Summary: Introduction to Probability

Daniele Cambria

2024

Contents

Chapter 1

Introduction: Part I

1.1 Sets

1.2 The principle of induction

1.3 Order structure of the real numbers

Exercise 1.1 (1.11 TOOL)**.** Let A be a set with n elements. Show that

1. the number of permutations of the elements from A is $n!$;
2. for any $0 \leq k \leq n$, the number of subsets of A having k elements if given by

$$\frac{n!}{(n-k)!k!}.$$

Chapter 2

Introduction: Part II

2.1 Functions

Definition 2.1. Let $f : A \to B$ be a function.
Surjective: if $f(A) = B$.
Injective: if $a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$.
Bijective: if it is surjective and injective.

Proposition 2.1. Let $f : I \to \mathbb{R}$, $I \subset \mathbb{R}$ be a strictly monotonic function. Then, $f : I \to f(I)$ is a bijection. Further, if f is strictly increasing (resp. strictly decreasing) on I, then an inverse of f is strictly increasing (resp. strictly decreasing) on $f(I)$.

Proposition 2.2. let $f : A \to B$ be a function. Let $B_* \subset B$. Then,

(a) $f^{-1}(B_c^*) = f^{-1}(B_*)^c$.

Let I and J be some sets and $A_i \subset A, i \in I$, and $B_j \subset B, j \in J$, be a collection of sets from A and B, respectively. Then,

(b) TODO

(c) TODO

(d) TODO

Proposition 2.3. TODO prop 2.12

2.2 Cardinality of Sets

2.3 Euclidean distance

Chapter 3

Introduction: Part III

3.1 Real valued sequences

Chapter 4

Measurable sets: Part I

4.1 Measurable spaces

Definition 4.1 (σ-field)**.** Let Ω be a nonempty set. A family of subsets $\mathcal{F}$ of Ω is called a σ-field on Ω if the following three itmes are statisfied:

(i) $\Omega \in \mathcal{F}$;

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;

(iii) if $\{A_i : i \in \mathbb{N}\}$ is a collection of sets s.t. $A_i \in \mathcal{F}$ for any $i \in \mathbb{N}$, then $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

Definition 4.2. 4.2 TODO

Definition 4.3 (Measurable space)**.** let $\Omega \neq \emptyset$ and $\mathcal{F}$ be a σ-field on Ω. The pair $(\Omega, \mathcal{F}$ is referred to as a measurable space. if $A \in \mathcal{F}$, then A is said to be measurable. if $A \subset \mathcal{F}$ and $\mathcal{A}$ is a σ-field on Ω, $\mathcal{A}$ is referred to as a sub-σ-field on Ω.

Chapter 5

Measurable sets: Part II

5.1 Measure spaces

Definition 5.1 (Measure on $\mathcal{F}$)**.** TODO

5.2 Semirings

Chapter 6

Measurable sets: Part III

6.1 Measure extensions

Proposition 6.1. Let $(a, b]$, $a < b \in \mathbb{R}$, be any left-open interval. Let I be countable and $(a_i, b_i]$, $i \in I$, be s.t., $(a, b] \subset \bigcup_{i \in I}(a_i, b_i]$, then

$$b - a \leq \sum_{i \in I}(b_i - a_i). \tag{10}$$

Proposition 6.2. Let $(a, b]$, $a < b \in \mathbb{R}$, be any left-open interval. let I be countable and $\{(a_i, b_i] : i \in I\}$ be a disjoint collection of left-open intervals s.t. $\bigcup_{i \in I}(a_i, b_i] \subset (a, b]$. Then

$$\sum_{i \in I}(b_i - a_i) \leq b - a.$$

Definition 6.1. Let $\Omega \neq \emptyset$ be a set and $\mathcal{A}$ be a collection of subsets from Ω. Let $A \in \mathcal{P}(\Omega)$ be any subset of Ω. A collection $\{U_i : i \in I\}$ is said to be a covering of A by sets from $\mathcal{A}$ if:

(i) $\{U_i : i \in I\} \subset \mathcal{A}$ (Set membership condition)

NOTE that (i) means $U_i \subset \mathcal{A}$ $\forall i \in I$, not $\bigcup_{i \in I} U_i \subset \mathcal{A}$.

(ii) $A \subset \bigcup_{i \in I} U_i$ (Covering condition)

A covering $\{\bigcup_i : i \in I\}$ of A by sets from $\mathcal{A}$ is referred as countable (resp. finite) if I is countable (resp. finite). We write $C_{\mathcal{A}}(A)$ for the set which contains all the countable covering of A by sets from $\mathcal{A}$, i.e.,

$$C_{\mathcal{A}}(A) = \{\xi : \xi \text{ is a countable covering of } A \text{ by sets from } \mathcal{A}\}.$$

Why do we say $A \in \mathcal{P}(\Omega)$ instead of $A \in \Omega$? When we use the notation $A \in \mathcal{P}(\Omega)$, it signifies that A is a subset of Ω, not an element of Ω. The power set $\mathcal{P}(\Omega)$ represents all possible subsets of Ω, including Ω itself, any subset of it, or even an empty set. Using $A \in \Omega$ would incorrectly imply that A is an individual element of Ω, which does not align with the context of covering subsets with subsets.

My Example 6.1 (Finite Covering)**.** Let $\Omega = \{1, 2, 3, 4, 5\}$, and let $\mathcal{A}$ be a collection of subests of Ω, such as $\mathcal{A} = \{\{1\}, \{2, 3\}, \{3, 5\}\}$, if we take $A = \{1, 2, 3\}$, a finite covering of A by sets from $\mathcal{A}$ could be $\{\{1\}, \{2, 3\}\}$. This covering is finite, as I can be $\{1, 2\}$, which is finite. The 2 conditions both hold. Each U_i is a subset of $\mathcal{A}$, and A is covered by the union of U_i. In this case, the possible countable coverings of A that can be formed using subsets of $\mathcal{A}$ are restricted to the one already provided. Therefore, $C_{\mathcal{A}}(A) = \{\{1\}, \{2, 3\}\}$

Important from Example 6.1 (Script) Let $\Omega = \mathbb{R}$ and $\mathcal{R} = \{A : A = (a, b], a, b \in \mathbb{R}\} \cup \{\emptyset\}$. We define the function $\ell : \mathcal{R} \to [0, \infty)$ s.t.

$$\ell(U) = \begin{cases} b - a, & \text{if } U = (a, b], \\ 0, & \text{if } U = \emptyset. \end{cases}$$

Given $A \in \mathcal{P}(\mathbb{R})$, we also define the function $v_\ell(\xi) : \mathcal{R} \to \mathbb{R}^+$, where $\xi \in C_{\mathcal{R}}(A)$ s.t.

$$v_\ell(\xi) = \sum_{U \in \xi} \ell(U).$$

We also show that

$$\inf\{v_\ell(\xi) : \xi \in C_{\mathcal{R}}((a, b])\} = \inf_{\xi \in C_{\mathcal{R}}((a,b])} v_\ell(\xi) = b - a, \tag{11}$$

i.e., $b-a$ is a lower bound for the values of $v_\ell(\xi)$, $\xi \in C_{\mathcal{R}}((a, b])$. We also saw that there exists $\xi \in C_{\mathcal{R}}((a, b])$ s.t. $b - a = v_\ell(\xi)$. Hence, the latter infimum is a minimum (Proposition 6.3).

Proposition 6.3. Given any left open interval $(a, b]$, $min_{\xi \in C_{\mathcal{R}}((a,b])} v_\ell(\xi) = b - a$

Define ℓ^* We build on the latter result and define the function

$$\ell^* = \inf_{\xi \in C_{\mathcal{R}}(A)} v_\ell(\xi), \quad A \in \mathcal{P}(\mathbb{R}).$$

Note, we know that if $A \in \mathcal{R}$, then $\ell^*(A) = b - a$

Chapter 7

Measurable functions

7.1 The concept of measurable functions

Definition 7.1 (Measurable function)**.** Let $(\Omega, \mathcal{F})$ and $(\Omega^*, \mathcal{F}^*)$ be two measurable spaces (cf. Definition 4.3). A function $f : \Omega \to \Omega^*$ is said to be measurable $\mathcal{F}/\mathcal{F}^*$ if for any $A^* \in \mathcal{F}^*, f^{-1}(A^*) \in \mathcal{F}$.

Proposition 7.1 (Measurable function)**.** let $(\Omega, \mathcal{F})$ and $(\Omega^*, \mathcal{F}^*)$ be two measurable spaces and $f : \Omega \to \Omega^*$ be a function. Suppose that $\mathcal{F}^* = \sigma(\mathcal{G})$ and for any $G \in \mathcal{G}$, $f^{-1}(G) \in \mathcal{F}$. Then, f is $\mathcal{F}/\mathcal{F}^*$ measurable.

Definition 7.2 (Borel function)**.** A function $f : \mathbb{R}^m \to \mathbb{R}^k$ is called Borel function if it is measurable $\mathfrak{B}(\mathbb{R}^m)/\mathfrak{B}(\mathbb{R}^k)$.

Proposition 7.2 (Continuous functions and Borel functions)**.** Any continuous function $f : \mathbb{R}^m \to \mathbb{R}^k$ is a Borel function.

Proposition 7.3 ($\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}$ be a real-valued function. Suppose that $\{\omega \in \Omega : f(\omega) \le x\} \in \mathcal{F}$ for any $x \in \mathbb{R}$, then f is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. In other words: f is a measurable function if the pre-image of any interval $(-\infty, x]$ under f is a measurable set in $\mathcal{F}$, or $f^{-1}((-\infty, x]) \in \mathcal{F}$. since $\mathfrak{B}(\mathbb{R}) = \sigma(\{(-\infty, x] : x \in \mathbb{R}\})$, we also clearly see the proof (cf. Proposition 7.1).

Thinking about $f^{-1}((-\infty, x))$ If $B \in \mathfrak{B}(\mathbb{R})$, then, $f^{-1}(B) = \{\omega \in \Omega : f(\omega) \in B\}$ Is the same as saying, $f^{-1}((-\infty, x)) = \{\omega \in \Omega : f(\omega) \le x\}$. $f^{-1}(B)$ will return ALL of the values $\omega \in \Omega$ for which $f(\omega) \in B$. See My Example 7.1 for further intuition.

Define $\mathbb{1}_A(\omega)$ TODO

Example 7.1 (Simple measurable function)**.** Let $\Omega = \{h, t\}$ and $\mathcal{F} = \mathcal{P}(\{h, t\}) = \{\emptyset, \{h\}, \{t\}, \{h, t\}\}$. Then, $\{h\} \in \mathcal{P}(\{h, t\})$. Thus

$$f(\omega) = \begin{cases} 1, & \text{if } \omega = h, \\ 0, & \text{if } \omega = t, \end{cases}$$

is $\mathcal{P}(\{h, t\})/\mathfrak{B}(\mathbb{R})$ measurable. In order for f to be $\mathcal{P}(\{h, t\})/\mathfrak{B}(\mathbb{R})$ measurable, the pre-image of every Borel set in $\mathbb{R}$ under f must be an element of $\mathcal{F}$. For any $x \in \mathbb{R}$, $f^{-1}((-\infty, x])$ will either be $\emptyset, \{h\}$, or $\{t\} \in \mathcal{F}$.

Proposition 7.4 ($\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}^k$, i.e.,

$$f(\omega) = (f_1(\omega), \ldots, f_k(\omega)).$$

Then, f is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable if and only if for any $i = 1, \ldots, k, f_i : \Omega \to \mathbb{R}$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Proposition 7.5 (Composite measurable function)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i : \Omega \to \mathbb{R}, i = 1, \ldots, k,$ be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Suppose that $g : \mathbb{R}^k \to \mathbb{R}$ is $\mathfrak{B}(\mathbb{R}^k)/\mathfrak{B}(\mathbb{R})$ measurable. Then,

$$w \mapsto g((f_1(\omega), \ldots, f_k(\omega))) = g(f_1(\omega), \ldots, f_k(\omega)).$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. (Composite function usually written without double brackets)

Proposition 7.6 (Continuity preserves measurability in function composition)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i : \Omega \to \mathbb{R}, i = 1, \ldots, k,$ be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Then, if $g : \mathbb{R}^k \to \mathbb{R}$ is continuous,

$$w \mapsto g(f_1(\omega), \ldots, f_k(\omega)).$$

is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Example 7.2 (Continuity preserves measurability)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i : \Omega \to \mathbb{R}, i = 1, \ldots, k,$ be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Then, $\sum_{i=1}^k f_i$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Proposition 2.3).

Example 7.3 (Continuity preserves measurability)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f_i : \Omega \to \mathbb{R}, i = 1, \ldots, k,$ be $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable. Then, $\prod_{i=1}^k f_i$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable (cf. Proposition 2.3).

Definition 7.3 (Simple functions)**.** A function $f : \Omega \to \mathbb{R}$ is called simple if there exists $n \in \mathbb{N}, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and sets $A_1, \ldots, A_n \subset \Omega$ s.t.

$$f(\omega) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(\omega) \quad \omega \in \Omega.$$

That is, a simple function is a finite linear combination of indicator functions.

Example 7.4 (Simple function)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and f be a simple function on Ω, i.e., $f(\omega) = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}(\omega)$. Then, if $A_i \in \mathcal{F}$ for any $i = 1, \ldots, n, f$ is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

My Example 7.1 (Simple function)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}$ be the function defined in 7.3. For this simplified setting, suppose $\Omega = \{1, 2, 3, 4\}$ and $\mathcal{F} = \{\emptyset, \{1, 2\}, \{3, 4\}, \Omega\}$. Moreover, we define our function with $n = 2$, where $\alpha_1 = 3, \alpha_2 = 5, A_1 = \{1, 2\}$ and $A_2 = \{3, 4\}$. Then,

$$f(\omega) = 3 \cdot \mathbb{1}_{\{1,2\}}(\omega) + 5 \cdot \mathbb{1}_{\{3,4\}}(\omega).$$

Now, let's consider two preimages of this function, $f^{-1}(\{3\})$ and $f^{-1}(\{12\})$. Note that both of these sets are Borel sets in $\mathbb{R}$. Also note that, if $B \in \mathfrak{B}(\mathbb{R})$, then,

$$f^{-1}(B) = \{\omega \in \Omega : f(\omega) \in B\}.$$

As seen in Thinking about 7.1. Since f takes the value 3 for $\omega \in \{1, 2\}$, $f^{-1}(\{3\}) = \{1, 2\} \in \mathcal{F}$. And, as f doesn't take any value for values $\notin \{\{1, 2\}, \{3, 4\}\}$, $f^{-1}(\{12\}) = \emptyset \in \mathcal{F}$. So indeed, f is $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Definition 7.4 (Simple functions in standard form)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}$ be a simple function, as defined in Definition 7.3. f is called standard if $\cup_{i=1}^n A_i = \Omega$ and $\{A_1, \ldots, A_n\} \subset \mathcal{F}$ is disjoint. if f is standard, we say that it is a simple function in standard form.

Proposition 7.7 (7.7)**.** TODO

Proposition 7.8 (7.8)**.** TODO

7.2 Functions taking values in the extended real numbers

Definition 7.5 (Measurable functions in $\overline{\mathbb{R}}$)**.** Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \overline{\mathbb{R}}$. We say that f is $\mathcal{F}$ measurable if for any $A \in \mathfrak{B}(\mathbb{R})$, $\{\omega \in \Omega : f(\omega) \in A\} \in \mathcal{F}$ and $\{\omega \in \Omega : f(\omega) = -\infty\} \in \mathcal{F}$ and $\{\omega \in \Omega : f(\omega) = \infty\} \in \mathcal{F}$. Or, in other words, $f^{-1}(A), f^{-1}(-\infty), f^{-1}(\infty) \in \mathcal{F}$.

Remark 7.2 As seen in the script, as, if $f : \Omega \to \mathbb{R}, f^{-1}(-\infty), f^{-1}(\infty) = \emptyset$, any results on $\mathcal{F}$ meeasurable functions $f : \Omega \to \overline{\mathbb{R}}$ also apply to $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable functions $f : \Omega \to \mathbb{R}$.

Remark 7.3 TODO, but important for notation, read it from the script.

Proposition 7.9 (7.9)**.** TODO

Proposition 7.10 (7.10)**.** TODO

Definition 7.6 (Positive and negative parts of a function)**.** TODO

Proposition 7.11. This proposition states that any $\mathcal{F}$-measurable function f can be approximated by a sequence of $\mathcal{F}$-measurable simple functions $(f_n)_{n\in\mathbb{N}}$ such that $f_n(\omega) \to f(\omega)$ for all $\omega \in \Omega$.

My Example 7.2. Consider $\Omega = [0, 1]$ and $\mathcal{F}$ be the Borel σ-field on $[0, 1]$. Let $f(x) = x$. Define the sequence of simple functions $f_n(x) = \frac{\lfloor nx \rfloor}{n}$. Each f_n is $\mathcal{F}$-measurable and $f_n(x) \to x$ as $n \to \infty$.

Proposition 7.12. This proposition extends 7.11 by specifying that if f is non-negative, the convergence of the simple functions can be made monotone, i.e., $f_n(\omega)$ increases with n and converges to $f(\omega)$.

My Example 7.3. Using the same function $f(x) = x$ on $\Omega = [0, 1]$, define $f_n(x) = \frac{\lfloor nx \rfloor}{n}$. Note that $f_n(x) \leq f_{n+1}(x)$ for all $x \in [0, 1]$ and $n \in \mathbb{N}$, ensuring that $f_n(x) \uparrow f(x)$ as $n \to \infty$.

7.3 Sequence of measurable functions

Chapter 8

Integration: Part I

8.1 The integral for non-negative functions

If $f : \Omega \to \overline{\mathbb{R}}$ is s.t. $f(\omega) \geq 0$ for any $\omega \in \Omega$, f is said to be nonnegative.

Definition 8.1 (Finite partitions)**.** Let Ω be a set. A partiton of Ω is a disjoint collection $\{A : A \in P\}, P \subset \mathcal{P}(\Omega)$, s.t. $\cup_{A \in P} A = \Omega$. That is, a partition of Ω is a disjoint collection of subets of Ω whose union is Ω. If ξ is a partition of Ω, a set $A \in \xi$ is referred to as an atom of ξ. A partition ξ of Ω is said to be finite, if it contains a finite number of atoms.

Example 8.1 (Finite partition)**.** Let $\Omega = \{0, 1, \ldots, N\}, N \in \mathbb{N}$. Then, $\xi = \{\{\omega\} : w \in \Omega\}$ is a finite partition of Ω. (Partition contains $N + 1$ elements).

Definition 8.2 ($Z_0^{\mathcal{F}}$)**.** Let $(\Omega, \mathcal{F})$ be a measurable space. We use the notation $Z_0^{\mathcal{F}}(\Omega) = Z_0^{\mathcal{F}}$ for the set which contains all the finite partitions of Ω with atoms from $\mathcal{F}$. That is,

$$Z_0^{\mathcal{F}} = \{\xi : \xi \text{ is finite partition of } \Omega \text{ s.t. for any A } \in \xi, A \in \mathcal{F}\}.$$

Definition 8.3 (Integral for a nonnegative standard simple function)**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \overline{\mathbb{R}}$ be nonnegative and $\mathcal{F}$ measurable. Then, we define

$$S_\mu^f(\xi) = \sum_{A \in \xi} (\inf_{\omega \in A} f(\omega))\mu(A), \quad \xi \in Z_0^{\mathcal{F}},$$

Essentially, $S_\mu^f(\xi)$ approximates the integral of f by considering the smallest value f takes on each piece of the partition and multiplying this by the measure of the piece. And

$$\int_\Omega f(\omega)\mu(d\omega) = \sup_{\xi \in Z_0^{\mathcal{F}}} S_\mu^f(\xi).$$

The integral of f over Ω with respect to μ, is the supremum of $S_\mu^f(\xi)$ over all possible partitions ξ of Ω in $Z_0^{\mathcal{F}}$. This definition captures the idea of the integral as the limit of finer and finer approximations of f by simple functions. Upon the latter definition, we deduce the integral for a (nonnegative) standard simple function (cf. Definition 7.4).

Proposition 8.1. TODO

My Example 8.1 (Integral of a nonnegative standard simple function)**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space with $\Omega = \{a, b, c, d\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, and μ is the counting measure, i.e., $\mu(A)$ is the number of elements in A. Let $f : \Omega \to \overline{\mathbb{R}}$,

$$f(\omega) = \begin{cases} 1 & \text{if } \omega = a, \\ 2 & \text{if } \omega = b, \\ 3 & \text{if } \omega = c, \\ 0 & \text{if } \omega = d \end{cases}$$

Consider the partition $\xi = \{\{a\},\{b\},\{c\},\{d\}\}$. $\inf_{\omega\in\{a\}} f(\omega) = 1$, $\inf_{\omega\in\{b\}} f(\omega) = 2$, $\inf_{\omega\in\{c\}} f(\omega) = 3$, $\inf_{\omega\in\{d\}} f(\omega) = 4$. Since each singleton set in ξ as measure of 1 under μ,

$$S^f_\mu(\xi) = (1\times 1) + (2\times 1) + (3\times 1) + (0\times 1) = 6$$

if $\sup_{\xi\in Z_0^{\mathcal{F}}} S^f_\mu = 6$, which I think it should be, then $\int_\Omega f(\omega)\mu(d\omega) = 6$.

Example 8.2. Example 8.2 interesting and clear, TODO.

Proposition 8.2 (Monotone convergence theorem)**.** Let $(\Omega,\mathcal{F},\mu)$ be a measure space and $f_n : \Omega \to \overline{\mathbb{R}}$, $n\in\mathbb{N}$, be a sequence of nonnegative $\mathcal{F}$ measurable functions s.t. for any $\omega\in\Omega, f_n(\omega)\uparrow f(\omega)$ for some $f:\Omega\to\overline{\mathbb{R}}$. Then,

$$\int_\Omega f_n(\omega)\mu(d\omega) \uparrow \int_\Omega f(\omega)\mu(d\omega).$$

Proposition 8.3 (The integral of nonnegative functions is linear)**.** Let $(\Omega,\mathcal{F},\mu)$ be a measurable space, $f,g:\Omega\to\overline{\mathbb{R}}$ be two nonnegative and $\mathcal{F}$ measurable functions. Given $\alpha,\beta\in[0,\infty)$ we have that

$$\int_\Omega (\alpha f + \beta g)(\omega)\mu(d\omega) = \alpha\int_\Omega f(\omega)\mu(d\omega) + \beta\int_\Omega g(\omega)\mu(d\omega).$$

As a consequence of the latter two proposition we have the following result:

Proposition 8.4. Let $(\Omega,\mathcal{F},\mu)$ be a measure space and $f_i:\Omega\to\overline{\mathbb{R}}, i\in\mathbb{N}$, be a sequence of nonnegative $\mathcal{F}$ measurable functions, then

$$\int_\Omega \left(\sum_{i\in\mathbb{N}} f_i\right)(\omega)\mu(d\omega) = \sum_{i\in\mathbb{N}}\left(\int_\Omega f_i(\omega)\mu(d\omega)\right).$$

Definition 8.4 (True almost everywhere ($a.e.$))**.** Let $(\Omega,\mathcal{F},\mu)$ be a measure space. Suppose that for any $\omega\in\Omega, S(\omega)$ is a statment on Ω. We say S is true μ almost everywhere ($a.e.$) if $\mu(\{\omega : S(\omega) \text{ is false}\}) = 0$.

Example 8.3 ($\mu(a.e.)$)**.** Interesting and clear. TODO.

Proposition 8.5. Let $(\Omega,\mathcal{F},\mu)$ be a measure space. Assume that $f,g:\Omega\to\overline{\mathbb{R}}$ be two nonnegatibe and $\mathcal{F}$ measurable functions.

(i) If $\mu(\{\omega : f(\omega) > 0\}) > 0$, then $\int_\Omega f(\omega)\mu(d\omega) > 0$;

(ii) If $\int_\Omega f(\omega)\mu(d\omega) < \infty$, then $f < \infty$ μ $a.e.$;

(iii) If $f \le g$ μ $a.e.$, then $\int_\Omega f(\omega)\mu(d\omega) \le \int_\Omega g(\omega)\mu(d\omega)$;

(iv) If $f = g$ μ $a.e.$, then $\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega g(\omega)\mu(d\omega)$.

8.2 Integrable functions

We recall the definiton of the positive (f^+) and negative (f^-) parts of a function (cf. Definition 7.6). Pay attention, f^- is basically the negative part of the function, but reflected by the x-axis. The result is positive. Also see 7.2

Definition 8.5 (Integral of an integrable function)**.** Let $(\Omega,\mathcal{F},\mu)$ be a measure space and $f:\Omega\to\overline{\mathbb{R}}$ be a $\mathcal{F}$ measurable function. The integral of f is defined by:

$$\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega f^+(\omega)\mu(d\omega) - \int_\Omega f^-(\omega)\mu(d\omega),$$

unless $\int_\Omega f^+(\omega)\mu(d\omega) = \int_\Omega f^-(\omega)\mu(d\omega) = \infty$, in which case $\int_\Omega f(\omega)\mu(d\omega)$ is not defined.
If both $\int_\Omega f^+(\omega)\mu(d\omega) < \infty$ and $\int_\Omega f^-(\omega)\mu(d\omega) < \infty$, f is said to be integrable.

(NOTE) This assumption is definied upon the measure μ, i.e., if one wants to further refer to the measure of integration one specifies that f is integrable with respect to μ.

Proposition 8.6 (Generalisation of the condition for f to be integrable)**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable. Then, f is integrable if and only if $\int_\Omega |f(\omega)| \mu(d\omega) < \infty$.

Proposition 8.7 (Extension (cf. (iii) Proposition 8.5))**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f, g : \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}$ measurable. If f and g are integrable and $f \le g$ *a.e.*, then, $\int_\Omega f(\omega)\mu(d\omega) \le \int_\Omega g(\omega)\mu(d\omega)$.

Proposition 8.8 (Extension (c.f. Proposition 8.3))**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measurable space, $f, g : \Omega \to \overline{\mathbb{R}}$ be two integrable and $\mathcal{F}$ measurable functions. Then, for any $\alpha, \beta \in \mathbb{R}$ we have that $\alpha f + \beta g$ is integrable and

$$\int_\Omega (\alpha f + \beta g)(\omega)\mu(d\omega) = \alpha \int_\Omega f(\omega)\mu(d\omega) + \beta \int_\Omega g(\omega)\mu(d\omega).$$

8.3 Fatou's lemma and Lebesgue's dominated convergence theorem

Proposition 8.9 (Fatou's lemma)**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f_n : \Omega \to \overline{\mathbb{R}}, n \in \mathbb{N}$, be a sequence of nonnegative and $\mathcal{F}$ measurable function. Then,

$$\int_\Omega \liminf_{n\to\infty} f_n(\omega)\mu(d\omega) \le \liminf_{n\to\infty} \int_\Omega f_n(\omega)\mu(d\omega).$$

8.4 Integration over measurable sets

Tool 8.1 (Integration over $\bigcup_{i \in I} A_i$)**.** (From Ex. 8.9). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \overline{\mathbb{R}}$ be a $\mathcal{F}$ measurable function. Suppose that either f is nonnegative of integrable and let $\{A_i : i \in I\} \subset \mathcal{F}$ be disjoint, where $I \subset \mathbb{N}$. Then

$$\int_{\bigcup_{i\in I} A_i} f(\omega)\mu(d\omega) = \sum_{i \in I} \int_{A_i} f(\omega)\mu(d\omega).$$

Chapter 9

Integration: Part II

9.1 Pushforward measure

Definition 9.1 (Pushforward function)**.** Let $(\Omega, \mathcal{F})$ and $(\Omega^*, \mathcal{F}^*)$ be two measurable spaces and $g : \Omega \to \Omega^*$ be $\mathcal{F}/\mathcal{F}^*$ measurable. Let μ be a measure on $\mathcal{F}$. Define the function

$$\mu g^{-1}(A^*) = \mu(g^{-1}(A^*)) = \mu(\{\omega \in \Omega : g(\omega \in A^*)\}), \quad A^* \in \mathcal{F}^*.$$

The measure μg^{-1} is referred to as the pushforward measure of μ. This means that μg^{-1} measures, in terms of μ, the pre-image of each set A^* under g. Hence, μ is a valid measure on $(\Omega^*, \mathcal{F}^*)$!! It provides a way to "transfer" the measure from $(\Omega, \mathcal{F})$ to $(\Omega^*, \mathcal{F}^*)$ via the function g.

Proposition 9.1. TODO

9.2 Densities

Proposition 9.2 (ν is a measure on $\mathcal{F}$)**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $\phi : \Omega \to \overline{\mathbb{R}}$ be a nonnegative and $\mathcal{F}$ measurable function. Then, ν defined by

$$\nu(A) = \int_A \phi(\omega)\mu(d\omega), \quad A \in \mathcal{F},$$

is a measure on $\mathcal{F}$

Definition 9.2 (ϕ, density of ν in respect to μ)**.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and ν be a measure on $\mathcal{F}$. A nonnegative and $\mathcal{F}$ measurable funtion $\phi : \Omega \to \overline{\mathbb{R}}$ is said to be a density of ν with respect to μ if for any $A \in \mathcal{F}, \nu(A) = \int_A \phi(\omega)\mu(d\omega)$.

Proposition 9.3. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Suppose that ν is a measure on $\mathcal{F}$ with density ϕ with respect to μ. Then

(i) for any nonnegative and $\mathcal{F}$ measurable function f,

$$\int_A f(\omega)\nu(d\omega) = \int_A f(\omega)\phi(w)\mu(d\omega), \quad A \in \mathcal{F};$$

(ii) f is integrable with respect to ν if and only if $f\phi$ (the product of the two functions) is integrable with respect to μ. This is clear in (i).

(iii) if $f\phi$ is integrable with respect to μ, then (i) holds.

9.3 Integration with respect to the Lebesgue measure on the real line

Definition 9.3. Consider the measure space $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \lambda)$, where λ is the Lebesgue measure on the Borel σ-field $\mathfrak{B}(\mathbb{R})$. In accordance with Definition 8.5, a $\mathfrak{B}(\mathbb{R})$ measurable function $f : \mathbb{R} \to \overline{\mathbb{R}}$ is Lebesgue integrable if $\int_{\mathbb{R}} |f(x)|\lambda(dx) < \infty$. The integral of f with respect to λ is denoted with $\int_{\mathbb{R}} f(x)dx$, i.e., $\int_{\mathbb{R}} f(x)dx = \int_{\mathbb{R}} f(x)\lambda(dx)$. If $E \subset \mathbb{R}$ and $\lambda|_E$ is the restriction of λ to $\mathfrak{B}(E)$ (cf. Definiton 4.2), then a $\mathfrak{B}(E)$ measurable function $f : E \to \overline{\mathbb{R}}$ is referred to as Lebesgue integrable if $\int_E |f(x)|\lambda|_E(dx) < \infty$. Also in this case we write $\int_E |f(x)|\lambda|_E(dx) = \int_E f(x)dx$.

In accordance with the fact that the Lebesgue measure of a single point is zero, we adapt the following definition.

Definition 9.4. TODO. Interesting but easy and well known.

We review the definition of a Riemann integrable function:

Definition 9.5 (title)**.**

Definition 9.6. Let $f : [a, b] \to \mathbb{R}$ be $\mathfrak{B}([a, b])$ measurable and Lebesgue integrable. The integral of f when the limits of integration are reverted is defined as follows

$$\int_a^b f(x)dx = -\int_b^a f(x)dx.$$

9.4 Change of variable

9.5 Integration on product spaces

Definition 9.7 (Product σ-field)**.** Let $(X, \mathscr{X})$ and $(Y, \mathscr{Y})$ be two measurable spaces. The product σ-field on the cartesian product $(X \times Y)$ is defined by

$$\mathscr{X} \otimes \mathscr{Y} = \sigma(\{A \times B : A \in \mathscr{X}, B \in \mathscr{Y}\}).$$

The definition extends to products of higher order.

Consider a collection of measure spaces $(X_1, \mathscr{X}_1), \ldots, (X_n, \mathscr{X}_n)$. We define

$$\otimes_{i=1}^n \mathscr{X}_i = \mathscr{X}_1 \otimes \ldots \otimes \mathscr{X}_n = \sigma(\{A_1 \times \ldots \times A_n : A_i \in \mathscr{X}_i, i = 1, \ldots, n\}).$$

One can also show that the latter product is associative.

9.6 Lecture

Partial integration and substitution TODO.

Exercise 9.1 (9.6)**.** ν is a measure with density ϕ with respect to μ. f nonnegative and $\mathcal{F}$ measurable. Prove:

(i) $\int_A f(\omega)\nu(d\omega) = \int_A f(\omega)\phi(w)\mu(d\omega)$

NOTE $\nu(d\omega) = \phi(\omega)\mu(d\omega)$ short notation for ν has density ϕ:

1. Definition of ν having a density ϕ with respect to μ: When we say that ν has a density ϕ with respect to μ, it means that for any measurable set $A \in \mathcal{F}$, the measure ν of A can be computed as:

$$\nu(A) = \int_A \phi(\omega)\mu(d\omega).$$

This is the integral of the function ϕ over the set A, with respect to the measure μ.

2. Notation $\nu(d\omega) = \phi(\omega)\mu(d\omega)$: This notation is shorthand and is used to express how ν acts on infinitesimal elements in a manner analogous to how μ acts, but scaled by the function ϕ. It is essentially saying that for a small element $d\omega$, the measure $\nu(d\omega)$ is given by $\phi(\omega)\mu(d\omega)$.

3. Clarification on $\int_{d\omega} \phi(\omega)\mu(d\omega)$: The correct notation or expression should not involve integrating over an "infinitesimal element" $d\omega$. The differential notation $\nu(d\omega) = \phi(\omega)\mu(d\omega)$ is symbolic and used to express the relationship between ν and μ at a small scale, rather than an actual operation.

In summary, $\nu(d\omega) = \phi(\omega)\mu(d\omega)$ is a concise way to denote that ν is derived by weighting μ by the density ϕ, and this relationship is used to transform integrals with respect to ν into integrals with respect to μ weighted by ϕ.

(ii) f integrable w.r.t. $\nu \iff f\phi$, $(f(\omega)\phi(\omega))$, integrable w.r.t. μ.

(iii) if either of the two statments in (ii) holds, then (i) holds.

Proof:
(i). Let f be a standard simple function, $f = \sum_{n=1}^{\mathbb{N}} \alpha_i \mathbb{1}_{Ai}$, then

$$\int_A f(\omega)\nu(d\omega) = \int_A (\sum_{n=1}^{\mathbb{N}} \alpha_i \mathbb{1}_{Ai}(\omega))\nu(d\omega) = \sum_{n=1}^{\mathbb{N}} \alpha_i \int_A \mathbb{1}_{Ai}(\omega)\nu(d\omega) = \sum_{n=1}^{\mathbb{N}} \alpha_i \int_\Omega \mathbb{1}_A(\omega)\mathbb{1}_{Ai}(\omega)\nu(d\omega)$$

$$= \sum_{n=1}^{\mathbb{N}} \alpha_i \int_\Omega \mathbb{1}_{A\cap A_i}(\omega)\nu(d\omega) = \sum_{n=1}^{\mathbb{N}} \alpha_i \nu(A \cap A_i) = \sum_{n=1}^{\mathbb{N}} \alpha_i \int_{A\cap A_i} \phi(\omega)\mu(d\omega) = \sum_{n=1}^{\mathbb{N}} \alpha_i \int_A \mathbb{1}_{A_i}(\omega)\phi(\omega)\mu(d\omega)$$

$$= \int_A \sum_{n=1}^{\mathbb{N}} \alpha_i \mathbb{1}_{A_i}(\omega)\phi(\omega)\mu(d\omega) = \int_A f(\omega)\phi(w)\mu(d\omega).$$

Hence we have verified (i) if f is standard and simple.
In order to verify it for nonnegative functions:
(IMPORTANT; TOOL, TO ADD) Recall (chapter 7): Any f nonnegative and $\mathcal{F}$ measurable can be approximated by a standard simple function, i.e., $\exists (f_n)_{n\in\mathbb{N}}$ s.t. $f_n(\omega) \uparrow f(\omega)$. By the monotone convergence theorem,

$$\int_\Omega f(\omega)\nu(d\omega) = \lim_{n\to\infty} \int_\Omega f_n(\omega)\nu(d\omega) = \lim_{n\to\infty} \int_\Omega f_n(\omega)\phi(\omega)\mu(d\omega)$$

f_n converges to f

$$\overset{\text{(again monotone convergence)}}{=} \int_\Omega f(\omega)\phi(\omega)\mu(d\omega).$$

This proves (i).

(ii). $\int_A |f(\omega)|\nu(d\omega) < \infty$ (definition of integrability), $= \int_A |f(\omega)\phi(\omega)|\mu(d\omega)$, and we know that the equality holds by (i). This shows (ii).

(iii). Recall $f^+ = max(f, 0)$, $f^- = max(-f, 0)$. Positive and negative parts of f. Cuts out all negative points. We know,

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

f integrable w.r.t. ν implies that,

$$\int_\Omega f(\omega)\nu(d\omega) = \int_\Omega f^+\nu(d\omega) - \int_\Omega f^-(\omega)\nu(d\omega).$$

By (i) applied to f^+ and f^-,

$$= \int_\Omega f^{(+)}(\omega)\phi(\omega)\mu(d\omega) - \int_\Omega f^-(\omega)\phi(\omega)\mu(d\omega) = \int_\Omega f(\omega)\phi(\omega)\mu(d\omega).$$

Exercise 9.2 (9.7)**.** b) TODO
$\frac{1}{2\pi}\int_{\mathbb{R}^2} e^{-(\frac{x^2+y^2}{2})} d(x, y)$, continuoius as composition of continuous functions, and nonnegative. Fobini - Tonelli Theorem:

$$= \frac{1}{2\pi}\int_\mathbb{R} e^{-\frac{x^2}{2}}(\int_\mathbb{R} e^{-\frac{y^2}{2}} dy)dx.$$

$$= \frac{1}{2\pi}(\int_\mathbb{R} e^{\frac{-x^2}{2}} dx)^2.$$

u $= \frac{x}{\sqrt{2}}$ substitute

$$= \frac{1}{2\pi}(\int_\mathbb{R} e^{-u^2}\sqrt{2}du)^2.$$

$$= \frac{1}{\pi}(\int_\mathbb{R} e^{-u^2} du)^2 = \frac{\pi}{\pi}.$$

Remember Gaussian integral:

$$\int_\mathbb{R} e^{-x^2} dx = \sqrt{\pi}.$$

Chapter 10

General notions in Probability

10.1 Probability spaces

Definition 10.1. Let $(\Omega, \mathcal{F})$ be a measurable space. A probability $\mathbb{P}$ on $\mathcal{F}$ is a measure on $\mathcal{F}$ s.t. $\mathbb{P}(\Omega) = 1$. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is referred to as a probability space.

Example 10.1. Let Ω be a finite and nonempety set. Define

$$\mathbb{P}(A) = \frac{\#A}{\#\Omega}, \quad A \in \mathcal{P}(\Omega), .$$

Where $\mathcal{P}(\Omega)$ is the power set on Ω. Then, $\mathbb{P}$ is a probability on $\mathcal{P}(\Omega)$.

Example 10.2. Let C be a set s.t. $\#C = 52$. Suppose that

$$C = S_1 \cup S_2 \cup S_3 \cup S_4,$$

with $\{S_1, S_2, S_3, S_4\}$ disjoint and s.t. $\#S_i = 13$ for all $i = 1, 2, 3, 4$. We remain in the setting of the previous example with

$$\Omega = \{A \subset C : \#A = 5\},$$

and $\mathbb{P}$ on $\mathcal{P}(\Omega)$ defined as in exercise 10.1. Upon exercise 1.1, we already know that $\#\Omega = \binom{52}{5}$. Let

$$A_i = \{A \subset S_i : \#A = 5\}, \quad i = 1, 2, 3, 4,$$

TODO

10.2 Random variables and random vectors

Definition 10.2 (Random variable)**.** Let $(\Omega, \mathcal{F})$ be a measurable space. A map $X : \Omega \to \mathbb{R}$ is referred to as a random variable on $(\Omega, \mathcal{F})$ if it if $\mathcal{F}/\mathfrak{B}(\mathbb{R})$ measurable.

Definition 10.3 (Random vector)**.** Let $(\Omega, \mathcal{F})$ be a measurable space. A map $X : \Omega \to \mathbb{R}^k$ is referred to as a random vector on $(\Omega, \mathcal{F})$ if it is $\mathcal{F}/\mathfrak{B}(\mathbb{R}^k)$ measurable.

Proposition 10.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X be a random vector on $(\Omega, \mathcal{F})$. A random variable Y on $(\Omega, \mathcal{F})$ is $\sigma(X)$ measurable if and only if there exists a function $f : \mathbb{R}^k \to \mathbb{R}$ which is $\mathfrak{B}(\mathbb{R}^k)$ measurable s.t. $Y = f(X)$.

Definition 10.4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. The distribution or law of a random vector on $(\Omega, \mathcal{F})$ is the pushforward measure $P_X = \mathbb{P}X^{-1}$ on $\mathfrak{B}(\mathbb{R}^k)$ (cf. Definition 9.1). In particular, for any $B \in \mathfrak{B}(\mathbb{R}^k)$, we use the simplified notation

$$\{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\},$$

and hence

$$P_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) = \mathbb{P}(X \in B).$$

For now, unless mentioned otherwise, if $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, any random vector X is a random vector on $(\Omega, \mathcal{F})$, i.e., a $\mathcal{F}$ measurable function with values in $\mathbb{R}^k$.

10.3 Discrete laws

Definition 10.5 (Discrete random vector)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random vector is referred to as discrete if there exists a countable set $E = E_1 \times \ldots \times E_k \subset \mathbb{R}^k$ s.t. $P_X(E) = 1$. That is to say that the law of X has a countable support.

Proposition 10.2 (Discrete random vector)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random vector X is discrete if and only if

$$P_X = \sum_{x \in E} p_x \delta_x, \quad p_x = \mathbb{P}(X = x),$$

for some countable set $E = E_1 \times \ldots \times E_k \subset \mathbb{R}^k$. In particular, for any $B \in \mathfrak{B}(\mathbb{R}^k)$, $P_X(B) = \sum_{x \in B \cap E} p_x$.

Proof of Proposition 10.2. Suppose that X is discrete. Let $B \in \mathfrak{B}(\mathbb{R}^k)$. We have that

$$P_X(B) = P_X(B \cap E) = \mathbb{P}(X \in B \cap E) = \mathbb{P}(\bigcup_{x \in B \cap E} \{X = x\})) = \sum_{x \in B \cap E} p_x = \sum_{x \in E} p_x \delta_x(B).$$

with respect to the other direction, if P_X is given as in Prop. 10.2, then

$$1 = P_X(\mathbb{R}^k) = \sum_{x \in E} p_x \delta_x(\mathbb{R}^k) = \sum_{x \in E} p_x = \mathbb{P}(X \in E) = P_x(E),$$

i.e., X is a discrete random vector according to Def. 10.5.

Example 10.3 (Tail, head)**.** Let $\Omega = \{t, h\}$ and

$$X(\omega) = \begin{cases} 0, & \text{if } w = t, \\ 1, & \text{if } w = h. \end{cases}$$

Then, X is a random variable on $(\Omega, \mathcal{P}(\Omega))$.

Explanation: for the cases $X^{-1}(0) = t$ and $X^{-1}(1) = h$ it is clear. Consider then $X^{-1}(\omega) = \emptyset \in \mathcal{P}(\Omega)$ for any $\omega \notin \{1, 2\}$.

Suppose that $\mathbb{P}$ is a probability on $\mathcal{P}(\Omega)$ s.t. $\mathbb{P}(X^{-1}(0)) = \mathbb{P}(X = 0) = 1 - p$ and $\mathbb{P}(X = 1) = p$. Clearly, $\mathbb{P}(X \in \{0, 1\}) = P_X(\{0, 1\}) = 1$. By Prop. 10.2, we deduce that the law of X is given by

$$P_X = (1 - p)\delta_0 + p\delta_1.$$

That is, for any $B \in \mathfrak{B}(\mathbb{R})$,

$$P_X(B) = (1 - p)\delta_0(B) + p\delta_1(B) = \begin{cases} 0, & \text{if } 0 \notin B \text{ and } 1 \notin B \\ 1 - p, & \text{if } 0 \in B \text{ and } 1 \notin B \\ p, & \text{if } 0 \notin B \text{ and } 1 \in B \\ 1, & \text{if } 0 \in B \text{ and } 1 \in B \end{cases}.$$

For example, for $B = (2, 4]$, $X^{-1}(B) = \emptyset \in \mathcal{P}(\Omega)$, and $P_X(B) = P_X(\emptyset) = 0$. Also note that, for example, $\mathbb{P}(X = 0) = \mathbb{P}(X^{-1}(0)) = \mathbb{P}(t) = 1 - p$.

My Example 10.1 (Examples of discrete probability distributions)**.**

Discrete uniform: $E \subset \mathbb{R}$ is a finite set s.t. $\#E = n$, and $p_x = \frac{1}{n}$ for any $x \in E$.

Bernoulli: $E = \{0, 1\}$ and $p_0 = 1 - p$ and $p_1 = p$, $p \in [0, 1]$.

Binomial: $E = \{0, 1, \ldots, n\}$, $n \in \mathbb{N}$ and $p_x = \binom{n}{x} p^x (1 - p)^{n - x}$, $p \in [0, 1]$.

Geometric: $E = \mathbb{N}$ and $p_x = (1 - p)^{x - 1} p$, $p \in (0, 1)$.

Poisson: $E = \mathbb{N} \cup \{0\}$ and $p_x = (\frac{\lambda}{x!}) e^{-\lambda}$, $\lambda > 0$.

Multinomial: TODO: Write multinomial discrete probability distribution.

Remark 10.1. If $X = (X_1, \ldots, X_k) : \Omega \to \mathbb{R}^k$ is discrete with support $E = E_1, \ldots, E_k$, we apply Prop. 10.2 and deduce that for any $i = 1, \ldots, k$,

$$\begin{aligned}\mathbb{P}(X_i = x) &= \mathbb{P}(X_1 \in \mathbb{R}, \ldots, X_{i-1} \in \mathbb{R}, X_i = x, X_{i+1} \in \mathbb{R}, \ldots, X_k \in \mathbb{R}) \\ &= P_X(\mathbb{R} \times \ldots \times \mathbb{R} \times x \times \mathbb{R} \times \ldots \times \mathbb{R}) \\ &= \sum_{(x_1,\ldots,x_k)\in E, x_i = x} p_{x_1}, \ldots, x_k.\end{aligned}$$

Given $i = 1, \ldots, k$, we apply the notation,

$$x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k),$$

(Every x apart from x_i). And,

$$E_{-i} = E_1 \times \ldots \times E_{i-1} \times E_{i+1} \times \ldots \times E_k.$$

Then, we obtain

$$\mathbb{P}(X_i = x) = \sum_{x_{-i}\in E_{-i}} p_{x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k}.$$

Notice that the sum is zero if $x \notin E_i$ (X_i has support E_i). For example, for $k = 3$,

$$\mathbb{P}(X_1 = x) = \sum_{(x_2,x_3)\in E_2\times E_3} p_{x,x_2,x_3},$$

where

$$p_{x,x_2,x_3} = \mathbb{P}(\{X_1 = x\} \cap \{X_2 = x_2\} \cap \{X_3 = x_3\}).$$

10.4 Continuous laws

Definition 10.6 (Continuous random vector)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random vector is referred to as continuous if the law of X has density $\phi : \mathbb{R}^k \to [0, \infty)$ with respect to the Lebesgue measure on $\mathfrak{B}(\mathbb{R}^k)$,

$$P_X(B) = \int_B \phi(x)dx.$$

The density ϕ of P_X is referred to as a probability density function.

My Example 10.2 (Classical examples of probability distributions with probability density function ϕ)**.** TODO: Write the probability distributions.

Continuous uniform:
Exponential:
Normal:
Multivariate Normal:

Definition 10.7. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Suppose that fo any $\omega \in \Omega$, $S(\omega)$ is a statement on Ω. We say S is true $\mathbb{P}$ almost surely (a.s.) if $\mathbb{P}(\{\omega : S(\omega) \text{ is true}\}) = 1$. (Cf. Def. 8.4).

10.5 Expectation

Definition 10.8 (Expectation of X)**.** TODO: Write definition

Proposition 10.3 (Expectation of $f(X)$)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X be a random vector. Then, for any nonnegative and $\mathfrak{B}(\mathbb{R}^k)$ measurable map $f : \mathbb{R}^k \to \overline{\mathbb{R}}$,

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}^k} f(x)P_X(dx).$$

In addition, if f is not necessarily nonnegative, this proposition holds if $\mathbb{E}[|f(X)|] < \infty$.

Remark 10.2. TODO: Write remark

Example 10.4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and U be a random variable with uniform law on $[0,1]$, i.e., $P_U(dx) = \mathbb{1}_{[0,1]}(x)dx$. Define the random variable $X = -2\log(U)$. By Prop. 10.3, for any $f : \mathbb{R} \to \mathbb{R}$, nonnegative and $\mathfrak{B}(\mathbb{R})$ measurable,

$$\mathbb{E}[f(X)] = \mathbb{E}[f(-2\log(U))] = \int_{\mathbb{R}} f(-2\log(u))\mathbb{1}_{[0,1]}(u)du = \int_{[0,1]} f(-2\log(u))du.$$

We then substitute $x = -2\log(u)$. Note that $u = e^{-x/2}$, and that $du = -\frac{e^{-x/2}}{2}$. We use Def. 9.6

$$\int_{[x(0),x(1)]} f(x)\left(-\frac{e^{-x/2}}{2}\right)dx = \int_{[x(1),x(0)]} f(x)\left(\frac{e^{-x/2}}{2}\right)dx = \int_{[0,\infty]} f(x)\frac{e^{-x/2}}{2}dx.$$

By Remark 10.2, $P_X(dx) = \frac{e^{-x/2}}{2}$, i.e., the law of X is exponential with $\lambda = \frac{1}{2}$.

10.6 Distribution function

Definition 10.9. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X be a random variable. The distribution function F of X is defined by

$$F_X(t) = \mathbb{P}(X \leq t) = P_X((-\infty, t]).$$

Remark 10.3. Using Prop. 10.2, if X is discrete , we have that for any $t \in \mathbb{R}$,

$$F_X(t) = \sum_{x \in E, x \leq t} p_x.$$

If X is continuous with law that has probability density function ϕ, we have upon Def. 10.6 that for any $t \in \mathbb{R}$,

$$F_X(t) \int_{(-\infty,t]} \phi(x)dx.$$

Chapter 11

Collection of random vectors

11.1 Independence

Definition 11.1 (Independent sub-σ-fields)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be n sub-σ-fields on Ω. $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are said to be independent if for any $A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n$,

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \times \ldots \times \mathbb{P}(A_n).$$

Definition 11.2. TODO: Understand if it's useful. If so, write it down and explain it.

Example 11.1. Quite clear. it is implied that after that the ball is drawn, it has to be put back into the urn.

Proposition 11.1. Let $X_1, \ldots, X_n$ be n random variables.

(i) Suppose that for any $i = 1, \ldots, n$, $P_{X_i}(dx) = \phi(x)dx$, i.e., P_{X_i} has probability density function ϕ_i. Then, if $X_1, \ldots, X_n$ are independent, the law of the random vector $X = (X_1, \ldots, X_n)$ has probability density function

$$\phi(x) = \prod_{i=1}^{n} \phi_i(x_i), \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

(ii) Suppose that the random vector $X = (X_1, \ldots, X_n)$ is s.t. $P_X(dx) = \phi(x)dx$, TODO: Finish writing

11.2 Sums of independent random vectors

11.3 Gauss vectors

Definition 11.3 (Gauss vector)**.** A random vector $X = (X_1, \ldots, X_k)$ is said to be a Gauss vector if and only if for any $v \in \mathbb{R}^k$, the random variable

$$v^t X = v_1 X_1 + \ldots + v_k X_k,$$

is Gaussian.

Remark 11.1. TODO

11.4 Lecture

Exercise 11.1. $X_1, \dots, X_n$ independent discrete uniform, on $\{1, \dots, p\}$, $p \in \mathbb{N}$. Meaning that each X_i can take any value in $\{1, \dots, p\}$ with equal probability $\frac{1}{p}$. Find the law of $M = max\{X_1, \dots, X_n\}$.

Note, Let X be discrete uniform on $\{1, \dots, p\}$, then the support of X is $\{1, \dots, p\}$. (The set of all values s.t. $\mathbb{P}(X = x) > 0$)

The support of M also is $\{1, \dots, p\}$. Why?

$$\mathbb{P}(M \notin \{1, \dots, p\}) \leq \mathbb{P}(\bigcup_{i=1}^{n} \{X_i \notin \{1, \dots, p\}\}) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \notin \{1, \dots, p\}) = 0.$$

$$\Rightarrow \mathbb{P}(M \notin \{1, \dots, p\}) = 0 \Rightarrow \mathbb{P}(M \in \{1, \dots, p\}) = 1.$$

Let $t \in \mathbb{R}$,

$$\mathbb{P}(M \leq t) = F_M(t) = \mathbb{P}(max\{X_1, \dots, X_n\} \leq t) = \mathbb{P}(\bigcap_{1=1}^{n} \{x_i \leq t\}).$$

$$\overset{\text{independence}}{=} \prod_{i=1}^{n} \mathbb{P}(X_i \leq t) = \prod_{i=1}^{n} P_{X_i}((-\infty, t]) = \mathbb{P}(X \leq t)^n.$$

Then, the law of X is,

$$\mathbb{P}(X \leq t) = \begin{cases} 0, & t < 1 \\ \frac{\#\{k : k \leq t\}}{p} & 1 \leq t \leq p \\ 1, & t \geq p \end{cases}.$$

And the law of M,

$$F_M(t) = \mathbb{P}(X \leq t)^n = \begin{cases} 0, & t < 1 \\ (\frac{\#\{k : k \leq t\}}{p})^n, & 1 \leq t \leq p \\ 1, & t > p \end{cases}.$$

Also note, as it is discrete,

$$F_M(i) - F_M(i+1) = \sum_{k=1}^{i} \mathbb{P}(M = k) - \sum_{k=1}^{i-1} \mathbb{P}(M = k) = \mathbb{P}(M = i).$$

Exercise 11.2. X_1, X_2 Poisson with parameters λ and μ respectively. What is the law of $X_1 + X_2$?

Note in general, for $X_1 + X_2 = z$:

Discrete case, $E_1 + E_2 = E_{sum}$, get support of X_1, X_2, and then, $\forall z \in Z_{SUM}$,

$$P_Z(\{z\}) = \sum_{X_2 \in E_2} P_{X_1}(\{z - x_2\}) P_{X_2}(\{x_2\}).$$

Continuous case, (densities ϕ_1, ϕ_2), density of Z

$$\phi(z) = \int_{\mathbb{R}} \phi_1(z - x_2) \phi_2(x_2) dx_2.$$

Let $E_1 = \mathbb{N} \cup \{0\}$, $E_2 = \mathbb{N} \cup \{0\}$, the support is

$$E_1 + E_2 = \mathbb{N} \cup \{0\} \overset{\text{by definition}}{=} \{x_1 + x_2 : x_1 \in E_1, x_2 \in E_2\}.$$

Knowing the support helps us, we know where we can sum. Here we know that for $k > z$, $P_{X_1} = 0$. We then apply the formula for the discrete case:

$$P_Z(\{z\}) = \sum_{k \in \mathbb{N} \cup \{0\}} P_{X_1}(\{z - k\}) P_{X_2}(\{k\}) \overset{\text{for } k > z,\ P_{X_1} = 0}{=} \sum_{k=0}^{z} P_{X_1}(\{z - k\}) P_{X_2}(\{k\})$$

$$\overset{\text{plug in distribution}}{=} \sum_{k=0}^{z} e^{-\lambda} \frac{\lambda^{(z-k)}}{(z-k)!} e^{-\mu} \frac{\mu^k}{k!} = e^{-(\lambda+\mu)} \sum_{k=0}^{z} \frac{1}{(z-k)!k!} \mu^k \lambda^{(z-k)}$$

Use binomial theorem, $(\mu+\lambda)^z = \sum_{k=0}^{z} \binom{z}{k} \mu^k \lambda^{(z-k)}$. Multiply by $z!$ inside and divide outside of the sum.

$$= \frac{e^{-(\lambda+\mu)}}{z!} \sum_{k=0}^{z} \frac{1 \times z!}{(z-k)!k!} \mu^k \lambda^{(z-k)} = e^{-(\lambda+\mu)} \frac{(\mu+\lambda)^z}{z!}.$$

Hence,

$$P_Z(\{z\}) = e^{-(\lambda+\mu)} \frac{(\mu+\lambda)^z}{z!}.$$

Note, we see that the law is the same as before, with the parameters added.

Chapter 12

Mock exam 1

Solve with the pdf of the mock exam on the side.

Notation: We recall some of the terminology:

- Given a nonempty set Ω, $\mathcal{P}(\Omega)$ is the power set on Ω;
- $\mathfrak{B}(\mathbb{R}^k)$ denotes the Borel σ-field on $\mathbb{R}^k$, $k \geq 1$;
- The measure

$$\mu(A) = \begin{cases} \#A, & \text{if } A \text{ is finite} \\ \infty, & \text{otherwise,} \end{cases} \qquad A \in \mathcal{P}(\Omega),$$

  is referred to as the counting measure on $\mathcal{P}(\Omega)$;
- Given a measurable space $(\Omega, \mathcal{F})$ and $x \in \Omega$, we write δ_x for the measure

$$\mathcal{F} \ni A \mapsto \delta_x(A) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{otherwise.} \end{cases}$$

Exercise 12.1.

(a) Refer to Def. 4.1.

(b) Measure on $\mathcal{F}$ (cf. Def 5.1).

- (i) $\mu_1(\emptyset) = C\mu(\emptyset) = 0$;

  (ii) We know that item ii holds for the counting measure by definition. For our redefined counting measure,

$$\mu_1(\bigcup_{i\in\mathbb{N}} A_i) = C\mu(\bigcup_{i\in\mathbb{N}} A_i) = C\sum_{i\in\mathbb{N}} \mu(A_i) = \sum_{i\in\mathbb{N}} C\mu(A_i) = \sum_{i\in\mathbb{N}} \mu_1(A_i).$$

- (i) $\mu_2(\emptyset) = \int_\emptyset f(\omega)\mu(d\omega) = 0$;

  (ii)

$$\mu_2(\bigcup_{i\in\mathbb{N}} A_i) = \int_{\bigcup_{i\in\mathbb{N}} A_i} f(\omega)\mu(d\omega) \overset{\text{Tool 8.1}}{=} \sum_{i\in\mathbb{N}} \int_{A_i} f(\omega)\mu(d\omega) = \sum_{i\in N} \mu_2(A_i).$$

- (i) $\mu_3(\emptyset) = \frac{1}{2} + \lambda(\emptyset) = \frac{1}{2}$.

We see that μ_3 is clearly not a measure on $\mathcal{F}$.

(c) Probability measure cf. Def. 10.1.

- $$P_1(\mathbb{R}) = \int_{\mathbb{R}} \mathbb{1}_{[0,\infty)}(x)e^{-x}dx = \int_{[0,\infty)} e^{-x}dx = (-e^{-x})|_0^\infty = (0-(-1)) = 1.$$

- $$P_2(\mathbb{N}) = \int_{\mathbb{N}} \mathbb{1}_{\{0,1\}}(x)x^2\mu(dx) = \int_{\{0,1\}} x^2\mu(dx) = 0^2 \cdot \mu(\{0\}) + 1^2 \cdot \mu(\{1\}) = 0 \cdot 1 + 1 \cdot 1 = 1.$$

- **Tool 12.1** (Integral with respect to a dirac measure)**.**

  $$\int_{\Omega} f(x)\delta_\omega(dx) = f(\omega).$$

  $$P_3(\mathbb{R}) = \int_{\mathbb{R}} x^2\mu(dx) = \int_{\mathbb{R}} x^2(\delta_{-1}(dx) + \delta_1(dx)) = \int_{\mathbb{R}} x^2\delta_{-1}(dx) + \int_{\mathbb{R}} x^2\delta_1(dx)$$
  $$= (-1)^2 + 1^2 = 2.$$

We see that P_3 is not a probability measure on $\mathcal{B}$.

(d) Calculate:

1. λ Lebesgue measure on $\mathfrak{B}(\mathbb{R})$.

$$\int_{\mathbb{R}} \mathbb{1}_{[-1,1]}(x)\lambda(dx) = \int_{[-1,1]} 1\lambda(dx) = 1 \cdot \lambda([-1,1]) = 1 \cdot 2 = 2.$$

2. $P(A) = (1-p)\delta_0(A) + p\delta_1(A)$, $A \in \mathfrak{B}(\mathbb{R})$, $p \in (0,1)$.

$$\int_{\mathbb{R}} (x-p)^2P(dx) = \int_{\mathbb{R}} (x-p)^2((1-p)\delta_0(dx)) + \int_{\mathbb{R}} (x-p)^2(p\delta_1(dx)) = (0-p)^2(1-p) + (1-p)^2p$$
$$= p^2 - p^3 + p + p^3 - 2p^2 = p - p^2.$$

3. λ Lebesgue measure on $\mathfrak{B}(\mathbb{R})$. As the Lebesgue measure of a singleton is equal to 0

$$\int_{\mathbb{N}} \log(x)\lambda(dx) = 0.$$

(e) Refer to Def. 10.5, Prop. 10.2.

1. The support is $E = \{0,1\}$, countable.

$$P_1(E) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 = 1.$$

2. As F_X is continuous, $\mathbb{P}(X = x) = 0$, $\forall x \in \mathbb{R}$. This means that there exists no countable set E s.t. $P_X(E) = 1$.
3. The support is $E = \{0,1\}$, countable.

$$P_3(A) = \mathbb{P}(X = 1) \cdot \delta_1(A) + \mathbb{P}(X = 0) \cdot \delta_0(A).$$

Where $\mathbb{P}(X = 1) = \mathbb{P}(X^{-1}(1)) = \mathbb{P}(h)$.

P_2 is not a discrete law.

(f) TODO: Understand and complete. Cf. Sec. 11.3.

Exercise 12.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X be a discrete random variable on Ω with support $\{-1, 1\}$ and law

$$P_X(A) = \frac{1}{2}\delta_{-1}(A) + \frac{1}{2}\delta_1(A).$$

(a) $\mathbb{P}(X = -1) = P_X(\{-1\}) = \frac{1}{2}$, $\mathbb{P}(X = 1) = \frac{1}{2}$

(b) We have that $f(X) = |X|^2$, cf. Prop. 10.3

$$\mathbb{E}(|X|^2) = \int_{\{-1,1\}} |x|^2 P_X(dx) = |-1|^2 \cdot P_X(\{-1\}) + |1|^2 \cdot P_X(\{1\}) = 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1.$$

(c) $\mathbb{E}[X] = -\frac{1}{2} + \frac{1}{2} = 0$. We than know that

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 1 - 0 = 1.$$

(d) We can find the support of $\frac{X+1}{2}$.

$$\mathbb{P}\left(\frac{X+1}{2} = \omega\right) \neq 0 \Rightarrow X = 2\omega - 1 = \{-1, 1\}.$$

For $2\omega - 1 = -1$, $\omega = 0$, and for $2\omega - 1 = 1$, $\omega = 1$. The support of $\frac{X+1}{2}$ is $\{0, 1\}$. TODO: Finish explaining

(e)

Exercise 12.3. Let

$$\phi(x) = \begin{cases} 0 & x < -3, \\ \frac{1}{3} + \frac{1}{9}x & -3 \leq x < 0 \\ \frac{1}{3} - \frac{1}{9}x & 0 \leq x < 3 \\ 0 & x \geq 3. \end{cases}$$

(a) Verify that $\int_{\mathbb{R}} \phi(x)dx = 1$

$$\begin{aligned} \int_{\mathbb{R}} \phi(x)dx &= \int_{[-3,0)} \left(\frac{1}{3} + \frac{1}{9}x\right) dx + \int_{[0,3)} \left(\frac{1}{3} - \frac{1}{9}x\right) dx \\ &= \left[\frac{x}{3} + \frac{x^2}{18}\right]_{-3}^{0} + \left[\frac{x}{3} - \frac{x^2}{18}\right]_{0}^{3} \\ &= 0 - \left(-1 + \frac{1}{2}\right) + \left(1 - \frac{1}{2}\right) - 0 = 1 \end{aligned}$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X be a random variable on Ω with law $P_X(dx) = \phi(x)dx$.

(b) Find the distribution function F_X of X.

$$F_X(t) = \int_{[-3,t]} \phi(x).$$

For $t \in [-3, 0)$

$$F_X(t) = \int_{[-3,t]} \left(\frac{1}{3} + \frac{1}{9}x\right) dx = \left[\frac{x}{3} + \frac{x^2}{18}\right]_{-3}^{t} = \frac{t}{3} + \frac{t^2}{18} + \frac{1}{2}.$$

For $t \in [0,3)$

$$\begin{aligned} F_X(t) &= \int_{[-3,0)} \left(\frac{1}{3}+\frac{x}{9}\right) dx + \int_{[0,t)} \left(\frac{1}{3}-\frac{x}{9}\right) dx \\ &= \left[\frac{x}{3}+\frac{x^2}{18}\right]_{-3}^{0} + \left[\frac{x}{3}-\frac{x^2}{18}\right]_{0}^{t} \\ &= \frac{1}{2}+\frac{t}{3}-\frac{t^2}{18} \end{aligned}$$

.

We obtain

$$F_X(t) = \begin{cases} 0, & t < -3 \\ \frac{t}{3}+\frac{t^2}{18}+\frac{1}{2}, & -3 \le t < 0 \\ \frac{t}{3}-\frac{t^2}{18}+\frac{1}{2}, & 0 \le t < 3 \\ 1, & t \ge 3 \end{cases}.$$

(c) Calculate the expected value and the variance of X.

- Expected value

$$\begin{aligned} \mathbb{E}(X) &= \int_{\mathbb{R}} x\phi(x)dx = \int_{[-3,0)} \left(\frac{x}{3}+\frac{x^2}{9}\right) dx + \int_{[0,3)} \left(\frac{x}{3}-\frac{x^2}{9}\right) dx \\ &= \left[\frac{x^2}{6}+\frac{x^3}{27}\right]_{-3}^{0} + \left[\frac{x^2}{6}-\frac{x^3}{27}\right]_{0}^{3} = 0-(\frac{3}{2}-1)+\frac{3}{2}-1-0=0 \end{aligned}$$

.

- Variance

$$\begin{aligned} Var(X) = \mathbb{E}(X^2) &= \int_{[-3,0)} \left(\frac{x^2}{3}+\frac{x^3}{9}\right) dx + \int_{[0,3)} \left(\frac{x^2}{3}-\frac{x^3}{9}\right) dx \\ &= \left[\frac{x^3}{9}+\frac{x^4}{36}\right]_{-3}^{0} + \left[\frac{x^3}{9}-\frac{x^4}{36}\right]_{0}^{3} = 0-(-3+\frac{9}{4})+3-\frac{9}{4}-0=\frac{3}{2} \end{aligned}$$

.

(d) Show that $F_X|_{(-3,3)} : (-3,3) \to (0,1)$ is a bijection. We know that F_X is continuous on $\mathbb{R}$.
For $t \in (-3,0)$

$$F'_X(t) = \frac{1}{3}+\frac{t}{9} > \frac{1}{3}+-\frac{3}{9} = 0.$$

For $t \in [0,3)$

$$F'_X(t) = \frac{1}{3}-\frac{t}{9} > \frac{1}{3}-\frac{3}{9} = 0.$$

$F'_X(t)|_{(-3,3)} > 0$ for every $t \in (-3,3) \Rightarrow F_X(t)|_{(-3,3)}$ is monotonically increasing for every $t \in (-3,3) \Rightarrow$ $F_X(t)|_{(-3,3)}$ is bijective.

Exercise 12.4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X_1 and X_2 be two random variables on Ω that are independent with common law that is continuous uniform on the interval $[0,1]$. what is the probability density function of the random vector $Y = (X_1, 2\sqrt{X_2})$.

Refer to Prop. 11.1

$$\phi(y) = \phi_1(y_1)\phi_2(y_2), \quad y = (y_1, y_2) \in \mathbb{R}^2.$$

We know that

$$\phi_1(y_1) = \mathbb{1}_{[0,1]}(y_1), \quad y_1 \in \mathbb{R}.$$

For the probability density function of $2\sqrt{X_2}$, refer to Example 10.4.

By Prop. 10.3, we know that

$$\mathbb{E}(f(Y_2)) = \mathbb{E}(f(2\sqrt{X_2})) = \int_0^1 f(2\sqrt{x_2})dx_2.$$

We then substitute $y_2 = 2\sqrt{x_2}$, $\left(x_2 = \left(\frac{y_2}{2}\right)^2\right)$, $\left(dy_2 = x_2^{-\frac{1}{2}}dx_2 \Rightarrow dy_2 = \frac{2}{y_2}dx_2 \Rightarrow dx_2 = 2^{-1}y_2\right)$, $(x_2 \in [0,1] \Rightarrow y_2 \in [0,2])$.

$$\mathbb{E}(f(Y_2)) = \int_0^2 f(y_2)2^{-1}y_2dy_2 = \int_{\mathbb{R}} f(y_2)\mathbb{1}_{[0,2]}(y_2)2^{-1}y_2dy_2.$$

Hence, the law of Y_2 is given by

$$\phi(y_2) = \mathbb{1}_{[0,2]}(y_2)2^{-1}y_2, \quad y_2 \in \mathbb{R}.$$

In conclusion

$$\phi(y_1, y_2) = \mathbb{1}_{[0,1]}(y_1)\mathbb{1}_{[0,2]}(y_2)2^{-1}y_2.$$

Exercise 12.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X be a discrete random variable on Ω with support $\{1, \ldots, N\}$, where $N \geq 2$ and N is even. Suppose that X has law defined upon:

$$\mathbb{P}(X = k) = C_N max\{k, N-k\}, \quad k = 1, \ldots, N,$$

Where $C_N \in \mathbb{R}$. Find C_N.

As N is even, we can find a middle point $m = \frac{N}{2}$. I will use $N = 2m$.

$$\begin{aligned}
\sum_{k=1}^{N} \mathbb{P}(X = k) &= C_N(\sum_{k=1}^{m}(2m-k) + \sum_{k=m+1}^{2m} k) \\
&= C_N(\sum_{k=1}^{m}(2m-k) + \sum_{j=1}^{m}(m+j)) \\
&= C_N(2m \cdot m - \sum_{k=1}^{m} k + m \cdot m + \sum_{j=1}^{m} j) \\
&= C_N(3m^2) \\
&= C_N(3(\frac{N}{2})^2)
\end{aligned}$$

By definition

$$C_N(\frac{3N^2}{4}) = 1 \Rightarrow C_N = \frac{4}{3N^2}.$$

Chapter 13

Mock exam 2

Solve with the pdf of the mock exam on the side.

Notation: We recall some of the terminology:

- Given a nonempty set Ω, $\mathcal{P}(\Omega)$ is the power set on Ω;
- $\mathfrak{B}(\mathbb{R}^k)$ denotes the Borel σ-field on $\mathbb{R}^k$, $k \geq 1$;
- The measure

$$\mu(A) = \begin{cases} \#A, & \text{if } A \text{ is finite} \\ \infty, & \text{otherwise,} \end{cases} \qquad A \in \mathcal{P}(\Omega),$$

  is referred to as the counting measure on $\mathcal{P}(\Omega)$;
- Given a measurable space $(\Omega, \mathcal{F})$ and $x \in \Omega$, we write δ_x for the measure

$$\mathcal{F} \ni A \mapsto \delta_x(A) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{otherwise.} \end{cases}$$

Exercise 13.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X_1 and X_2 be two random variables on Ω that are independetn with common law that is continuous uniform on the interval $[0,1]$. what is the probability density function of the random vector $Y = \left(\frac{1+X_1}{2}, X_2\right)$?

We know that

$$\phi_2(y_2) = \mathbb{1}_{[0,1]}(y_2), \quad y_2 \in \mathbb{R}.$$

To find the probability density function of $\frac{1+X_1}{2}$, we know by Prop. 10.3 that

$$\mathbb{E}\left[f(Y_1)\right] = \mathbb{E}\left[f\left(\frac{1+X_1}{2}\right)\right] = \int_{[0,1]} f\left(\frac{1+x_1}{2}\right) dx_1.$$

We substitue $y_1 = \frac{1+x_1}{2}$. We also note that $x_1 = 2y_1 - 1$, and that $dx_1 = 2dy_1$. We also know that $x_1 \in [0,1] \Rightarrow y_1 \in \left[\frac{1}{2}, 1\right]$.

$$\int_{\left[\frac{1}{2},1\right]} f(y_1) 2dy_1 = \int_{\mathbb{R}} f(y_1) \mathbb{1}_{\left[\frac{1}{2},1\right]}(y_1) 2dy_1.$$

Hence, the law of Y_1 is given by

$$\phi_1(y_1) = 2 \times \mathbb{1}_{\left[\frac{1}{2},1\right]}(y_1) \quad y_1 \in \mathbb{R}.$$

In conclusion

$$\phi(y_1, y_2) = 2(\mathbb{1}_{\left[\frac{1}{2},1\right]}(y_1) \mathbb{1}_{[0,1]}(y_2)).$$