Contents

Color code: highlight, key concepts and definitions, to be added.

Reference: PTE, Folland, Tao, Cinlar, Claodio Landim.

Chapter 1

Measure and Probability

1.1 Measure and Integration

1.1.1 Measure and Measurable Functions

First of all, why should we discuss measure theory. It's for a better integration theory than the usual Riemann integral. Because we need strong conditions for arguments like:

$$\lim_n \int f_n = \int \lim_n f_n$$

to hold for Riemann integral.

Integration provides a tool to analyse things we are interested in, for example, we can control $P(X > \varepsilon)$ by the integration/moment of X.

TBA: Relationship with Riemann Integral We'll talk about

1. Measure space, construction of measure, Lebesgue measures;
2. Integration
3. Convergence concepts
4. Differentiation

Let's start with an abstract measure space. Let Ω be a space. A σ-algebra $\mathcal{F}$ is a collection of subsets called *measurable sets* of Ω such that it's closed under complements and countable unions. Also the $\varnothing \in \mathcal{F}$. We can assign *measure* μ to members in $\mathcal{F}$. The reason we need to restrict attention from the power set to $\mathcal{F}$ is that there are examples that violate some properties we want a measure to have. μ satisfies:

$$\mu(\varnothing) = 0 \quad \text{and} \quad \mu\left(\sqcup_i^\infty A_i\right) = \sum_i \mu(A_i) \quad \text{for} \quad (A_i) \subset \mathcal{F}$$

The triplet $(\Omega, \mathcal{F}, \mu)$ is called a *measure space*.

A *measurable function* is $X : (\Omega, \mathcal{F}, \mu) \to (Y, \mathcal{H})$ such that for all $H \in \mathcal{H}$, $X^{-1}(H) \in \mathcal{F}$. Meaning that we can assign a measure to the subsets of $\mathcal{H}$ based on the measure μ and X. Given two measurable spaces, we can check whether a function X is measurable.

1.1.1 Lemma. *Suppose $\mathcal{H} = \sigma(\mathcal{G})$, if for all $B \in \mathcal{G}$, $X^{-1}(B) \in \mathcal{F}$, then X is measurable.*

Proof. Let $\mathcal{H}' = \left\{B \subset S \mid X^{-1}(B) \in \mathcal{F}\right\}$, then we can show $\mathcal{H}'$ is a σ-algebra because X^{-1} preserves complements and union. Then we have $\mathcal{G} \subset \sigma(\mathcal{G}) = \mathcal{H} \subset \mathcal{H}'$, meaning that all $H \in \mathcal{H}$ satisfies the condition. ♦

There are useful lemmas for checking a collection of subsets is indeed a σ-algebra.

We say $\mathcal{P}$ is a π-system, if $A, B \in \mathcal{P}$ implies $A \cap B \in \mathcal{P}$. $\mathcal{L}$ is a λ-system if 1. $\Omega \in \mathcal{L}$ 2. $A, B \in \mathcal{L}$ and $A \subset B$ will imply $A - B \in \mathcal{L}$ 3. If $A_n \in \mathcal{L}$, $A_n \uparrow A$, then $A \in \mathcal{L}$.

1.1.2 Dykin's $\pi - \lambda$. *If a p-system $\mathcal{P}$ is contained in a λ-system $\mathcal{D}$, then $\sigma(\mathcal{P}) \subset \mathcal{D}$.*

1.1.3 Monotone Class Theorem. *A collection of subsets $\mathcal{M}$ is a monotone class if $(A_n) \subset \mathcal{A}$*

1. *$A_n \uparrow A$ implies $A \in \mathcal{A}$.*
2. *$A_n \downarrow A$ implies $A \in \mathcal{A}$.*

The monotone class generated by an algebra $\mathcal{A}$, $\mathcal{M}(\mathcal{A}) = \sigma(\mathcal{A})$.

A very useful property of the topology on $\mathbb{R}$ is the following.

1.1.4 Open Sets on $\mathbb{R}$. *Every open set $U \subset \mathbb{R}$ is an at most countable union of disjoint open intervals.*

Proof. Equivalence argument: if $x, y \in U$, $x \sim y$ if there exists an open interval I_x containing x, y. $U/\!\sim$ is a set of disjoint open intervals and each contains a different rational q, so it's at most countable. ♦

1.1.5 Measurability. *For countable collections of random variables* $X_n, Y_n, X+Y, XY$ *and* $\sup X_n$, $\inf X_n$, $\limsup X_n$, $\liminf X_n$ *are measurable. If* $\lim_n X_n$ *exists, it's also measurable.*

PROOF. $+, *$ are continuous in the topological vector space $\mathbb{R}, \mathbb{C}$. For $\sup_n X_n$, let $t \in \mathbb{R}$, $\sup_n X_n \in (-\infty, t) \iff X_n < t$, for all t, hence $(\sup_n X_n)^{-1}(-\infty, t) = \cap_n X_n^{-1}(-\infty, t)$. ♦

Probability Space and Random Variables

A measure space is a *probability space* if $\mu = 1$, we write it as $(\Omega, \mathcal{F}, P)$. A random variable/vector is a measurable function on probability space $X : \Omega \to \mathbb{R}/\mathbb{R}^d/\mathbb{C}$ with the Borel σ-algebra, the σ-algebra generated by the open sets.

1.1.2 Construction of Measures

Motivation

Now we want to show that we can actually find measure on $\mathbb{R}$ and $\mathbb{R}^d$ that are the most useful spaces. We start from abstract construction method inspired by the following observation of measure on $\mathbb{R}$.

In order to have a natural measure λ on $\mathbb{R}$, we want the following properties:

1. $\lambda_0(a, b] = b - a$;
2. Measure of union of disjoint intervals should be the sum;(we can't have uncountable sum, hence restricted to countable union)
3. Apply to all intervals(we can't apply to the power set, hence restrict to the measurable sets).

Now consider the set $\mathcal{S}_{\mathbb{R}} = \{(a, b] : a \le b \in \mathbb{R}\}$, it's too small, we want to extend the natural measure funtion λ_0 to a larger set(we can extend to $\mathcal{B}_{\mathbb{R}}$, even to $\mathcal{L}_{\mathbb{R}}$). It can be shown that $\mathcal{B} = \sigma(\mathcal{S})$.

Construction of Abstract Measure

We follow the steps of extensions

$$(\mathcal{S}, \mu_0) \to (\mathcal{A}, \mu_1) \to (\mathcal{M}, \mu^*) \to (\mathcal{B}, \mu)$$

Semi-algebra $\mathcal{S}$ is collection of subsets of a space Ω that satisfies the following conditions: closed under intersection and the complements are union of finite disjoint sets in $\mathcal{S}$. Let μ_0 be a what I call *semi-pre-measure* be a set function on $\mathcal{S}$ such that if $S_N = \sqcup^N S_j, S_\infty = \sqcup^\infty S_j \in \mathcal{S}$, then

$$\mu_0(\varnothing) = 0; \quad \mu_0(S_N) = \sum^N \mu_0 S_j; \quad \mu_0(S_\infty) \leq \sum^\infty \mu_0 S_j$$

There is a unique extension of μ_0 to a *premeasure* on the *algebra* $\mathcal{A}$ generated by $\mathcal{S}$, that is the finite disjoint union of members of $\mathcal{S}$, satisfying

$$\mu_1(\varnothing) = 0; \quad \mu_1(\sqcup_j A_j) = \sum_j \mu_1(A_j)$$

if $\sqcup_j^\infty A_j \in \mathcal{A}$.

1.1.6 $(\mathcal{S}, \mu_0) \longrightarrow (\mathcal{A}, \mu_1)$. *We have that* $\mathcal{A} = \left\{\bigsqcup_j^N S_j : S_j \in \mathcal{S}\right\}$ *is an algebra. And* $\mu_1(A) = \sum_j^N \mu_0 S_j$ *is a premeasure.*

PROOF. $\mu_1(\varnothing) = 0$. Suppose $A = \bigsqcup^\infty A_i \in \mathcal{A}$, then there exists $T_l \in \mathcal{S} : l = 1, \ldots N_T$, such that $A = \sqcup_l T_l$, also for each $A_j = \sqcup_k^{N_j} S_{j,k}$. ♦

Let an *outer measure* be a function over the power set such that:

$$\mu^*(\varnothing) = 0 \quad \text{and} \quad \mu^*(\sqcup^\infty A_j) \leq \sum_j \mu^* A_j \quad \text{and} \quad \mu^*(A) \leq \mu^*(B) \quad \text{for} \quad A \subset B$$

Also let $\mathcal{M}$ be the *outer-measurable sets*, that is the collections of A such that

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \text{for all} \quad E \subset X$$

1.1.7 $(\mathcal{A}, \mu_1) \longrightarrow (\mathcal{M}, \mu^*)$ **Caratheodory's Theorem applied to premeasures.** *We can find the outer measure induced by* μ_1*:*

$$\mu_1^*(B) = \inf\left\{\sum_j^\infty \mu_1 A_j : B \subset \cup_j A_j\right\}$$

we show μ_1^* *is indeed an outer measure, Caratheodory states that for outer measures,* $\mathcal{M}$ *of* μ_1^* *is a* σ*-algebra and* $\mu^*|_{\mathcal{M}}$ *is complete.*

PROOF.

1. Monotone property is easy, for countable subadditivity, we use ε-room method, find a cover C_{jk} of A_j such that :
$$\mu_1^*(A_j) > \sum_k \mu_1^* C_{jk} - \varepsilon$$
2. To show that $\mathcal{M}$ is a σ-algebra we only need to show that for $E_j \in \mathcal{M}$ and any $B \subset \Omega$, we have:
$$\mu^*(B) \geq \mu^*(B \cap E) + \mu^*(B \cap E)$$
which can be seen by considering finite sum, taking out one E_j at a time and then take limit.
3. Show that the restriction is complete.

♦

REMARK. As a result, there exists an extension of a premeasure on $\mathcal{A}$ to a complete measure on $\sigma(\mathcal{A})$, and if the premeasure is σ-finite, then the extension is unique.

1.1.8 $(\mathcal{M}, \mu^*) \longrightarrow (\sigma(\mathcal{S}), \mu)$. *We can restrict μ^* to define $\mu = \mu^*|_{\sigma(\mathcal{S})}$, and the resulting measure is complete.*

PROOF. We show that $\sigma(\mathcal{S}) \subset \mathcal{M}$ and the restriction is complete. ♦

1.1.3 Borel and Lebesgue-Stietjes Measure on $\mathbb{R}$

We verify that for any nondecreasing right-continuous function $F : \mathbb{R} \longrightarrow \mathbb{R}$, the semi-algebra $\mathcal{S}_\mathbb{R}$ together with the semi-pre-measure defined by

$$\lambda_0(a, b] = F(b) - F(a)$$

satisfies the conditions set in the previous section. Let $\mathcal{A}_\mathbb{R}$ be the algebra generated by $\mathcal{S}_\mathbb{R}$ in turn will generate $\mathcal{B}_\mathbb{R}$. We can extend it to a something larger, a complete measure whose domain contains $\mathcal{B}_\mathbb{R}$. It will be called the *Lebesgue-Stietjes measure* and *Lebesgue measurable sets* $\mathcal{M}_\lambda$. The restriction on $\mathcal{B}$ is called *Borel measure*. It's the *completion* of $\mathcal{B}$.

A measure space is *complete* if for all $E \subset N$ where $\mu(N) = 0$, $E \subset \mathcal{F}$. A completion of of $\mathcal{F}$, denoted $\bar{\mathcal{F}}$ will be

$$\bar{\mathcal{F}} = \{E \cup F : F \in \mathcal{F}, E \subset N, \mu(N) = 0\}$$

Regularity of Lebesgue Measurable Sets

1.1.9 Regularity of $\mathcal{M}_\lambda$. *Lebesgue measurable sets are of simple form if you allow for a small error.*

1. $\lambda(A) = \inf\{\lambda U : A \subset U, U \text{ is open}\} = \sup\{\lambda K : K \subset A, K \text{ is compact}\}$
2. $A = G_\delta \setminus N_1 = F_\sigma \cup N_2$ *where* N *are null sets.*
3. *Littlewood's First Principle: Suppose* $\lambda(A) < \infty$ *then for all* $\varepsilon > 0$, *there exists* B *that is a finite union of open intervals such that* $A \triangle B < \varepsilon$.

PROOF. ♦

REMARK. As a result, $\mathcal{L} \subset \bar{\mathcal{B}}$.

1.1.4 Integration

Definition

Integration and measure are the two sides of same coin: integration is the extension of measure from the space of *indicator functions* $\mathbf{1}_A$ of measurable sets to a larger set of functions.

For *simple functions* $f_s = \sum_j^N \mathbf{1}_{A_j}$, we can define $\mu f_s = \sum_j^N \mu(A_j)$

Then for *nonnegative measurable* functions f, we can define

$$\mu f = \sup\{\mu f_s : f_s \leq f \quad \text{pointwise}\}$$

If $f = f^+ - f^-$ and the integrals of two nonnegative parts are not ∞, then we say f is *integrable* or in L^1 and $\mu f = \mu f^+ - \mu f^-$.

Characterisations of Integral

1.1.10 Properties of Integral. *1. Monotone $\mu f \leq \mu g$ if $f \leq g$ pointwise.*

2. Linearity for $a > 0$, $\mu(af) = a\mu f$, $\mu(f+g) = \mu f + \mu g$.

3. Monoton Convergence: if $0 \leq f_n \uparrow f$ a.e., then $\mu f_n \uparrow \mu f$.

PROOF. Folland proves by choosing simple function and a scaling factor α. Tao's proof is essentially the same. PTE proves via Fatou's lemma.

- Suppose $f_n \uparrow f$ a.e., if μf exists, we know μf_n must converge because it's a nondecreasing bounded sequence of numbers. Also $\mu f_n \leq \mu f$ by monotonicity.
- Suppose f_n are indicator functions $f_n = \mathbf{1}\{A_n\}$, then $f_n \uparrow f$ implies that A_n is an increasing sequence of sets and $f = \mathbf{1}\{\cup A_n\}$, it follows from countable additivity of measure(continuity from below).
- To show that $\mu f \leq \lim_n \mu f_n$, where $\mu f = \sup\{\mu g : g \text{ is simple}, g \leq f\}$, we only need to show that for all simple $g \leq f$, we have $\mu g \leq \lim \mu f_n$. We can use ε-room method, for any g simple, let $E_n = \{\omega : f_n(\omega) \geq (1-\varepsilon)g(\omega)\}$, $E_n \uparrow \Omega$, and we have $\int f_n \geq \int_{E_n}(1-\varepsilon)g$, taking lim , we have $\lim \int f_n \geq (1-\varepsilon)\int g$ for all $\varepsilon > 0$.
- For the last assertion we can simply show a simple function induces a new measure.

♦

REMARK. Let g be a simple function, $\mu_g(A) = \int_A g \,\mathrm{d}\mu$ defines a measure $\longrightarrow$ Monotone convergence $\longrightarrow$ for any $f \in \mathcal{F}^+$, μ_f defines a measure. Later we will show by Radon-Nikodym that L^1 functions are related to signed measures.

Regularity of Integrable Functions

1.1.11 Littlewood's Second Principle. *Measurable and integrable functions can be approximated by simple well-behaved functions.*

1.1.12 Lusin's Theorem.

REMARK. These three properties fully characterise the integral. That is given a functional that satisfies the three properties we can find a measure for which the function is the integral.

Integral Inequalities

1.1.13 Inequalities and Controls. *Useful inequalities and controls, suppose that $f \in L^1$;*

1. *Jensen's inequality: suppose we have a convex function $\Phi(x) : (a, b) \longrightarrow \mathbb{R}$, suppose $\mu(\Omega) = 1$, and $f \in L^1 : \Omega \longrightarrow (a, b)$ then $\mu(\Phi(f)) \geq \Phi\mu f$*
2. *Holder's inequality. Suppose $1 \leq p \leq \infty$ For $q : \frac{1}{p} + \frac{1}{q} = 1$ and f, g being measurable, we have $\|fg\|_1 \leq \|f\|_p \|g\|_q$ with equality iff $\alpha, \beta \neq 0$, $\alpha|f|^p = \beta|g|^q$ a.e.*
3. *Minkowski's inequality: $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.*
4. *Chebyshev's Inequality:*
5. *For nonnegative measurable functions, $\mu f < \infty \implies f < \infty$ a.e., $\mu f = 0$ implies $f = 0$ a.e.*
6. *Moment and tail behaviour, if μ is finite,*

$$\mu f^r < \infty \implies \sum x^{r-1}\mu(f > x) < \infty$$

PROOF. We prove the above inequalities.

1. First we show that if Φ is convex and $t_0 \in (a, b)$ then there exists β such that $F(t) - F(t_0) \geq \beta(t - t_0)$ for all $t \in (a, b)$. Second Φ is measurable, because for any t, $\Phi^{-1}(x : x < t)$ is an interval.
2. Suppose $\|f\|_p = 0$, then $f = 0$, *a.e.*, then LHS is also 0. If $\|f\|_p = \infty$ then it obviously is true. Then one proof is based on the lemma:

1.1.14 Lemma. *If $a, b \geq 0$ and $0 < \lambda < 1$ then*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1 - \lambda)b$$

with equality iff $a = b$.

♦

1.1.5 Product Space and Lebesgue Measure on $\mathbb{R}^d$

Product Measurable Space

Let $(Y_\alpha, \mathcal{H}_\alpha) : \alpha \in \mathcal{A}$ be a collection of measurable spaces, then we can define a *product σ-algebra* $\otimes_\alpha \mathcal{H}_\alpha = \sigma\{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{H}_\alpha, \alpha \in \mathcal{A}\}$ on the product space $\prod_\alpha \Omega_\alpha$,

1.1.15 **Proposition.** *Let $\Omega, \mathcal{F}$ and $(Y_\alpha, \mathcal{H}_\alpha)$ be measurable spaces, let $(Y, \mathcal{H}) = (\prod Y_\alpha, \otimes_\alpha \mathcal{H}_\alpha)$. $f : \Omega \to Y$ is $(\mathcal{F}, \mathcal{H})$-measurable if and only if $f_\alpha = \pi_\alpha f$ is $\mathcal{F}, \mathcal{H}_\alpha$-measurable for all α.*

1.1.16 **Proposition.** *Product σ-algebra is generally smaller than the σ-algebra on the product space itself. On separable metric space, they coincide.*

A measure space is a *probability space* if $\mu = 1$, we write it as $(\Omega, \mathcal{F}, P)$. A random variable/vector is a measurable function on probability space $X : \Omega \to \mathbb{R}/\mathbb{R}^d/\mathbb{C}$ with the Borel σ-algebra, the σ-algebra generated by the open sets. Now consider finite vector of random variables $X^d := (X_1, \ldots, X_d) : \Omega \to \mathbb{R}^d$. Then X^d is measurable iff all X_i are measurable. *Product measurable* and *product measure*.

Finite Product Measure

Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two σ-finite measure spaces. We can definte a *semi-pre-measure* on the *rectangles*: $E_1 \times E_2 \in \mathcal{F}_1 \times \mathcal{F}_2$:

$$\tilde{\mu}(E_1 \times E_2) = \mu_1(E_1)\mu_2(E_2)$$

we can show that the collection of rectangles $\mathcal{S} = \mathcal{F}_1 \times \mathcal{F}_2$ a semi-algebra on $\Omega_1 \times \Omega_2$: they are closed under intersection and $\Omega_1 \times \Omega_2 \in \mathcal{S}$, and $(E_1 \times E_2)^c$ is a finite disjoint union of members of $\mathcal{S}$.

The product σ-algebra on $\Omega_1 \times \Omega_2$ is $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{S})$, if we can show that $\tilde{\mu}$ is σ-additive, then we can uniquely extend it to a measure on $\mathcal{F}_1 \otimes \mathcal{F}_2$ by Caratheodory's theorem.

1.1.17 **Theorem.** *Under the assumptions above, we have a unique extension of $\tilde{\mu}$ to a measure μ on $(\mathcal{F}_1 \otimes \mathcal{F}_2)$.*

PROOF. 1. as a lemma we show that for any $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$, the *section* of A at any $x \in \Omega_1$,

$$A_x := \{y \in \Omega_2 : (x, y) \in A\} \in \mathcal{F}_2$$

with *monotone class argument*, defining $\mathcal{E} := \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2 : \forall x \in \Omega_1, A_x \in \mathcal{F}_2\}$.

2. Then we show μ is additive on $\mathcal{F}_1 \times \mathcal{F}_2$. Suppose (A_j) are disjoint rectangles and $A = \sqcup_j A_j \in \mathcal{F}_1 \times \mathcal{F}_2$. Let $A_j = E_j \times F_j$ and $A = E \times F$.

   We can show that $A_x = F\mathbf{1}_E(x)$ is also equal to $\sum_j F_j \mathbf{1}_{E_j}(x)$ and since it's $\mathcal{F}_2$-measurable, we have $\mu_2(F)\mathbf{1}_E(x) = \sum_j \mu_2(F_j)\mathbf{1}_{E_j}(x)$.

   This as a function of x is $\mathcal{F}_1$-measurable, because it's a simple function. Taking integral with restpect to μ_1 will give the result.

   Then we show $\tilde{\mu}$ is also σ-additive by replacing the finite sum with infinite sum and use monotone convergence.

♦

Countable Product Measure

Let $(\Omega_j, \mathcal{F}_j, P_j)_{j=1}^{\infty}$ be a countable collection of finite-measure spaces, we fix $P_j(\Omega_j) = 1$. We want to construct a measure on the product σ-algebra $\otimes_j \mathcal{F}_j$ based on the set function $\tilde{\mu}$ defined on the seim-algebra of *cylinder sets* $A \in \mathcal{C}$ such that for some n,

$$A = E_1 \times E_2 \times \cdots \times E_n \times \Omega_{n+1} \times \dots \quad \text{where} \quad E_k \in \mathcal{F}_k$$

and $\tilde{\mu} : \mathcal{C} \to \mathbb{R}_+ \cup \{\infty\}$,

$$\tilde{\mu}(A) = \prod_{j=1}^{n} \mu_j(E_j)$$

1.1.18 Countable Product Measure. *We can uniquely extend $\tilde{\mu}$ to a measure on $\otimes \mathcal{F}_j$.*

PROOF. We need to show that

1. $\mathcal{C}$ is a semi-algebra.

2. $\tilde{\mu}$

♦

Fubini-Tonelli

1.1.19 Fubini-Tonelli. *Suppose we have a product space $\Omega_1 \times \Omega_2$, let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a measurable function with respect to $(\mathcal{F}_1 \otimes \mathcal{F}_2, \mathcal{B}_{\mathbb{R}})$. If Ω_1, Ω_2 are σ-finite under μ_1, μ_2, then*

Tonelli *If $f \geq 0$,*

1.1.6 Convergence Concepts

1.1.20 Uniqueness of Limits. *For different convergence modes, they agree on the limit, if exists, a.e.*

1.1.21 Fatou's Lemma. *Suppose f_n is measurable on Ω, then we have the following inequality:*

$$\int \liminf f_n \leq \liminf \int f_n$$

Proof. Based on Monotone Convergence. Because for each fixed n and $m \geq n$, $\inf_{m \geq n} f_m \leq f_m$, we have $\int \inf_{m\geq n} f_m \leq \int f_m$, then $\int \inf_{m \geq n} f_m \leq \inf_{m\geq n} \int f_m$. The integrand on the left is an increasing sequence of functions, we apply MCT and get the result. ♦

1.1.22 Dominated Convergence. *Suppose we have measurable $f_n \to f$ a.e., and a function $g \in L^1$ such that $|f_n| < g$ a.e. for all n, then $\lim \int f_n = \int \lim f_n$.*

Proof. Because $|f_n| < g$,we know that $|f| < g$ as well and hence integrable. Apply Fatou's Lemma twice. $|f_n - g + g - f| \leq |f_n - g| + |g - f|$. We know that $\left|\int f_n - f\right| \leq \int |f_n - f|$, which is integrable because ♦

1.1.23 Integration to the limit. *Suppose $f_n \to f$, there exists $g(x)$ such that*

1.1.24 Defect version of Fatou's Lemma. *From Tao Exercise 1.4.48, which I find interesting. $f_n \to f$ where f_n, f are all integrable. Show that*

$$\int f_n - \int f - \|f_n - f\|_{L^1} \to 0$$

Proof. Apply dominated convergence to $\min(f_n, f)$. ♦

Uniform integrability is a substitute for *dominance condition*. A sequence of $f_n \in L^1$ is uniformly integrable if

1. $\sup_n \|f_n\|_{L^1} < \infty$;
2. $\sup_n \|f_n \mathbf{1}(|f_n| > M)\| \to 0$ as $M \to \infty$;
3. $\sup_n \|f_n \mathbf{1}(|f_n| \le \delta)\| \to 0$ as $\delta \to 0$.

For functions $f_n : E \to \mathbb{R}$ there are a lot of ways to define $f_n \to f$. $f_n \to f$ *pointwise* if for each $x \in E$, $f_n(x)$ as a sequence of number converges to $f(x)$. $f_n \rightrightarrows f$ if for all x, the convergence rate is bounded from below.

With E equipped with a measure, we can have a relaxation of *uniform convergence*. $f_n \to f$ *almost uniformly* if for any $\varepsilon > 0$, there exists a $N_\varepsilon \subset \mathcal{E}$ with measure $< \varepsilon$ such that on $E \setminus N_\varepsilon$, $f_n \rightrightarrows f$. We have the following

We introduce *convergence in measure* and *convergence in distribution*. $f_n \to f$ in measure if $\mu(\|f_n - f(x)\| > \varepsilon) \to 0$ for any $\varepsilon > 0$.

Convergence in distribution has a *Skorokhod's representation* as independent random variables with corresponding distributions.

We relate the convergence of sets to the sum of measures.

1.1.25 **Borel-Cantelli Lemma.** *Let $E_1, E_2, \ldots, E_3$ be a sequence of $\mathcal{B}$-measurable sets such that $\sum_n \mu(E_n) < \infty$, show that $\mu(\limsup E_n) = 0$.*

Proof. $\sum \mu(E_n) < \infty$, meaning that $\sum_{n>m} \mu E_n \to 0$ as $m \to \infty$. Then for any $\varepsilon > 0$ there exists m such that $\mu(\limsup E_n) < \mu\left(\sup_{n>m} E_n\right) < \mu(\sum_{n>m} E_n) < \varepsilon$. ♦

Remark. Borel-Cantelli is useful for stong law of large numbers which turns the statement of $\mu\{x : \lim |f_n - f|(x) > \varepsilon\} = 0$ to $\sum_n \mu(x : |f_n - f| > \varepsilon) < \infty$.

Borel-Cantelli's Lemmas

1.1.26 **Borel-Cantelli's Lemma.** *If $\sum P(A_n) < \infty$, then $P(\limsup A_n) = P(A_n$ i.o.$) = 0$*

As a result, we can upgrade convergence in probability to convergence a.s.

1.1.27 **Theorem.** *$X_n \to X$ in probability iff for every subsequence of X_n, there exists a further subsequence that converges to X almost surely.*

PROOF. If $X_n \to X$ in probability, then $P(|X_n - X| \geq \varepsilon) \to 0$, for any subsequence $X_{n(k)}$ of X_n, we can select a further subsequence $X_{n(k(m))}$ such that $P(|X_n - X| \geq \varepsilon) \leq \frac{1}{m^2}$. By Borel-Cantelli, we know $X_{n(k(m))} \to X$ a.s.

For the other direction, it's easy to show that $X_n \to X$ a.s. implies $X_n \to X$ in probability, then use thinning in topological space method, Method C.2.2. Because convergence in probability can be put in a metric. ♦

REMARK. Notice that convergence almost surely is not a convergence in a topology.

1.1.28 Egoroff's Theorem. *Suppose $f_n \to f$ a.e. on a finite measure space, then $f_n \to f$ almost uniformly.*

PROOF. Suppose $E_{m,n} := \left\{x : |f_n(x) - f(x)| > \frac{1}{m}\right\}$. We know for fixed m, $E_{m,n} \to \varnothing$ a.e. The set $E_m(k) := \left\{x : \sup_{n>k} |f_n(x) - f(x)| > \frac{1}{m}\right\}$ ♦

1.1.7 Signed Measure

Definition and Hahn-Jordan

We relax the assumption that measure takes only non-negative values. Let $\Omega, \mathcal{F}$ be a measurable space, a *signed measure* is

$$\nu : \mathcal{F} \to [-\infty, +\infty]$$

such that

$$\nu(\varnothing) = 0; \quad \nu\left(\bigsqcup_{j}^{\infty} E_j\right) = \sum_j \nu(E_j)$$

where we always require the infinite sum is well-defined.

1.1.29 Hahn-Jordan Decomposition. *Let ν be σ-finite, then there exist P, N such that $P = N^c$ and for all $E \subset P, F \subset N, \nu(E) \geq 0$ and $\nu(F) \leq 0$.*

Radon-Nikodym

1.2 Independence

1.2.1 Distribution and Density

Distribution is the probability induced by a random variable X, such that $\mu(A) = P(X \in A)$. *Density* is defined via some methods(To be added). It's used to change measure, for example,

$$\int g(X)\,\mathrm{d}P = \int g(x)f(x)\,\mathrm{d}x$$

the latter being the Lebesgue integral over $\mathbb{R}$ to be defined later.

1.2.2 Independence

Definition

Independence is the first concept that has nontrivial meaning in probability other than inheriting measure theoretical names.

The independence is defined via the σ-algebras. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A collection of sub-σ-algebras $(\mathcal{F}_\alpha)$ are independent if for all finite subcollection, for all $F_1, F_2, \ldots, F_n \in \mathcal{F}_1, \mathcal{F}_2, \ldots \mathcal{F}_n$

$$P(F_1F_2) = P(F_1)P(F_2)\ldots P(F_n)$$

Can we extend this concept? For countable collection of sub-σ-algebras, $(\mathcal{F}_n)$, it's *independency* if for any finite collection $F_j \in \mathcal{F}_j$ we have

$$P\left(\bigcap_j^J F_j\right) = \prod_j P(F_j)$$

What about uncountable, let's assume it can be defined in the same way.

There are several questions: can we find independency, how can we show that two σ-algebras are independent.

Sufficient Condition for Independence

Let $\mathcal{P}_1, \ldots, \mathcal{P}_n$ be collections of measurable subsets, we say they are *independent* if for all $A_k \in \mathcal{P}_k$

$$P(\cap A_k) = \prod_k^n P(A_k)$$

1.2.1 **π-system determines independence.** *If $\mathcal{P}_k$ are independent π-systems, then $\sigma(\mathcal{P}_k)$ are independent.*

PROOF. Uses $\pi - \lambda$ theorem. ♦

The use of independence is that if $\sigma(X_n)$ are independent, then we have for any measurable functions of X_n, then we have

$$\int f_1(X_1) f_2(X_2) \ldots f_n(X_n) \, \mathrm{d}P = \int f_1(X_1) \, \mathrm{d}P \cdots \int f_n(X_n) \, \mathrm{d}P$$

This is related to the Fubini-Tonelli theorem in that we know the joint distribution will be the product measure of each distribution.

The Kolmogorov's Extension Theorem asserts that we can find infinite sequence of independent random variables that are consistent on **nice** spaces.

For a sequence of sub-σ-algebras $\mathcal{F}_n \subset \mathcal{F}$, let $\mathcal{F}'_m = \sigma(\cup_{n \geq m} \mathcal{F}_n)$ the *tail σ-algebra* $\mathcal{T}$ is defined as $\cap_m^\infty \mathcal{F}'_m$. It's the σ-algebra that contains the events about the limits.

For independent σ-algebras, the tail σ-algebra is simple, by the Kolmogorov's 0-1 law.

1.2.2 **Kolmogorov's 0-1 Law.**

1.2.3 Law of Large Numbers

Weak Law of Large Numbers

Truncation or *equivalent sequence* will be important for our purpose because it guarantees the existence of moments. Let $X_1, \ldots, X_n$ be a sequence of random variables. Let $Y_n = X_n \mathbf{1}_{X_n \leq M}$. Then Chebyshev's inequlaity will be used.

1.2.3 **Weak Law of Large Numbers.** *Suppose X_n are a sequence of random variables, satisfying that X_n*

1. $\sum_k P(X_n \geq k) < \infty$;

2.

<u>**Strong Law of Large Numbers**</u>

1.2.4 Convergence of Random Series

The *Kronecker's Lemma* provides a bridge between convergence of series and the strong law of large numbers.

1.2.4 **Kronecker's Lemma.** *If* $\sum \frac{a_n}{b_n} < \infty$ *and* $b_n \uparrow \infty$, *then* $\frac{1}{b_n} \sum a_n \to 0$ *a.s.*

1.2.5 Central Limit Theorem

Chapter 2

Stochastic Process over Time

Will be based on Dexter's Notes, Durrett and Probability II.

2.1 Conditional Expectation

2.2 Random Walk and Stopping Time

2.3 Conditional Expectation and Martingale

Conditional expectation is very important, it's a way to capture the dependence among random variables.

We begin with the definition of *filtration*. A filtration is a collection of sub-σ-algebra of $(\Omega, \mathcal{F}, P)$, such that

$$\mathcal{F}_n \uparrow \mathcal{F}$$

It represents the accumulation of information over time.

2.4 Markov Chain

2.5 Ergodic Theory

Ergodicity is about the intrinsic law. Let's consider discrete time, let T be a *measure-preserving* transformation on $(\Omega, \mathcal{F}, P)$, $T : \omega \mapsto T(\omega)$ such that $\forall A \in \mathcal{F}$, $P(T(A)) = P(A)$.

2.5.1 von Neumann Mean Ergodic Theorem. *Let $f \in L^p(\Omega, \mathcal{F}, P)$, if T is a measure preserving trasformation, then as $n \to \infty$*

$$\frac{1}{n+1}\sum_{i=0}^{n} f \circ T^i \to_{L_p} E(f \mid \mathcal{I})$$

If T is ergodic, then $\mathcal{I}$ is trivial and the conditional expectation is equal to the unconditional expectation almost surely.

Proof.
1. von Neumann proved the convergence as a property in Hilbert space and projections.
2. L^2 is a Hilbert space and the projection is conditional expecation.
3. We can extend to L^p because we have a finite measure space, L^∞ is dense subset of L^2.

♦

2.5.2 Birkhoff Individual Ergodic Theorem. *The convergence holds also almost surely.*

Appendix A

Fourier and Complex Analysis, Characteristic Functions

Appendix B

Functional Analysis

B.1 L^p Spaces

Let $f : (\Omega, \mathcal{F}, \mu) \to \mathbb{C}$ be a $\mathcal{F}$-measurable function, let $[f] := \{g \in \mathcal{F}, g = f, a.e.\}$, for $1 \leq p < \infty$, define $\|g\|_p = \left(\int_\Omega |g|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} < \infty$

$$L^p(\Omega, \mathcal{F}, \mu) := \left\{[g] : \|g\|_p < \infty\right\}$$

We define $\|f\|_\infty = \operatorname{esssup} |f| = \inf \{c : \mu\{|f| > c\} = 0\}$

That $\|f\|_p$ is indeed a norm is given by the Minkowski's inequality.

B.2 Dual Space of L^p

For any *bounded linear functional* $T : L^p \to \mathbb{F}$, there exists $g \in L^q$, where q is the conjugate, such that

$$T(f) = \int fg \,\mathrm{d}\mu$$

and $\|g\|_q = \|T\|$.

Appendix C

Methods

This will collect the methods we used in the probability book.

C.1 Property Determining Class

C.1.1 **Measurability Determining.**

C.1.2 **Monotone Class Theorem.**

C.1.3 $\pi - \lambda$ **Theorem.**

C.1.4 **Independence determined by** π**-systems.**

C.1.5 **Weak Convergence.**

C.2 Thinning and Truncating

The purpose is to have a sequence that has better property than the original sequence.

C.2.1 **Cauchy's Condensation Test.** *Let* $(f(n))$ *be a sequence of real numbers, then* $\sum f(n)$ *converges iff* $\sum 2^n f(2^n)$ *converges.*

Proof. ♦

C.2.2 **Thinning in Convergence in Topological Space.** *Let* x_n *be a sequence in a topological space* $\mathcal{X}$*, then* $x_n \longrightarrow x$ *in that topology, iff for all subseqeunce of* x_n *there exists a further subsequence that converges to* x*.*

PROOF. $\Longrightarrow$ is obvious. For the other direction, suppose x_n is not converging, then there exists an open neighbourhood $\mathcal{N}$ of x such that for all N, there exists $n > N$, $x_n \notin \mathcal{N}$, which will constitute a subsequence that has no further subsequence that converges. ♦

C.2.3 Thinning and Control. *Let $c_n \uparrow \infty$ be a sequence of constants, then $X_n/c_n \to 1$ if we can find a subsequence that converges and $c_{n+1}/c_n \to 1$.*

C.2.4 Truncating. *Let $Y_n = X_n \mathbf{1}(X_n \leq M)$, then Y_n and X_n are equivalent iff $P(Y_n \neq X_n i.o.) = 0$.*

Appendix D

Fun Facts

D.1 Convexity

Appendix E

Problems