Part II — Probability and Measure
Definitions

Based on lectures by J. Miller
Notes taken by Dexter Chua

Michaelmas 2016

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Analysis II is essential

Measure spaces, σ-algebras, π-systems and uniqueness of extension, statement *and proof* of Carathéodory's extension theorem. Construction of Lebesgue measure on $\mathbb{R}$. The Borel σ-algebra of $\mathbb{R}$. Existence of non-measurable subsets of $\mathbb{R}$. Lebesgue-Stieltjes measures and probability distribution functions. Independence of events, independence of σ-algebras. The Borel–Cantelli lemmas. Kolmogorov's zero-one law. [6]

Measurable functions, random variables, independence of random variables. Construction of the integral, expectation. Convergence in measure and convergence almost everywhere. Fatou's lemma, monotone and dominated convergence, differentiation under the integral sign. Discussion of product measure and statement of Fubini's theorem. [6]

Chebyshev's inequality, tail estimates. Jensen's inequality. Completeness of L^p for $1 \leq p \leq \infty$. The Hölder and Minkowski inequalities, uniform integrability. [4]

L^2 as a Hilbert space. Orthogonal projection, relation with elementary conditional probability. Variance and covariance. Gaussian random variables, the multivariate normal distribution. [2]

The strong law of large numbers, proof for independent random variables with bounded fourth moments. Measure preserving transformations, Bernoulli shifts. Statements *and proofs* of maximal ergodic theorem and Birkhoff's almost everywhere ergodic theorem, proof of the strong law. [4]

The Fourier transform of a finite measure, characteristic functions, uniqueness and inversion. Weak convergence, statement of Lévy's convergence theorem for characteristic functions. The central limit theorem. [2]

Contents

0 Introduction

1 Measures

1.1 Measures

Definition (σ-algebra)**.** Let E be a set. A *σ-algebra* $\mathcal{E}$ on E is a collection of subsets of E such that

(i) $\emptyset \in \mathcal{E}$.

(ii) $A \in \mathcal{E}$ implies that $A^C = X \setminus A \in \mathcal{E}$.

(iii) For any sequence (A_n) in $\mathcal{E}$, we have that

$$\bigcup_n A_n \in \mathcal{E}.$$

The pair $(E, \mathcal{E})$ is called a *measurable space*.

Definition (Measure)**.** A *measure* on a measurable space $(E, \mathcal{E})$ is a function $\mu : \mathcal{E} \to [0, \infty]$ such that

(i) $\mu(\emptyset) = 0$

(ii) Countable additivity: For any disjoint sequence (A_n) in $\mathcal{E}$, then

$$\mu\left(\bigcup_n A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Definition (Generator of σ-algebra)**.** Let E be a set, and that $\mathcal{A} \subseteq P(E)$ be a collection of subsets of E. We define

$$\sigma(\mathcal{A}) = \{A \subseteq E : A \in \mathcal{E} \text{ for all } \sigma\text{-algebras } \mathcal{E} \text{ that contain } \mathcal{A}\}.$$

In other words $\sigma(\mathcal{A})$ is the smallest sigma algebra that contains $\mathcal{A}$. This is known as the sigma algebra *generated by* $\mathcal{A}$.

Definition (Borel σ-algebra)**.** Let $E = \mathbb{R}$, and $\mathcal{A} = \{U \subseteq \mathbb{R} : U \text{ is open}\}$. Then $\sigma(\mathcal{A})$ is known as the *Borel σ-algebra*, which is *not* the set of all subsets of $\mathbb{R}$.

We can equivalently define this by $\tilde{\mathcal{A}} = \{(a, b) : a < b, a, b \in \mathbb{Q}\}$. Then $\sigma(\tilde{\mathcal{A}})$ is also the Borel σ-algebra.

Definition (π-system)**.** Let $\mathcal{A}$ be a collection of subsets of E. Then $\mathcal{A}$ is called a *π-system* if

(i) $\emptyset \in A$

(ii) If $A, B \in \mathcal{A}$, then $A \cap B \in A$.

Definition (d-system)**.** Let $\mathcal{A}$ be a collection of subsets of E. Then $\mathcal{A}$ is called a *d-system* if

(i) $E \in \mathcal{A}$

(ii) If $A, B \in \mathcal{A}$ and $A \subseteq B$, then $B \setminus A \in \mathcal{A}$

(iii) For all increasing sequences (A_n) in $\mathcal{A}$, we have that $\bigcup_n A_n \in \mathcal{A}$.

Definition (Ring)**.** A collection of subsets $\mathcal{A}$ is a *ring* on E if $\emptyset \in A$ and for all $A, B \in \mathcal{A}$, we have $B \setminus A \in \mathcal{A}$ and $A \cup B \in \mathcal{A}$.

Definition (Algebra)**.** A collection of subsets $\mathcal{A}$ is an *algebra* on E if $\emptyset \in A$, and for all $A, B \in \mathcal{A}$, we have $A^C \in \mathcal{A}$ and $A \cup B \in \mathcal{A}$.

Definition (Set function)**.** Let $\mathcal{A}$ be a collection of subsets of E with $\emptyset \in \mathcal{A}$. A *set function* function $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(\emptyset) = 0$.

Definition (Increasing set function)**.** A set function is *increasing* if it has the property that for all $A, B \in \mathcal{A}$ with $A \subseteq B$, we have $\mu(A) \leq \mu(B)$.

Definition (Additive set function)**.** A set function is *additive* if whenever $A, B \in \mathcal{A}$ and $A \cup B \in \mathcal{A}$, $A \cap B = \emptyset$, then $\mu(A \cup B) = \mu(A) + \mu(B)$.

Definition (Countably additive set function)**.** A set function is *countably additive* if whenever A_n is a sequence of disjoint sets in $\mathcal{A}$ with $\cup A_n \in \mathcal{A}$, then

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n).$$

Definition (Countably subadditive set function)**.** A set function is *countably subadditive* if whenever (A_n) is a sequence of sets in $\mathcal{A}$ with $\bigcup_n A_n \in \mathcal{A}$, then

$$\mu\left(\bigcup_n A_n\right) \leq \sum_n \mu(A_n).$$

Definition (Borel σ-algebra)**.** Let E be a topological space. We define the *Borel σ-algebra* as

$$\mathcal{B}(E) = \sigma(\{U \subseteq \mathcal{E} : U \text{ is open}\}).$$

We write $\mathcal{B}$ for $\mathcal{B}(\mathbb{R})$.

Definition (Borel measure and Radon measure)**.** A measure μ on $(E, \mathcal{B}(E))$ is called a *Borel measure.* If $\mu(K) < \infty$ for all $K \subseteq E$ compact, then μ is a *Radon measure*.

Definition (Lebesgue measure)**.** The *Lebesgue measure* is the unique Borel measure μ on $\mathbb{R}$ with $\mu([a, b]) = b - a$.

Definition (σ-finite measure)**.** Let $(E, \mathcal{E})$ be a measurable space, and μ a measure. We say μ is *σ-finite* if there exists a sequence (E_n) in $\mathcal{E}$ such that $\bigcup_n E_n = E$ and $\mu(E_n) < \infty$ for all n.

1.2 Probability measures

Definition (Probability measure and probability space)**.** Let $(E, \mathcal{E})$ be a measure space with the property that $\mu(E) = 1$. Then we often call μ a *probability measure*, and $(E, \mathcal{E}, \mu)$ a *probability space*.

Definition (Sample space)**.** In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we often call Ω the *sample space*.

Definition (Events)**.** In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we often call the elements of $\mathcal{F}$ the *events*.

Definition (Probaiblity)**.** In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, if $A \in \mathcal{F}$, we often call $\mathbb{P}[A]$ the *probability* of the event A.

Definition (Independence of events)**.** A sequence of events (A_n) is said to be *independent* if

$$\mathbb{P}\left[\bigcap_{n\in J} A_n\right] = \prod_{n\in J} \mathbb{P}[A_n]$$

for all finite subsets $J \subseteq \mathbb{N}$.

Definition (Independence of σ-algebras)**.** A sequence of σ-algebras $(\mathcal{A}_n)$ with $\mathcal{A}_n \subseteq \mathcal{F}$ for all n is said to be independent if the following is true: If (A_n) is a sequence where $A_n \in \mathcal{A}_n$ for all n, them (A_n) is independent.

Definition (limsup and liminf)**.** Let (A_n) be a sequence of events. We define

$$\limsup A_n = \bigcap_n \bigcup_{m\geq n} A_m$$
$$\liminf A_n = \bigcup_n \bigcap_{m\geq n} A_m.$$

2 Measurable functions and random variables

2.1 Measurable functions

Definition (Measurable functions)**.** Let $(E, \mathcal{E})$ and $(G, \mathcal{G})$ be measure spaces. A map $f : E \to G$ is *measurable* if for every $A \in \mathcal{G}$, we have

$$f^{-1}(A) = \{x \in E : f(x) \in E\} \in \mathcal{E}.$$

If $(G, \mathcal{G}) = (\mathbb{R}, \mathcal{B})$, then we will just say that f is measurable on E.

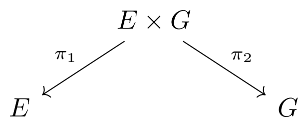
If $(G, \mathcal{G}) = ([0, \infty], \mathcal{B})$, then we will just say that f is *non-negative measurable.*

If E is a topological space and $\mathcal{E} = \mathcal{B}(E)$, then we call f a *Borel function*.

Definition (σ-algebra generated by functions)**.** Now suppose we have a set E, and a family of real-valued functions $\{f_i : i \in I\}$ on E. We then define

$$\sigma(f_i : i \in I) = \sigma(f_i^{-1}(A) : A \in \mathcal{B}, i \in I).$$

Definition (Product measurable space)**.** Let $(E, \mathcal{E})$ and $(G, \mathcal{G})$ be measure spaces. We define the *product measure space* as $E \times G$ whose σ-algebra is generated by the projections



.

More explicitly, the σ-algebra is given by

$$\mathcal{E} \otimes \mathcal{G} = \sigma(\{A \times B : A \in \mathcal{E}, B \in \mathcal{G}\}).$$

More generally, if $(E_i, \mathcal{E}_i)$ is a collection of measure spaces, the *product measure space* has underlying set $\prod_i E_i$, and the σ-algebra generated by the projection maps $\pi_i : \prod_j E_j \to E_i$.

Notation. We will write

$$f \wedge g = \min\{f, g\}, \quad f \vee g = \max\{f, g\}.$$

2.2 Constructing new measures

Definition (Image measure)**.** Let $(E, \mathcal{E})$ and $(G, \mathcal{G})$ be measure spaces. Suppose μ is a measure on $\mathcal{E}$ and $f : E \to G$ is a measurable function. We define the *image measure* $\nu = \mu \circ f^{-1}$ on G by

$$\nu(A) = \mu(f^{-1}(A)).$$

2.3 Random variables

Definition (Random variable)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $(E, \mathcal{E})$ a measurable space. Then an *E-valued random variable* is a measurable function $X : \Omega \to E$.

By default, we will assume the random variables are real.

Definition (Distribution/law)**.** Given a random variable $X : \Omega \to E$, the *distribution* or *law* of X is the image measure $\mu_x : \mathbb{P} \circ X^{-1}$. We usually write

$$\mathbb{P}(X \in A) = \mu_x(A) = \mathbb{P}(X^{-1}(A)).$$

Definition (Distribution function)**.** A *distribution function* is a non-decreasing, right continuous function $f : \mathbb{R} \to [0, 1]$ satisfying

$$F_X(x) \to \begin{cases} 0 & x \to -\infty \\ 1 & x \to +\infty \end{cases}.$$

Definition (Independence of random variables)**.** A family (X_n) of random variables is said to be *independent* if the family of σ-algebras $(\sigma(X_n))$ is independent.

2.4 Convergence of measurable functions

Definition (Convergence almost everywhere)**.** Suppose that $(E, \mathcal{E}, \mu)$ is a measure space. Suppose that $(f_n), f$ are measurable functions. We say $f_n \to f$ *almost everywhere (a.e.)* if

$$\mu(\{x \in E : f_n(x) \not\to f(x)\}) = 0.$$

If $(E, \mathcal{E}, \mu)$ is a probability space, this is called *almost sure convergence*.

Definition (Convergence in measure)**.** Suppose that $(E, \mathcal{E}, \mu)$ is a measure space. Suppose that $(f_n), f$ are measurable functions. We say $f_n \to f$ *in measure* if for each $\varepsilon > 0$, we have

$$\mu(\{x \in E : |f_n(x) - f(x)| \geq \varepsilon\}) \to 0 \text{ as } n \to \infty,$$

then we say that $f_n \to f$ *in measure*.

If $(E, \mathcal{E}, \mu)$ is a probability space, then this is called *convergence in probability*.

Definition (Convergence in distribution)**.** Let $(X_n), X$ be random variables with distribution functions F_{X_n} and F_X, then we say $X_n \to X$ *in distribution* if $F_{X_n}(x) \to F_X(x)$ for all $x \in \mathbb{R}$ at which F_X is continuous.

2.5 Tail events

Definition (Tail σ-algebra)**.** Let (X_n) be a sequence of random variables. We let

$$\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \cdots),$$

and

$$\mathcal{T} = \bigcap_n \mathcal{T}_n.$$

Then $\mathcal{T}$ is the *tail σ-algebra*.

3 Integration

3.1 Definition and basic properties

Definition (Simple function). A *simple function* is a measurable function that can be written as a finite non-negative linear combination of indicator functions of measurable sets, i.e.

$$f = \sum_{k=1}^{n} a_k \mathbf{1}_{A_k}$$

for some $A_k \in \mathcal{E}$ and $a_k \geq 0$.

Definition (Integral of simple function). The integral of a simple function

$$f = \sum_{k=1}^{n} a_k \mathbf{1}_{A_k}$$

is given by

$$\mu(f) = \sum_{k=1}^{n} a_k \mu(A_k).$$

Definition (Integral). Let f be a non-negative measurable function. We set

$$\mu(f) = \sup\{\mu(g) : g \leq f, g \text{ is simple}\}.$$

For arbitrary f, we write

$$f = f^+ - f^- = (f \vee 0) + (f \wedge 0).$$

We put $|f| = f^+ + f^-$. We say f is *integrable* if $\mu(|f|) < \infty$. In this case, set

$$\mu(f) = \mu(f^+) - \mu(f^-).$$

If only one of $\mu(f^+), \mu(f_-) < \infty$, then we can still make the above definition, and the result will be infinite.

3.2 Integrals and limits

3.3 New measures from old

Definition (Restriction of measure space). Let $(E, \mathcal{E}, \mu)$ be a measure space, and let $A \in \mathcal{E}$. The *restriction* of the measure space to A is $(A, \mathcal{E}_A, \mu_A)$, where

$$\mathcal{E}_A = \{B \in \mathcal{E} : B \subseteq A\},$$

and μ_A is the restriction of μ to $\mathcal{E}_A$, i.e.

$$\mu_A(B) = \mu(B)$$

for all $B \in \mathcal{E}_A$.

Definition (Pushforward/image of measure). Let $(E, \mathcal{E})$ and $(G, \mathcal{G})$ be measure spaces, and $f : E \to G$ a measurable function. If μ is a measure on $(E, \mathcal{E})$, then

$$\nu = \mu \circ f^{-1}$$

is a measure on $(G, \mathcal{G})$, known as the *pushforward* or *image* measure.

Definition (Density)**.** Let $(E, \mathcal{E}, \mu)$ be a measure space, and f be a non-negative measurable function. We define

$$\nu(A) = \mu(f\mathbf{1}_A).$$

Then ν is a measure on $(E, \mathcal{E})$.

Definition (Density)**.** Let X be a random variable. We say X has a density if its law μ_X has a density with respect to the Lebesgue measure. In other words, there exists f_X non-negative measurable so that

$$\mu_X(A) = \mathbb{P}[X \in A] = \int_A f_X(x)\;\mathrm{d}x.$$

In this case, for any non-negative measurable function, for any non-negative measurable g, we have that

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) f_X(x)\;\mathrm{d}x.$$

3.4 Integration and differentiation

3.5 Product measures and Fubini's theorem

Definition (Product σ-algebra)**.** Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be finite measure spaces. We let

$$\mathcal{A} = \{A_1 \times A_2 : A_2 \times \mathcal{E}_1, A_2 \times \mathcal{E}_2\}.$$

Then $\mathcal{A}$ is a π-system on $E_1 \times E_2$. The *product σ-algebra* is

$$\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2 = \sigma(\mathcal{A}).$$

4 Inequalities and L^p spaces

Definition (L^p spaces)**.** Let $(E, \mathcal{E}, \mu)$ be a measurable space. For $1 \leq p < \infty$, we define $L^p = L^p(E, \mathcal{E}, \mu)$ to be the set of all measurable functions f such that

$$\|f\|_p = \left(\int |f|^p \mathrm{d}\mu\right)^{1/p} < \infty.$$

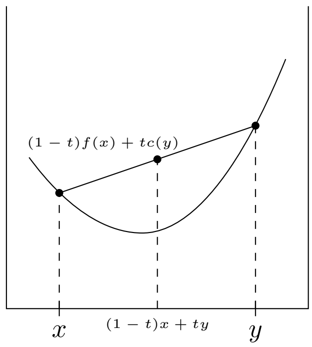
For $p = \infty$, we let $L^\infty = L^\infty(E, \mathcal{E}, \mu)$ to be the space of functions with

$$\|f\|_\infty = \inf\{\lambda \geq 0 : |f| \leq \lambda \text{ a.e.}\} < \infty.$$

4.1 Four inequalities

Definition (Convex function)**.** Let $I \subseteq \mathbb{R}$ be an interval. Then $c : I \to \mathbb{R}$ is convex if for any $t \in [0, 1]$ and $x, y \in I$, we have

$$c(tx + (1 - t)y) \leq tc(x) + (1 - t)c(y).$$



Definition (Conjugate)**.** Let $p, q \in [1, \infty]$. We say that they are *conjugate* if

$$\frac{1}{p} + \frac{1}{q} = 1,$$

where we take $1/\infty = 0$.

4.2 L^p spaces

Definition (Norm of vector space)**.** Let V be a vector space. A *norm* on V is a function $\| \cdot \| : V \to \mathbb{R}_{\geq 0}$ such that

(i) $\|u + v\| \leq \|u\| + \|v\|$ for all $U, v \in V$.

(ii) $\|\alpha v\| = |\alpha|\|v\|$ for all $v \in V$ and $\alpha \in \mathbb{R}$

(iii) $\|v\| = 0$ implies $v = 0$.

Definition (L^p spaces)**.** Let $(E, \mathcal{E}, \mu)$ be a measurable space. For $1 \leq p < \infty$, we define $L^p = L^p(E, \mathcal{E}, \mu)$ to be the set of all measurable functions f such that

$$\|f\|_p = \left(\int |f|^p \mathrm{d}\mu\right)^{1/p} < \infty.$$

For $p = \infty$, we let $L^\infty = L^\infty(E, \mathcal{E}, \mu)$ to be the space of functions with

$$\|f\|_\infty = \inf\{\lambda \geq 0 : |f| \leq \lambda \text{ a.e.}\} < \infty.$$

Definition ($\mathcal{L}^p$ space)**.** We define

$$\mathcal{L}^p = \{[f] : f \in L^p\},$$

where

$$[f] = \{g \in L^p : f - g = 0 \text{ a.e.}\}.$$

This is a normed vector space under the $\|\cdot\|_p$ norm.

Definition (Complete vector space/Banach spaces)**.** A normed vector space $(V, \|\cdot\|)$ is *complete* if every Cauchy sequence converges. In other words, if (v_n) is a sequence in V such that $\|v_n - v_m\| \to 0$ as $n, m \to \infty$, then there is some $v \in V$ such that $\|v_n - v\| \to 0$ as $n \to \infty$. A complete vector space is known as a *Banach space*.

4.3 Orthogonal projection in $\mathcal{L}^2$

Definition (Hilbert space)**.** A *Hilbert space* is a vector space with a complete inner product.

Definition (Orthogonal functions)**.** Two functions $f, g \in \mathcal{L}^2$ are *orthogonal* if

$$\langle f, g\rangle = 0,$$

Definition (Orthogonal complement)**.** Let $V \subseteq L^2$. We then set

$$V^\perp = \{f \in L^2 : \langle f, v\rangle = 0 \text{ for all } v \in V\}.$$

Definition (Closed subspace)**.** Let $V \subseteq L^2$. Then V is closed if whenever (f_n) is a sequence in V with $f_n \to f$, then there exists $v \in V$ with $v \sim f$.

Definition (Conditional expectation)**.** Suppose we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and (G_n) is a collection of pairwise disjoint events with $\bigcup_n G_n = \Omega$. We let

$$\mathcal{G} = \sigma(G_n : n \in \mathbb{N}).$$

The *conditional expectation* of X given $\mathcal{G}$ is the random variable

$$Y = \sum_{n=1}^\infty \mathbb{E}[X \mid G_n]\mathbf{1}_{G_n},$$

where

$$\mathbb{E}[X \mid G_n] = \frac{\mathbb{E}[X\mathbf{1}_{G_n}]}{\mathbb{P}[G_n]} \text{ for } \mathbb{P}[G_n] > 0.$$

4.4 Convergence in $L^1(\mathbb{P})$ and uniform integrability

Definition (Uniformly integrable)**.** Let $\mathcal{X}$ be a family of random variables. Define

$$I_{\mathcal{X}}(\delta) = \sup\{\mathbb{E}[|X|\mathbf{1}_A] : X \in \mathcal{X}, A \in \mathcal{F} \text{ with } \mathcal{P}[A] < \delta\}.$$

Then we say $\mathcal{X}$ is *uniformly integrable* if $\mathcal{X}$ is L^1-bounded (see below), and $I_{\mathcal{X}}(\delta) \to 0$ as $\delta \to 0$.

Definition (L^p-bounded)**.** Let $\mathcal{X}$ be a family of random variables. Then we say $\mathcal{X}$ is L^p-bounded if

$$\sup\{\|X\|_p : X \in \mathcal{X}\} < \infty.$$

5 Fourier transform

5.1 The Fourier transform

Definition (Fourier transform)**.** The *Fourier transform* $\hat{f} : \mathbb{R}^d \to \mathbb{C}$ of $f \in L^1(\mathbb{R}^d)$ is given by

$$\hat{f}(u) = \int_{\mathbb{R}^d} f(x) e^{i(u,x)}\, \mathrm{d}x,$$

where $u \in \mathbb{R}^d$ and (u, x) denotes the inner product, i.e.

$$(u, x) = u_1 x_1 + \cdots + u_d x_d.$$

Definition (Fourier transform of measure)**.** The Fourier transform of a *finite measure* μ on $\mathbb{R}^d$ is the function $\hat{\mu} : \mathbb{R}^d \to \mathbb{C}$ given by

$$\hat{\mu}(u) = \int_{\mathbb{R}^d} e^{i(u,x)} \mu(\mathrm{d}x).$$

Definition (Characteristic function)**.** Let X be a random variable. Then the *characteristic function* of X is the Fourier transform of its law, i.e.

$$\phi_X(u) = \mathbb{E}[e^{i(u,X)}] = \hat{\mu}_X(u),$$

where μ_X is the law of X.

5.2 Convolutions

Definition (Convolution of random variables)**.** Let μ, ν be probability measures. Their *convolution* $\mu * \nu$ is the law of $X + Y$, where X has law μ and Y has law ν, and X, Y are independent. Explicitly, we have

$$\begin{aligned} \mu * \nu(A) &= \mathbb{P}[X + Y \in \mathcal{A}] \\ &= \iint \mathbf{1}_A(x + y)\, \mu(\mathrm{d}x)\, \nu(\mathrm{d}y) \end{aligned}$$

Definition (Convolution of function with measure)**.** Let $f \in L^p$ and ν a probability measure. Then the *convolution* of f with μ is

$$f * \nu(x) = \int f(x - y)\, \nu(\mathrm{d}y) \in L^p.$$

5.3 Fourier inversion formula

Definition (Gaussian density)**.** The *Gaussian density* with variance t is

$$g_t(x) = \left(\frac{1}{2\pi t}\right)^{d/2} e^{-|x|^2/2t}.$$

This is equivalently the density of $\sqrt{t}Z$, where $Z = (Z_1, \cdots, Z_d)$ with $Z_i \sim N(0, 1)$ independent.

Definition (Gaussian convolution)**.** Let $f \in L^1$. Then a *Gaussian convolution* of f is a function of the form $f * g_t$.

5.4 Fourier transform in $\mathcal{L}^2$

5.5 Properties of characteristic functions

Definition (Weak convergence of measures)**.** Let $\mu, (\mu_n)$ be Borel probability measures. We say that $\mu_n \to \mu$ *weakly* if and only if $\mu_n(g) \to \mu(g)$ for all bounded continuous g.

Definition (Weak convergence of random variables)**.** Let $X, (X_n)$ be random variables. We say $X_n \to X$ weakly iff $\mu_{X_n} \to \mu_X$ weakly, iff $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$ for all bounded continuous g.

5.6 Gaussian random variables

Definition (Gaussian random variable)**.** Let X be a random variable on $\mathbb{R}$. This is said to be *Gaussian* if there exists $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ such that the density of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

A constant random variable $X = \mu$ corresponds to $\sigma = 0$. We say this has *mean* μ and *variance* σ^2.

When this happens, we write $X \sim N(\mu, \sigma^2)$.

Definition (Gaussian random variable)**.** Let X be a random variable. We say that X is a *Gaussian on* $\mathbb{R}^n$ if (u, X) is Gaussian on $\mathbb{R}$ for all $u \in \mathbb{R}^n$.

6 Ergodic theory

Definition (Invariant subset)**.** We say $A \in \mathcal{E}$ is invariant for Θ if $A = \Theta^{-1}(A)$.

Definition (Invariant function)**.** A measurable function f is invariant if $f = f \circ \Theta$.

Definition ($\mathcal{E}_\Theta$)**.** We write

$$\mathcal{E}_\Theta = \{A \in \mathcal{E} : A \text{ is invariant}\}.$$

Definition (Ergodic)**.** We say Θ is *ergodic* if $A \in \mathcal{E}_\Theta$ implies $\mu(A) = 0$ or $\mu(A^C) = 0$.

6.1 Ergodic theorems

7 Big theorems

7.1 The strong law of large numbers

7.2 Central limit theorem