}{2}(tX)^2, 2|tX|\}) \\
(10.13) & \quad |\varphi(t) - (1 + itE(X) - \frac{t^2}{2} E(X^2))| \leq E(\min\{\frac{1}{6}|tX|^3, (tX^2)\})
\end{align*}
\]

### 2.2. Moments and derivatives.

**Theorem 10.7.** If \( E(|X|^n) < \infty \) then \( \varphi(t) \) has the \( n \)-th derivative and \( i^n E X^n = \varphi^{(n)}(0) \).

We note that the theorem has partial converse: if \( \varphi(t) \) has a derivative of even order \( 2k \) then \( E(|X|^{2k}) < \infty \).

**Proof.** We verify only the property for the first moment:

\[
\frac{\varphi(h) - \varphi(0)}{h} - iE(X) = E\left(\frac{e^{ihX} - 1 - ihX}{h}\right)
\]

and we note that

\[
\left|\frac{e^{ihX} - 1 - ihX}{h}\right| \leq 2|X|
\]

by (10.8). So we can apply the dominated convergence theorem

\[
\lim_{h \to 0} E\left(\frac{e^{ihX} - 1 - ihX}{h}\right) = E\left(\lim_{h \to 0} \frac{e^{ihX} - 1 - ihX}{h}\right) = 0
\]

**Theorem 10.8.** If \( F \) has a density then \( \varphi(t) \to 0 \text{ as } t \to \infty \).

**Proof.** Suppose \( f \) is the density. Then for every \( \varepsilon > 0 \) there is a step function \( g = \sum \alpha_k I_{(a_k, b_k]} \) such that \( \int |f - g| \, dx < \varepsilon \), so \( |\varphi(t) - \int e^{itx} g(x) \, dx| < \varepsilon \) for all \( t \). Now \( \int e^{itx} g(x) \, dx = \sum_k \alpha_k \frac{e^{itb_k} - e^{ita_k}}{it} \to 0 \) as \( t \to \infty \). \( \square \)
3. Uniqueness

The fact that characteristic function determines distribution uniquely can be proved in many ways. For example, one can use Weierstrass theorem — for a sketch of such a proof, see [Billingsley, Exercise 26.19]. Or one can use convolutions – such a proof appears e.g. in [Resnik]. We will follow [Billingsley].

3.1. Inversion formula.

**Theorem 10.9 (Inversion Formula).** If a cumulative distribution function \( F \) has characteristic function \( \varphi(t) \) then for points of continuity of \( F \) we have

\[
F(b) - F(a) = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt
\]

(10.14)

Since points of continuity are dense, we have uniqueness.

**Corollary 10.10 (uniqueness).** If \( F, G \) have the same characteristic function \( \varphi(t) \) then \( F(x) = G(x) \) for all \( x \).

**Proof.** Denote by \( I_T \) the right hand side of (10.14). Fubini’s theorem gives

\[
I_T = \frac{1}{2\pi} \int_{\mathbb{R}} \left( \int_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{it} dt \right) F(dx)
\]

We then re-write the inner integral using the fact that sin is odd while cos is even:

\[
\int_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{it} dt = \frac{1}{\pi} \int_{0}^{T} \frac{\sin(t(x-a))}{t} dt - \frac{1}{\pi} \int_{0}^{T} \frac{\sin(t(x-b))}{t} dt
\]

We need the following non-obvious fact (see Example 7.2 on page 83)

\[
\lim_{T \to \infty} \int_{0}^{T} \frac{\sin t}{t} dt = \frac{\pi}{2}.
\]

Noting that

\[
\lim_{T \to \infty} \int_{0}^{T} \frac{\sin tx}{t} dt = \lim_{T \to \infty} \int_{0}^{T} \frac{\sin u}{u} du = \begin{cases} 
\pi/2 & \text{if } x > 0 \\
0 & \text{if } x = 0 \\
-\pi/2 & \text{if } x < 0 
\end{cases}
\]

we get

\[
\lim_{T \to \infty} \frac{1}{\pi} \int_{0}^{T} \frac{\sin(t(x-a))}{t} dt - \frac{1}{\pi} \int_{0}^{T} \frac{\sin(t(x-b))}{t} dt = \begin{cases} 
-\frac{1}{2} + \frac{1}{2} = 0 & \text{if } x < a \\
0 + \frac{1}{2} = 1/2 & \text{if } x = a \\
\frac{1}{2} + \frac{1}{2} = 1 & \text{if } a < x < b \\
\frac{1}{2} - 0 = 1/2 & \text{if } x = b \\
\frac{1}{2} - \frac{1}{2} = 0 & \text{if } x > b 
\end{cases}
\]

\[\square\]

The following application of inversion formula deals with the densities.
Theorem 10.11. If \( \varphi(t) \) is integrable then \( F \) has density

\[
f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt
\]

Proof. Taking \( T \to \infty \) in (10.14) we get

\[
\frac{F(x+h) - F(x)}{h} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} - e^{-it(x+h)} \varphi(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-ith}}{it} e^{-itx} \varphi(t) dt
\]

and the formula for \( f(x) = F'(x) \) follows from Lebesgue’s dominated convergence theorem, as \( \lim_{h \to 0} \frac{1 - e^{-ith}}{it} = 1 \).

As an application of Theorem 10.11 we deduce the following.

Proposition 10.12. The characteristic function of the Cauchy distribution with \( F(x) = \frac{1}{2} + \frac{\arctan(x)}{\pi} \) is \( \varphi(t) = e^{-|t|} \)

Proof. Since \( \varphi \) is integrable, we compute the density of the Cauchy distribution:

\[
f(x) = \frac{1}{2\pi} \int_{-\infty}^{0} e^{-itx} e^{-|t|} dt = \frac{1}{2\pi} \int_{-\infty}^{0} e^{-itx} e^t dt + \frac{1}{2\pi} \int_{0}^{\infty} e^{-itx} e^{-t} dt
\]

\[
= \frac{1}{2\pi} \left( \frac{1}{1 - ix} + \frac{1}{1 + ix} \right) = \frac{1}{\pi} \frac{1}{1 + x^2}
\]

4. The continuity theorem

Theorem 10.13. Let \( F_n, F \) be cumulative distribution functions with characteristic functions \( \varphi_n \) and \( \varphi \). Then the following conditions are equivalent:

(i) \( F_n \xrightarrow{D} F \)

(ii) \( \varphi_n(t) \to \varphi(t) \) for each \( t \).

Note that \( \varphi_n(t) \) may converge without \( F_n \xrightarrow{D} F \), see Exercise 10.6.

Proof. If \( X_n \xrightarrow{D} X \) then \( \varphi_n(t) \xrightarrow{D} \varphi(t) \) for all \( t \) by Portmanteau Theorem 9.7.

The difficult part of proof is to show that the convergence of \( \varphi_n(t) \to \varphi(t) \) for each \( t \) implies \( F_n \xrightarrow{D} F \). The plan of proof for the converse implications is as follows:

- Show that \( \varphi_n(t) \xrightarrow{D} \varphi(t) \) implies tightness.
- Use Prokhorov’s theorem (Theorem 9.9) to deduce that \( F_n \) has convergent subsequences
- From the uniqueness theorem (Corollary 10.10) we verify that all limiting distributions \( F' \) are the same.
- By Theorem 9.10 we deduce convergence \( F_n \xrightarrow{D} F \).

Clearly, the first step is the where the difficulty lies.

Lemma 10.14. If \( \varphi_n(t) \xrightarrow{D} \varphi(t) \) for all \( t \) in a neighborhood of 0 then \( \{X_n\} \) is tight.
Proof. Since \( \varphi(0) = 1 \), and \( \varphi \) is continuous at \( t = 0 \), for all \( u \) small enough \( \frac{1}{u} \int_{-u}^{u} (1 - \varphi(t)) dt < \varepsilon \).
(Note that the integral is real.) Since \( \varphi_n(t) \to \varphi(t) \) and \( |1 - \varphi_n(t)| \leq 2 \) by Lebesgue’s dominated convergence theorem, there exists \( n_0 \) such that \( \frac{1}{u} \int_{-u}^{u} (1 - \varphi_n(t)) dt < 2\varepsilon \) for all \( n > n_0 \).

Now we use Fubini’s theorem:
\[
\frac{1}{u} \int_{-u}^{u} (1 - \varphi_n(t)) dt = E \left( \int_{-u}^{u} \frac{1 - e^{itX_n}}{u} dt \right) = E \left( \int_{-u}^{u} \frac{1 - \cos tX_n}{u} dt \right) = E \left( 2 - \int_{-u}^{u} \frac{\cos tX_n}{u} dt \right)
\]
\[
= 2E \left( 1 - \frac{\sin(u|X_n|)}{u|X_n|} \right) \geq 2 \int_{|X_n| \geq 2} \left( 1 - \frac{1}{|X_n|} \right) dP \geq 2 \int_{|X_n| \geq 2} \left( 1 - \frac{1}{|X_n|} \right) dP \geq P(|X_n| \geq 2).
\]

So with \( K = 2/u \) we have \( P(|X_n| \geq K) \leq 2\varepsilon \) for all \( n > n_0 \). Increasing \( K \) if necessary we can ensure that \( P(|X_n| \geq K) \leq 2\varepsilon \) holds also for the finitely many \( n \) preceding \( n_0 \).

As an application we give another proof of Slutski’s theorem (Theorem 9.4).

Corollary 10.15. If \( X_n \overset{D}{\to} X \) and \( Y_n \overset{P}{\to} 0 \) then \( X_n + Y_n \overset{D}{\to} X \)

Proof.
\[
|\varphi_{X_n+Y_n}(t) - \varphi_{X_n}(t)| = \left| E(e^{itX_n}(e^{itY_n} - 1)) \right| \leq E(|e^{itY_n} - 1|)
\]

Since \( |e^{itY_n} - 1| \overset{P}{\to} 0 \) by Exercise 4.23 and is bounded by 2, Lebesgue’s dominated convergence theorem (Theorem 6.10) shows that \( \lim_{n \to \infty} \varphi_{X_n+Y_n}(t) - \varphi_{X_n}(t) = 0 \). So both sequences have the same limit.

Required Exercises

Exercise 10.3. Let \( Z \) be the standard normal \( N(0,1) \) r.v. Use Example 10.2 to compute the characteristic function of \( X = \mu + \sigma Z \). (This is called the general normal r.v.)

Exercise 10.4. Let \( X \) be a Poisson random variable with parameter \( \lambda \). That is, \( P(X = k) = e^{-\lambda} \lambda^k/k! \), \( k = 0, 1, \ldots \). Compute the characteristic function of \((X - \lambda)/\sqrt{\lambda}\) and find its limit as \( \lambda \to \infty \).

Exercise 10.5. Suppose \( X_1, X_2, \ldots \) are i.i.d. with \( P(X = \pm 1) = 1/2 \). Let \( S_n = X_1 + \cdots + X_n \). Compute the characteristic function of \( \frac{1}{\sqrt{n}} S_n \) and find its limit as \( n \to \infty \).

Exercise 10.6. Suppose \( U_n \) are uniform on \((-n,n)\). Compute the characteristic function \( \varphi_n(t) \) and find its limit as \( n \to \infty \).

Exercise 10.7. Prove the case \( n = 2 \) of Theorem 10.7.

Exercise 10.8. Suppose \( X, Y \) are independent exponential (i.e. with density \( e^{-x} \) for \( x > 0 \)). Compute the characteristic function of \( X - Y \).
Additional Exercises

Exercise 10.9. Suppose $X_n \xrightarrow{D} X$ and $a_n \to a$, $b_n \to b$. Use characteristic functions to show that $a_nX_n + b_n \to aX + b$.

Exercise 10.10. Suppose $X_1, X_2$ are independent and take values $\pm 1$ with equal probabilities. Show that the characteristic function of $X_1 + X_1X_2$ is $\cos^2 t$.

Exercise 10.11. Show that

$$P(X = a) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} e^{-ita} \varphi(t) dt$$

Exercise 10.12. Suppose $P(X = x_k) > 0$. Show that

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |\varphi(t)|^2 dt = \sum_k (P(X = x_k))^2$$

*Hint* See [Billingsley, Exercise 26.13]
Bibliography

[Gut] A. Gut, Probability: a graduate course
Index

$L_1$ metric, 11
$L_2$ metric, 11
$L_p$-norm, 59, 76
$\lambda$-system, 25
$\pi$-system, 25
$\sigma$-field, 16
$\sigma$-field generated by $X$, 41
distribution of a random variable, 42
Bernoulli random variables, 46
Binomial distribution, 17, 73
bivariate cumulative distribution function, 30
Bonferroni’s correction, 18
Boole’s inequality, 18
Borel $\sigma$-field, 41
Borel sigma-field, 16
Cantelli’s inequality, 65
cardinality, 9
Cauchy distribution, 113
Cauchy-Schwarz inequality, 58
centered, 63
Central Limit Theorem, 117
characteristic function, 110
characteristic function – continuity theorem, 113
Characteristic functions – uniqueness, 112
Characteristic functions – inversion formula, 112
Chebyshev’s inequality, 57
complex numbers, 108
conjugate exponents, 60
continuity condition, 14
Continuous mapping theorem, 103
converge in $L_p$, 61
converge in mean square, 61
convergence in distribution, 49, 99
converges in distribution, 125
converges in probability, 46
converges pointwise, 7
converges uniformly, 7
converges with probability one, 47
convex function, 58
correlation coefficient, 59
countable additivity, 14
covariance matrix, 128
cumulative distribution function, 26, 43
cylindrical sets, 32, 33
cylindrical sets, 32
de Moivre formula, 108
DeMorgan’s law, 8
density function, 29, 74
diadic interval, 133
discrete random variable, 73
discrete random variables, 45
equal in distribution, 43
events, 13, 17
expected value, 54, 69
Exponential distribution, 74
exponential distribution, 29
Fatou’s lemma, 71
field, 13
finite dimensional distributions, 32
finitely-additive probability measure, 14
Fubini’s Theorem, 82
Geometric distribution, 73
Hölder’s inequality, 60, 76
inclusion-exclusion, 18
independent $\sigma$-fields, 35
independent events, 35
independent identically distributed, 46
independent random variables, 44
indicator functions, 9
induced measure, 42
infinite number of tosses of a coin, 133
integrable, 69
intersection, 8
Jensen’s inequality, 58
joint cumulative distribution function, 30
joint distribution of random variables, 43
Kolmogorov’s maximal inequality, 91
Kolmogorov’s one series theorem, 91
Kolmogorov’s three series theorem, 92
Kolmogorov’s two series theorem, 92
Kolmogorov’s zero-one law, 90
Kolmogorov-Smirnov metric, 10, 11
Kronecker’s Lemma, 94

Lévy distance, 106
law of $X$, 42
Lebesgue’s dominated convergence theorem, 71, 72
Lebesgue’s dominated convergence theorem – used, 73, 89, 102, 114
Levy’s metric, 11
Levy’s theorem, 93
Lindeberg condition, 119
Lyapunov’s condition, 120
Lyapunov’s inequality, 58

marginal cumulative distribution functions, 30
Markov’s inequality, 57
maximal inequality, Etemadi’s, 95
maximal inequality, Kolmogorov’s, 91
mean square convergence, 76
measurable function, 41
measurable rectangle, 81
metric, 10
metric space, 10
Minkowski’s inequality, 59
Minkowski’s inequality, 76
moment generating function, 56, 77
moments, 55
Monotone Convergence Theorem, 69
multivariate normal, 127
multivariate normal distribution, 128
multivariate random variable, 41

negative binomial distribution, 18
normal distribution, 29

Poisson distribution, 18, 73
Polya’s distribution, 18
Portmanteau Theorem, 102
power set, 7
probability, 13
probability measure, 14
probability space, 13, 17
product measure, 82
quantile function, 44, 101

random element, 41
random variable, 41
random vector, 41

sample space, 13
Schéffe’s theorem, 100
section, 81
semi-algebra, 15
semi-ring, 15
sigma-field generated by $\mathcal{A}$, 16
simple random variable, 53

simple random variables, 45
Skorohod’s theorem, 101
Slutsky’s Theorem, 101
Standard normal density, 74
stochastic process with continuous trajectories, 43
stochastic processes, 42
stochastically bounded, 51
symmetric distribution, 97
tail $\sigma$-field, 36
Tail integration formula, 84
Taylor polynomials, 107
tight, 51
tight probability measure, 19
Tonelli’s theorem, 82
total variation metric, 11
truncation of r.v., 50
uncorrelated, 63
uniform continuous, 28
Uniform density, 74
uniform discrete, 28
uniform singular, 28
uniformly integrable, 72, 105
union, 8

variance, 56

Wasserstein distance, 11
weak convergence, 49
weak law of large numbers, 63
zero-one law, 36, 90