Random variables

1. Measurable mappings

Suppose $\Omega$ and $E$ are two sets. Often $E = \mathbb{R}$ or $E = \mathbb{R}^d$. Suppose $X : \Omega \rightarrow E$ i.e. $X$ is a function with domain $\Omega$ and target set $E$. Then $X$ induces a mapping

$$X^{-1} : 2^E \rightarrow 2^\Omega$$

defined by $X^{-1}(U) = \{ \omega \in \Omega : X(\omega) \in U \}$, where $U \subset E$.

**Proposition 4.1.** Properties of induced mapping:

(i) $X^{-1}(\emptyset) = \emptyset$, $X^{-1}(E) = \Omega$

(ii) $X^{-1}(U^c) = (X^{-1}(U))^c$

(iii) $X^{-1}(\bigcup_{t \in T} U_t) = \bigcup_{t \in T} X^{-1}(U_t)$

**Proof.** For (iii), $\omega \in X^{-1}(\bigcup_{t \in T} U_t)$ iff $\exists_{i \in T} X(\omega) \in U_i$. \qed

**Corollary 4.2.** If $B$ is a $\sigma$-field of subsets of $E$ then $X^{-1}(B)$ is a $\sigma$-field of subsets of $\Omega$.

**Proof.** This is based on the identities for inverse images under functions, see Proposition 4.1. \qed

**Definition 4.1.** A $\sigma$-field generated by $X$ is $\sigma(X) = X^{-1}(B)$.

Exercise 4.17 says that this is the smallest $\sigma$-field of subsets of $\Omega$ which makes $X$ measurable.

1.1. Random elements and random variables. Suppose $(\Omega, F, P)$ is a probability space and $E$ is a set with distinguished $\sigma$-field $B$. In most applications, $E$ is a separable complete metric space and $B$ is the Borel $\sigma$-field which is generated by the countable collection of open balls.

**Definition 4.2.** In analysis, $X$ is called a measurable function if $X^{-1}(B) \subset F$. In probability, $X$ is then called a random element of $E$.

If we want to indicate the $\sigma$-fields, we will write $X : (\Omega, F) \rightarrow (E, B)$.
The most important special cases are $\mathbb{E} = \mathbb{R}$ and $\mathbb{E} = \mathbb{R}^d$. When $\mathbb{E} = \mathbb{R}$, we say that $X$ is a random variable. When $\mathbb{E} = \mathbb{R}^d$, we say that $X$ is a random vector or that $(X_1, \ldots, X_d)$ is a multivariate random variable. In such cases, measurability can be verified somewhat easier.

**Proposition 4.3.** To verify whether $X : \Omega \to \mathbb{R}$ is a random variable we only need to verify that the sets $A_x = \{ \omega : X(\omega) \leq x \}$ are in $\mathcal{F}$ for every real $x$.

Similarly, to verify whether $(X, Y) : \Omega \to \mathbb{R}^2$ is measurable, we only need to verify whether for all $x, y \in \mathbb{R}$ we have $\{ \omega : X(\omega) \leq x, Y(\omega) \leq y \}$ is in $\mathcal{F}$.

**Proof.** Consider the set $\mathcal{U}$ of all sets $U \subset \mathbb{R}$ such that $X^{-1}(U) \in \mathcal{F}$. In view of Proposition 4.1, this is a sigma-field.

For $x \in \mathbb{R}$, the inverse image of the set $(-\infty, x]$ is in $\mathcal{F}$, so $(-\infty, x] \in \mathcal{U}$. Hence the generated sigma field $\sigma\{(-\infty, x] : x \in \mathbb{R}\} = \mathcal{B}$ is in $\mathcal{U}$. \hfill \square

**Remark 4.1.** The collection $X_1, \ldots, X_d$ of random variables (on the same probability space) defines random vector $(X_1, \ldots, X_d)$. (For $d = 2$, this is Exercise 4.16.)

We also remark that random elements of spaces of functions, such as $\mathbb{E} = C[0, 1]$, the space of all continuous functions on $[0, 1]$, or $\mathbb{E} = D[0, \infty)$, the space of right-continuous functions with left limits, are called stochastic processes rather than random functions. So we say "Wiener process" or "Poisson process", rather than random continuous function, or random piecewise-linear function.

1.2. Induced probability measures.

**Definition 4.3.** The distribution of a random variable $X : (\Omega, \mathcal{F}) \to (\mathbb{E}, \mathcal{B})$ is a probability measure $Q$ on $(\mathbb{E}, \mathcal{B})$ defined by

$$Q(U) = P(X^{-1}(U))$$

Sometimes $Q$ is called an induced measure and some authors use notation $Q = P \circ X^{-1}$.

If $X$ is a random variable, then its distribution is uniquely determined by the corresponding cumulative distribution function

$$F(x) = Q((-\infty, x]) = P(\{ \omega : X(\omega) \leq x \}) \quad (4.1)$$

In probability and statistics the latter is usually abbreviated to $F(x) = P(X \leq x)$ but this abbreviated notation is just the shorthand for the right hand side of (4.1).

**Definition 4.4.** We say that random variables $X, Y$, defined perhaps on different probability spaces, are equal in distribution, if they induce the same probability measure on $(\mathbb{R}, \mathcal{B})$.

In view of Proposition 2.14, this is equivalent to $X, Y$ having the same cumulative distribution function.

If $X, Y$ are two random variables on the same probability space $(\Omega, \mathcal{F}, P)$ then the pair $(X, Y)$ is a measurable mapping $\Omega \to \mathbb{R}^2$. The joint distribution of random variables is just the induced measure on $\mathbb{R}^2$ and is uniquely determined by the joint cumulative distribution function

$$F(x, y) = P(X \leq x, Y \leq y) \quad (4.2)$$

(Note the abbreviated notation for $P(\{ \omega : X(\omega) \leq x, Y(\omega) \leq y \}$ If $\mathbb{E} = C[0, 1]$ then a measurable mapping $X : \Omega \to C[0, 1]$ is called a stochastic process with continuous trajectories. The standard
notation is $X = (X_t)_{t \in [0,1]}$. The distribution of $X$ is uniquely determined by the family of finite dimensional distributions

$$F_{t_1,t_2,\ldots,t_k}(x_1,x_2,\ldots,x_k) = P(X_{t_1} \leq x_1, \ldots, X_{t_k} \leq k)$$

that satisfy natural consistency conditions. The converse is not as simple here: consistent families of finite-dimensional distributions

$$\{F_{t_1,t_2,\ldots,t_k}(x_1, x_2, \ldots, x_k) : 0 \leq t_1 \leq t_2 \leq \cdots \leq t_k \leq 1\}$$

define a probability measure on Borel sets of the product space $\mathbb{R}^{[0,1]}$ of all (including nonmeasurable) functions $[0,1] \rightarrow \mathbb{R}$ with pointwise convergence, see [Billingsley, Theorem 36.1] but not necessarily on Borel subsets of $C[0,1]$. (In fact, $C[0,1] \subset \mathbb{R}^{[0,1]}$ is not a Borel subset, see the discussion that follows [Billingsley, Theorem 36.3].)

The following properties are sometimes useful.

**Proposition 4.4.** If $AF$ then $I_A : \Omega \rightarrow \mathbb{R}$ is measurable. A non-decreasing right-continuous function $\mathbb{R} \rightarrow \mathbb{R}$ is measurable. A strictly increasing function $\mathbb{R} \rightarrow \mathbb{R}$ is measurable. A continuous function $\mathbb{R} \rightarrow \mathbb{R}$ is measurable. Sum of two measurable functions is measurable. Product of two measurable functions is measurable. A pointwise limit of a sequence of measurable functions $\mathbb{R} \rightarrow \mathbb{R}$ is a measurable function.

For example, $\mathbb{R} \ni x \mapsto e^x I_{(a,b)}(x)$ is measurable as a product of two measurable functions.

2. Random variables with prescribed distributions

This section is based on [Billingsley, Section 14] or [Durrett, Theorem 1.2.2].

**Theorem 4.5.** If $F$ is a cumulative distribution function\(^1\), then there exists on some probability space $\omega$ random variable $X$ for which $P(X \leq x) = F(x)$.

**First proof.** Proposition 2.14 gives a probability measure $P$ on $(\mathbb{R}, B)$ such that $F(x) = P((-\infty,x])$. Take $(\mathbb{R}, B, P)$ for the probability space $(\Omega, F, P)$. Define $X(\omega) = \omega$ (the identity mapping). Then $X$ has distribution $P$. \hfill $\square$

**Second proof.** [This is independent of Proposition 2.14, and in fact can be used to prove it.]

Let $\Omega = (0,1)$ with Lebesgue measure $\lambda$ on Borel sigma-field. Since $F$ is non-decreasing right-continuous with limits 0, 1, for $0 < u < 1$, the set $\{x : u \leq F(x)\}$ is\(^2\) a closed\(^3\) half-line\(^4\) of the form $[\varphi(u), \infty)$ and its complement is $\{x : u > F(x)\} = (-\infty, \varphi(u))$. This shows that for every real $x$, we have $\varphi(u) \leq x$ iff $F(x) \geq u$. This also defines the quantile function

$$\varphi(u) = \inf\{x : u \leq F(x)\} = \sup\{x : F(x) < u\}$$

Define $X(\omega) = \varphi(\omega)$. Then $\lambda(\{\omega : X(\omega) \leq x\}) = \lambda(\{\omega : \omega \leq F(x)\}) = \lambda([0,F(x)]) = F(x)$. \hfill $\square$

---

\(^1\)See Definition 2.3
\(^2\)Can you see why isn’t it $\mathbb{R}$ or $\emptyset$?
\(^3\)Why?
\(^4\)Why?
Corollary 4.6 (Proposition 2.14). If $F$ is a CDF then there exists a unique probability measure $P$ on the Borel sets of $\mathbb{R}$ such that $P((\infty, a]) = F(a)$.

Proof. Existence: Take Lebesgue measure on Borel sigma-field of $(0,1)$, and $X$ as in the second proof above. Then $P$ is the induced probability measure. (Uniqueness follows from Theorem 2.9, see Proof of Proposition 2.14). from $\pi - \lambda$ theorem, see □

Example 4.1. Write $X = X_+ - X_-$, i.e.

$$X_+(\omega) = \begin{cases} X(\omega) & \text{if } X(\omega) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$X_-(\omega) = (-X)_+ = \begin{cases} -X(\omega) & \text{if } X(\omega) \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

If $X$ has CDF $F(x)$, what are the CDFs of $X_+$ and $X_-$?

Solution.

$$P(X_+ \leq x) = \begin{cases} P(X \leq x) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

So $F_+(x) = F(x)I_{[0,\infty)}(x)$. □

2.1. Independent random variables. The second proof of Theorem 4.5 lets us construct a finite or an infinite sequence $X_1, X_2, \ldots$ of random variables with prescribed distributions. However, this gives only very special measures on $\mathbb{R}^\infty$, see Exercise 4.18. We now consider another special construction that gives joint distributions that are more often of interest.

Random elements $X_1, X_2, \ldots$ are independent if $\sigma$-fields $\sigma(X_1), \sigma(X_2), \ldots$ are independent.

Example 4.2. Suppose discrete random variables $X = \sum x_j I_{A_j}, Y = \sum y_k B_k$. Then $X, Y$ are independent if $A = \{A_1, A_2, \ldots\}$ and $B = \{B_1, B_2, \ldots\}$ are independent $\pi$-systems. Thus

$$P(X = x, Y = y) = P(X = x)P(Y = y) \text{ for all } x, y \in \mathbb{R}$$

Similarly, discrete random variables $X, Y, Z$ are independent iff

$$P(X = x, Y = y, Z = z) = P(X = x)P(Y = y)P(Z = z) \text{ for all } x, y, z \in \mathbb{R}$$

Example 4.3. Suppose $X_1, X_2, \ldots$ take only values 0, 1 and $p_k = P(X_k = 1), q_k = 1 - p_k$. Then $X_1, X_2, \ldots$ are independent iff

$$P(X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_n = \varepsilon_n) = \prod_{k=1}^n p_k^{\varepsilon_k} q_k^{1-\varepsilon_k}$$

for all choices of $\varepsilon_1, \ldots, \varepsilon_n \in \{0, 1\}$.

Independence is often assumed in the theorems. So it is of some interest to make sure that such random variables exist.

Theorem 4.7. If $F_1, F_2, \ldots$ are cumulative distribution functions then there exists a probability space $(\Omega, \mathcal{F}, P)$ and a sequence $X_1, X_2, \ldots$, of independent random variables such that $X_n$ has cumulative distribution function $F_n$. 

Sketch of First Proof. In this proof we take \( \Omega = \mathbb{R}^\infty \) with (infinite!) product measure\(^5\) \( P = P_1 \otimes P_2 \otimes \ldots \) where \( P_k \) is the probability measure on \( \mathbb{R} \) with cumulative distribution function \( F_k \). For \( \omega = (\omega_1, \omega_2, \ldots) \in \mathbb{R}^\infty \) we define \( X_k(\omega) = \omega_k \).

Sketch of Second Proof. \(^6\) We use \( \Omega = (0,1] \) with Lebesgue measure \( \lambda \) and with binary digits function \( d_n : (0,1] \to \{0,1\} \).

We first note that random variables \( d_1, d_2, \ldots \) are independent. Indeed, as noted in the proof of Proposition A.1 we have \( \lambda(d_1 = \varepsilon_1, \ldots, d_m = \varepsilon_m) = 1/2^m \). By Example 4.3 this proves independence.

Next, we arrange all of these random variables into an infinite array \( d_{i,j} \). Then random variables \( U_i(\omega) = \sum_{j=1}^{\infty} d_{i,j}(\omega)/2^j \) are independent. On the other hand, \( \lambda(\omega : U_i(\omega) \leq x) = x \); this is easiest to see for diadic rational numbers\(^7\) of the form \( x = k/2^n \).

Now take \( X_k = \phi_k(U_k) \), where \( \phi_k(u) \) is the quantile transform (4.3) of \( F_k \).

Definition 4.5. We say that \( X_1, X_2, \ldots \) are independent identically distributed (i. i. d.) random variables, if they are independent and have the same CDF.

2.2. Elementary examples.

Proposition 4.8. If \( f: \mathbb{R}^d \to \mathbb{R} \) is measurable (say, continuous) and \( X_1, \ldots, X_d : \Omega \to \mathbb{R} \) are random variables on \( \Omega, \mathcal{F}, P \), then \( Y = f(X_1, \ldots, X_d) \) is a random variable.

Proof. If \( B \) is a Borel subset of \( \mathbb{R} \) then \( U = f^{-1}(B) \subset \mathbb{R}^d \) is a Borel subset of \( \mathbb{R}^d \). So \( Y^{-1}(B) = (X_1, \ldots, X_d)^{-1}(U) \in \mathcal{F} \). \( \square \)

Here are some examples of such functions:

Proposition 4.9 (Sum theorems). Suppose \( X_1, X_2, \ldots \) are independent and \( S = X_1 + X_2 + \cdots + X_n \).

(i) If \( X_1, \ldots, X_n \) are i. i. d. Bernoulli random variables, i.e., \( P(X_j = 1) = p, P(X_j = 0) = 1 - p \), then \( S \) is Binomial \( \text{Bin}(n, p) \) (see Example 1.6)

(ii) If \( X_1, X_2, \ldots \) are Poisson random variables with parameters \( \lambda_1, \lambda_2, \ldots \) then \( S \) is Poisson with parameter \( \lambda = \lambda_1 + \cdots + \lambda_n \) (see Example 1.7)

(iii) If \( X_1, X_2, \ldots \) are i. i. d. Normal \( N(0,1) \) random variables (see Example 2.5) then \( Y = X_1 + \cdots + X_n \) is normal with mean zero and variance \( n \) (i.e., has same law as \( \sqrt{n}Z \) for some \( N(0,1) \) r.v. \( Z \).)

Proof. Omitted\(^8\) \( \square \)

\(^5\)We did not show how to construct infinite product measure!

\(^6\)This proof is from [Billingsley, Theorem 20.4]. It also answers Exercise 3.3.

\(^7\)Observe that intervals of the form \( [0, k/2^n] \) are a \( \pi \)-system that generates \( \mathcal{B} \)

\(^8\)These are “elementary” facts covered in undergraduate courses.
3. Convergence of random variables

**Definition 4.6.** A sequence of random variables converges in probability, \( X_n \xrightarrow{P} X \), if\(^9\)

\[
\lim_{n \to \infty} P(|X_n - X| \geq \varepsilon) = 0
\]

for every \( \varepsilon > 0 \)

**Example 4.4.** On \( \Omega = [0, 1] \) consider \( X_n = I_{[0,n/(2n+1)]} \). Then \( X_n \xrightarrow{P} X \).

Suppose \( X_n, X \) are random variables on some probability space \((\Omega, \mathcal{F}, P)\). Then we have

**Proposition 4.10.**

\[
\{ \omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega) \} \in \mathcal{F}
\]

**Proof.** First we note that for a fixed \( \varepsilon > 0 \), the set \( A_n = \{ \omega : |X_n(\omega) - X(\omega)| < \varepsilon \} \in \mathcal{F} \). This is a consequence of Exercise 4.16.

Next, we note that

\[ A_\varepsilon = \{ \omega : \forall \varepsilon \exists k > n |X_k(\omega) - X(\omega)| > \varepsilon \}
\]

is in \( \mathcal{F} \). Indeed, \( A_\varepsilon = \bigcap_n \bigcup_{k \geq n} A_k^{\varepsilon} \). Finally, we note that

\[
\bigcap_{\varepsilon > 0} A_\varepsilon = \bigcap_{n \in \mathbb{N}} A_{1/n} \in \mathcal{F}
\]

□

**Definition 4.7.** A sequence of random variables converges with probability 1 if

\[
P\left( \{ \omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega) \} \right) = 1
\]

**Proposition 4.11.** If \( X_n \to X \) with probability 1, then \( X_n \xrightarrow{P} X \)

**Proof.** The discussion of measurability shows that \( P(\forall \varepsilon > 0 \exists N \forall n > N \{ \omega : |X_n - X| < \varepsilon \}) = 1 \) iff for every rational \( \varepsilon > 0 \)

\[
P(\exists N \forall n > N |X_n - X| < \varepsilon) = P\left( \bigcup_{N} \bigcap_{n > N} |X_n - X| < \varepsilon \right) = 1
\]

This is the same as

\[
P(\bigcap_{N} \bigcup_{n > N} |X_n - X| > \varepsilon) = P(|X_n - X| > \varepsilon \text{ i.o.}) = 0
\]

Now \( P(\bigcap_{N} \bigcup_{n > N} |X_n - X| > \varepsilon) = \lim_{N \to \infty} P(\bigcup_{n > N} |X_n - X| > \varepsilon) \). So convergence with probability 1 is equivalent to

\[
\forall \varepsilon > 0 \lim_{N \to \infty} P(\sup_{n > N} |X_n - X| > \varepsilon) = 0.
\]

Of course, \( P(|X_n - X| > \varepsilon) \leq P(\sup_{n > N} |X_n - X| > \varepsilon) \).

\(^9\)We already switched to abbreviated notation:

\[
\lim_{n \to \infty} P(\{ \omega : |X_n(\omega) - X(\omega)| > \varepsilon \}) = 0
\]
Proposition 4.12. Suppose $X_n \xrightarrow{P} X$. Then there exists a subsequence $n_k$ such that $X_{n_k} \to X$ with probability 1.

Proof. Choose positive $\varepsilon_k \to 0$. Given $k$, choose $n_k > k$ so that $P(|X_{n_k} - X| > \varepsilon_k) < 1/2^k$. Since $\sum_k 1/2^k < \infty$, by the first Borel-Cantelli Lemma,

$$P(|X_{n_k} - X| > \varepsilon_k \text{ i.o.}) = 0$$

Therefore, for any $\varepsilon > 0$, $P(|X_{n_k} - X| > \varepsilon \text{ i.o.}) \leq P(|X_{n_k} - X| > \varepsilon_k \text{ i.o.}) = 0$

Details: Choose $N_0$ such that $\varepsilon_{n_k} < \varepsilon$ for $k > N_0$. Then

$$\bigcap_{N=1}^{\infty} \bigcup_{k>N} |X_{n_k} - X| > \varepsilon \subset \bigcap_{N>N_0} \bigcup_{k>N} |X_{n_k} - X| > \varepsilon \subset \bigcap_{N>N_0} \bigcup_{k>N} |X_{n_k} - X| > \varepsilon_k$$

Remark 4.2. Convergence in probability is a metric convergence. Convergence with probability 1 is not a "metric convergence".

Remark 4.3. Suppose $X_n$ are random variables such that $X_n(\omega)$ converges for all $\omega \in \Omega$. Then $X(\omega) := \lim_{n \to \infty} X_n(\omega)$ is a random variable.

Proof. $\{\omega : X(\omega) \leq x\} = \bigcap_{j \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} \bigcap_{k>n} \{\omega : X_k(\omega) \leq x + 1/j\}$

The third type of convergence, the so called convergence in distribution, is somewhat different, as it is really convergence of the induced probability measures, not random variables. This topic will appear later, but we can give a definition here:

Definition 4.8. We say that a sequence of $\mathbb{R}$-valued random variables $X_n$ with CDFs $F_n$ converges in distribution to a random variable $Y$ with CDF $F$, if $F_n(x) \to F(x)$ for all continuity points $x$ of $F$.

Convergence in probability is a “metric convergence” (the metric cannot be yet written), so it has the following property.

Proposition 4.13. If every subsequence of $\{X_n\}$ has a $P$-convergent subsequence, then all the limits must be equal (with probability one) and $X_n$ converges in probability.

Proof. If different subsequences converge to say $X'$ and $X''$, then by choosing a subsequence that alternates between the two subsequences we can check that $\Pr(|X' - X''| > \varepsilon) = 0$ for every $\varepsilon > 0$, so $X' = X''$ with probability one. Let's denote the common limit by $X$.

To prove convergence, to the above $X$, we proceed by contradiction. Suppose that $X_n$ does not converge in probability. Then there exists a subsequence $\{X_{n_k}\}_{k \in \mathbb{N}}$ such that $P(|X_{n_k} - X| > \varepsilon) > \delta > 0$. This subsequence cannot have a further subsequence that would converge to $X$. □
4. Random variables

To see that this is quite useful, try solving Exercise 4.21 without using Proposition 4.13. (Yes, it can be done!)

Every convergent sequence of numbers is bounded. An analog of this involves a separate concept which is introduced in Exercise 4.13.

---

**Required Exercises**

**Measurability.**

**Exercise 4.1.** Suppose that $\varphi : (0,1) \rightarrow \mathbb{R}$ is strictly increasing. Prove that $\varphi$ is measurable with respect to Borel sigma-fields.

**Exercise 4.2.** Suppose that $\varphi : (0,1) \rightarrow \mathbb{R}$ is continuous. Prove that $\varphi$ is measurable with respect to Borel sigma-fields.

**Exercise 4.3.** Prove one/some/all of the statements in Proposition 4.4.

**Cumulative distribution functions.**

**Exercise 4.4.** Consider probability space $((0,1), \mathcal{B}, \lambda)$. Suppose $X : (0,1) \rightarrow \mathbb{R}$ is given by $X(\omega) = \ln(\omega)$. Find the CDF of $X$.

**Exercise 4.5.** Suppose $X : \Omega \rightarrow \mathbb{R}$ has CDF $F$. Let $Y = X^2$. What is the CDF of $Y$?

**Exercise 4.6.** Suppose $X : \Omega \rightarrow \mathbb{R}$ has CDF $F$. Let $Y = X I_{|X| \leq M}$ be the truncation of r.v. $X$ at level $M$. What is the CDF of $Y$?

**Exercise 4.7.** Suppose $U$ is uniform on $(0,1)$. Let $X = U^2$, $Y = U^3$. What is their joint CDF? (See (4.2).)

**Exercise 4.8 (Statistics).** Use the second proof of Theorem 4.5 to describe how to simulate exponential random variables (see Example 2.4) using a random number generator that produces uniform $U(0,1)$ random variables.

**Independence.**

**Exercise 4.9.** Consider $\omega = [0,1]$ with Lebesgue measure and the measure-preserving map $f$ defined in Exercise 2.3. Show that events $A := [0,1/2]$, $B := f^{-1}(A)$ and $C := f^{-1}(B)$ are independent. (This is one of the possible answers to Exercise 3.3.)

**Convergence.**

**Exercise 4.10.** Suppose $F, G$ are two cumulative distribution functions. Define $\varphi_G(u) X$ has CDF $F$. Let $Y = X^2$. What is the CDF of $Y$?

**Exercise 4.11.** Suppose random variables

$$X_n = \begin{cases} n & \text{with probability } p_n \\ 0 & \text{with probability } 1 - p_n \end{cases}$$

Prove that
(i) if \( p_n \to 0 \) then \( X_n \xrightarrow{P} 0 \).
(ii) if \( \sum p_n < \infty \) then \( X_n \to 0 \) with probability 1.
(iii) if \( X_n \) are independent then \( X_n \to 0 \) with probability 1 iff and only if \( \sum p_n < \infty \)

Exercise 4.12. Prove that if \( X_n \xrightarrow{P} X \) and \( Y_n \xrightarrow{P} Y \) then \( X_n + Y_n \xrightarrow{P} X + Y \).

Exercise 4.13. Suppose \( X_n \xrightarrow{P} X \). Show that \( \{X_n\} \) is stochastically bounded (which is the same as the sequence of laws being tight, compare Exercise 1.13), i.e. for every \( \varepsilon > 0 \) there exists \( K > 0 \) such that for all \( n \) we have \( P(|X_n| > K) < \varepsilon \).

Exercise 4.14. Use the result from Exercise 4.13 to prove that if \( X_n \xrightarrow{P} X \) and \( Y_n \xrightarrow{P} Y \) then \( X_nY_n \xrightarrow{P} XY \).

Exercise 4.15. Suppose \( U_1, U_2, \ldots, U_n, \ldots \) are independent identically distributed \( U(0,1) \) random variables (i.e. with cumulative distribution function \( F(x) = x \) for \( 0 < x < 1 \), see Example 2.2). Show that the sequence \( Z_n = U_1U_2 \ldots U_n \) converges with probability 1.

Exercise 4.16. Suppose \( X : \Omega \to \mathbb{R} \) and \( Y : \Omega \to \mathbb{R} \) are two measurable functions (with respect to the Borel \( \sigma \)-field \( B(\mathbb{R}) \)). Prove that \( (X,Y) : \Omega \to \mathbb{R}^2 \) is measurable (with respect to the Borel \( \sigma \)-field \( B(\mathbb{R}^2) \)). (Hint: Proposition 4.3.)

Exercise 4.17. Prove that \( \sigma(X) \) as defined in the notes (as \( X^{-1}(B) \)) is in fact the smallest \( \sigma \)-field for which \( X \) is measurable. (This is the definition of \( \sigma(X) \) in [Billingsley].)

Exercise 4.18. Suppose \( X, Y \) are random variables with cumulative distribution functions \( F(x) \) and \( G(y) \), constructed as in the second proof of Theorem 4.5. Find the joint cumulative distribution function of \( X, Y \).

Exercise 4.19 (Statistics). Suppose \( X, Y \) are independent \( N(0,1) \) random variables. Verify that \( X^2 + Y^2 \) is exponential. Hint: use polar coordinates.

Exercise 4.20. Suppose that \( X_1 \leq X_2 \leq \cdots \leq X_n \leq X_{n+1} \leq \cdots \). If \( X_n \xrightarrow{P} X \), show that \( X_n \to X \) with probability one.

Exercise 4.21. Suppose \( f : \mathbb{R} \to \mathbb{R} \) is a continuous function and \( X_n \xrightarrow{P} X \). Prove that \( Y_n = f(X_n) \) converges in probability to \( Y = f(X) \).

Exercise 4.22. Suppose \( X_n \xrightarrow{P} X \) and \( X_n \) are independent. Show that there is \( a \in \mathbb{R} \) such that the cumulative distribution of \( X \) is \( F(x) = \begin{cases} 0 & x < a \\ 1 & x \geq a \end{cases} \).
Bibliography

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

$L - p$-norm, 50
diadic interval, 105
discrete random variable, 56
discrete random variables, 42
equal in distribution, 40
events, 11, 15
expected value, 47
Exponential distribution, 57
exponential distribution, 27
Fatou’s lemma, 55
field, 11
finite dimensional distributions, 30
finitely-additive probability measure, 12
Fubini’s Theorem, 62
Geometric distribution, 57
 Hölder’s inequality, 50, 58
inclusion-exclusion, 16
independent $\sigma$-fields, 33
independent events, 33
independent identically distributed, 42
indicator functions, 8
induced measure, 40
infinite number of tosses of a coin, 105
integrable, 54
intersection, 8
Jensen’s inequality, 49
joint cumulative distribution function, 28
joint distribution of random variables, 40
Kolmogorov’s maximal inequality, 67
Kolmogorov’s one series theorem, 68
Kolmogorov’s three series theorem, 68
Kolmogorov’s two series theorem, 68
Kolmogorov’s zero-one law, 67
Kronecker’s Lemma, 69
Lévy distance, 79
Lebesgue’s dominated convergence theorem, 55
Lebesgue's dominated convergence theorem – used, 56, 66, 75, 88
Levy’s theorem, 70
Lindeberg condition, 93
Lyapunov’s condition, 94
Lyapunov’s inequality, 49
marginal cumulative distribution functions, 28
Markov’s inequality, 49
maximal inequality, Etemadi’s, 70
maximal inequality, Kolmogorov’s, 67
measurable function, 39
measurable rectangle, 61
Minkowski’s inequality, 50
Minkowski’s inequality, 58
moment generating function, 60
Monotone Convergence Theorem, 55
multivariate normal, 99
multivariate normal distribution, 100
negative binomial distribution, 16
normal distribution, 27
Poisson distribution, 16, 57
Polya’s distribution, 16
Portmanteau Theorem, 75
power set, 7
probability, 11
probability measure, 12
probability space, 11, 15
product measure, 62
quantile function, 41, 75
random element, 39
random variable, 40
random vector, 40
sample space, 11
Scheffe’s theorem, 73
section, 61
semi-algebra, 13
semi-ring, 13
sigma-field generated by \( A \), 14
simple random variable, 47
Skorohod’s theorem, 75
Slutsky’s Theorem, 74
Standard normal density, 57
stochastic process with continuous trajectories, 40
stochastically bounded, 46
symmetric distribution, 71
tail \( \sigma \)-field, 34
Tail integration formula, 63
Taylor polynomials, 81
tight, 46
tight probability measure, 17
Tonelli’s theorem, 62
truncation of r.v., 45
uncorrelated, 51
uniform continuous, 26
Uniform density, 57
uniform discrete, 26
uniform singular, 26
uniformly integrable, 56, 77
union, 8
variance, 48
zero-one law, 34, 67