Tools of Modern Probability
Imre Péter Tóth
Exercise sheet 2, fall 2020

2.1 Define a σ-algebra as follows:

Definition 1 *For a nonempty set Ω, a family $\mathcal{F}$ of subsets of ω (i.e. $\mathcal{F} \subset 2^\Omega$, where $2^\Omega := \{A : A \subset \Omega\}$ is the power set of Ω) is called a σ-algebra over Ω if*

- $\emptyset \in \mathcal{F}$
- *if $A \in \mathcal{F}$, then $A^C := \Omega \setminus A \in \mathcal{F}$ (that is, $\mathcal{F}$ is closed under complement taking)*
- *if $A_1, A_2, \dots \in \mathcal{F}$, then $(\cup_{i=1}^\infty A_i) \in \mathcal{F}$ (that is, $\mathcal{F}$ is closed under countable union).*

Show from this definition that a σ-algebra is closed under countable intersection, and under finite union and intersection.

2.2 *Continuity of the measure*

(a) Prove the following:

Theorem 1 *(Continuity of the measure)*

i. *If $(\Omega, \mathcal{F}, \mu)$ is a measure space and $A_1, A_2, \dots$ is an increasing sequence of measurable sets (i.e. $A_i \in \mathcal{F}$ and $A_i \subset A_{i+1}$ for all i), then $\mu(\cup_{i=1}^\infty A_i) = \lim_{i\to\infty} \mu(A_i)$ (and both sides of the equation make sense).*

ii. *If $(\Omega, \mathcal{F}, \mu)$ is a measure space, $A_1, A_2, \dots$ is a decreasing sequence of measurable sets (i.e. $A_i \in \mathcal{F}$ and $A_i \supset A_{i+1}$ for all i) and $\mu(A_1) < \infty$, then $\mu(\cap_{i=1}^\infty A_i) = \lim_{i\to\infty} \mu(A_i)$ (and both sides of the equation make sense).*

(b) Show that in the second statement the condition $\mu(A_1) < \infty$ is needed, by constructing a counterexample for the statement when this condition does not hold.

2.3 (a) We toss a biased coin, on which the probability of heads is some $0 \le p \le 1$. Define the random variable ξ as the indicator function of tossing heads, that is

$$\xi := \begin{cases} 0, & \text{if tails} \\ 1, & \text{if heads} \end{cases} .$$

i. Describe the distribution of ξ (called the Bernoulli distribution with parameter p) in the "classical" way, listing possible values and their probabilities,

ii. and also by describing the distribution as a measure on $\mathbb{R}$, giving the weight $\mathbb{P}(\xi \in B)$ of every Borel subset B of $\mathbb{R}$.

iii. Calculate the expectation of ξ.

(b) We toss the previous biased coin n times, and denote by X the *number of heads* tossed.

i. Describe the distribution of X (called the Binomial distribution with parameters (n, p)) by listing possible values and their probabilities.

ii. Calculate the expectation of X by integration (actually summation in this case) using its distribution,

iii. and also by noticing that $X = \xi_1 + \xi_2 + \dots + \xi_n$, where ξ_i is the indicator of the i-th toss being heads, and using linearity of the expectation.

2.4 The *ternary* number $0.a_1a_2a_3\dots$ is the analogue of the usual decimal fraction, but writing numbers in base 3. That is, for any sequence $a_1, a_2, a_3, \dots$ with $a_n \in \{0,1,2\}$, by definition

$$0.a_1a_2a_3\cdots := \sum_{n=1}^{\infty} \frac{a_n}{3^n}.$$

Now let us construct the ternary fraction form of a random real number X via a sequence of fair coin tosses, such that we rule out the digit 1. That is,

$$a_n := \begin{cases} 0, & \text{if the } n\text{-th toss is tails,} \\ 2, & \text{if the } n\text{-th toss is heads} \end{cases},$$

and setting $X = 0.a_1a_2a_3\dots$ (ternary). In this way, X is a "uniformly" chosen random point of the famous *middle-third Cantor set* C defined as

$$C := \left\{ \sum_{n=1}^{\infty} \frac{a_n}{3^n}\,,\ a_n \in \{0,2\}\ (n = 1, 2, \dots) \right\}.$$

Show that

(a) The distribution of X gives zero weight to every point – that is, $\mathbb{P}(X = x) = 0$ for every $x \in \mathbb{R}$. (As a consequence, the cumulative distribution function of X is continuous.)

(b) The distribution of X is not absolutely continuous w.r.t the Lebesgue measure on $\mathbb{R}$.

2.5 Let V be a random vector in $\mathbb{R}^n$ with an n-dimensional standard Gaussian distribution, meaning that it has density

$$f(v_1, \dots, v_n) = \frac{1}{\sqrt{2\pi}^n} e^{-\frac{v_1^2 + \cdots + v_n^2}{2}}.$$

Think of V as the velocity vector of a particle with mass m, so the energy is $E = \frac{m}{2}V^2$. Calculate the distribution of the random variable E. (Meaning: calculate the distribution function and/or the density, and tell the name of the distribution.)

2.6 *Usefulness of the linearity of the expectation.* A building has 10 floors, not including the ground floor. On the ground floor, 10 people get into the elevator, and every one of them chooses a destination at random, uniformly out of the 10 floors, independently of the others. Let X denote the number of floors *on which the elevator stops* – i.e. the number of floors that were chosen by at least one person. Calculate the expectation and the variance of X. *(hint: First notice that the distribution of X is hard to calculate. Find a way to calculate the expectation and the variance without that.)*

2.7 Let $X = [0, 1]$ and let μ be Lebesgue measure on X. Let $f(x) = x^2$. Describe the measure $f_*\mu$

a.) by calculating $(f_*\mu)([a, b])$ for every interval $[a, b] \subset \mathbb{R}$

b.) by giving the density of $f_*\mu$ with respect to Lebesgue measure.

2.8 Let $X = \{(a_1, a_2, \dots) \,|\, a_k \in \{0, 1\} \text{ for every } k\}$ be the set of $\{0, 1\}$-sequences. Let μ be the measure on X for which

$$\mu(\{(a_1, a_2, \dots) \in X \,|\, a_1 = b_1, \dots, a_N = b_N\}) = \frac{1}{2^N}$$

for every $b_1, \dots, b_N \in \{0, 1\}$. Let $f : X \to \mathbb{R}$ be defined as

$$f(a_1, a_2, \dots) := \sum_{k=1}^{\infty} \frac{a_k}{2^k}.$$

Describe the measure $f_*\mu$

a.) by calculating $(f_*\mu)([a,b])$ for every interval $[a,b] \subset \mathbb{R}$

b.) by giving the density of $f_*\mu$ with respect to Lebesgue measure.

2.9 Consider the following measure spaces (X, μ):

I. $X = [0,1]$, μ is Lebesgue measure.

II. $X = [0, \infty)$, μ is Lebesgue measure.

III. $X = \{1, 2 \ldots, N\}$, μ is counting measure.

IV. $X = \{1, 2 \ldots\}$, μ is counting measure.

Show examples of functions $f_1, f_2, \ldots$ and f from X to $\mathbb{R}$ such that f_n converges to f

a.) almost everywhere, but not in L^1,

b.) in L^1, but not almost everywhere,

c.) in L^1, but not in L^2,

d.) in L^2, but not in L^1.

2.10 The characteristic function of a random variable X is the function $\Psi : \mathbb{R} \to \mathbb{C}$ defined as $\Psi(t) = \mathbb{E}(e^{itX})$. Calculate the characteristic function of

(a) The Bernoulli distribution $B(p)$

(b) The "pessimistic geometric distribution with parameter p" – that is, the distribution μ on $\{0, 1, 2 \ldots\}$ with weights $\mu(\{k\}) = (1-p)p^k$ $(k = 0, 1, 2 \ldots)$.

(c) The "optimistic geometric distribution with parameter p" – that is, the distribution ν on $\{1, 2, 3, \ldots\}$ with weights $\nu(\{k\}) = (1-p)p^{k-1}$ $(k = 1, 2 \ldots)$.

(d) The Poisson distribution with parameter λ – that is, the distribution η on $\{0, 1, 2 \ldots\}$ with weights $\eta(\{k\}) = e^{-\lambda}\frac{\lambda^k}{k!}$ $(k = 0, 1, 2 \ldots)$.

(e) The exponential distribution with parameter λ – that is, the distribution on $\mathbb{R}$ with density (w.r.t. Lebesgue measure)

$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0 \\ 0, & \text{if not} \end{cases} .$$

2.11 For a real values random variable X, the characteristic function of X is $\psi_X : \mathbb{R} \to \mathbb{C}$ defined as $\psi_X(t) := \mathbb{E}\left(e^{itX}\right)$, where $i \in \mathbb{C}$ is the imaginary unit. Show that $\psi_X(t)$ exists for every $t \in \mathbb{R}$.

2.12 For a probability distribution ν on $\mathbb{R}$, the characteristic function of ν is $\psi_\nu : \mathbb{R} \to \mathbb{C}$ defined as $\psi_\nu(t) := \int_{\mathbb{R}} e^{itx}\, \mathrm{d}\nu(x)$, where $i \in \mathbb{C}$ is the imaginary unit. Show that $\psi_\nu(t)$ exists for every $t \in \mathbb{R}$.

2.13 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $X : \Omega \to \mathbb{R}$ be a random variable and let $\nu = X_*\mathbb{P}$ be its distribution. Show that $\psi_X = \psi_\nu$, where ψ_X and ψ_μ are the characteristic functions defined in exercises 11 and 12.

2.14 *Dominated convergence and continuous differentiability of the characteristic function.* The Lebesgue dominated convergence theorem is the following

Theorem 2 (dominated convergence) *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f_1, f_2, \ldots$ measurable real valued functions on Ω which converge to the limit function pointwise, μ-almost everywhere. (That is, $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in \Omega$, except possibly for a set of x-es with μ-measure zero.) Assume furthermore that the f_n admit a common* integrable dominating function*: there exists a $g : \Omega \to \mathbb{R}$ such that $|f_n(x)| \le g(x)$ for every $x \in \Omega$ and $n \in \mathbb{N}$, and $\int_\Omega g \,\mathrm{d}\mu < \infty$. Then (all the f_n and also f are integrable and)*

$$\lim_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu = \int_\Omega f \,\mathrm{d}\mu.$$

Use this theorem to prove the following:

a.) **Theorem 3 (Continuity of the characteristic function, 1)** *For any real valued random variable X, its characteristic function $\psi_X(t) = \mathbb{E}(e^{itX})$ is continuous.*

b.) **Theorem 4 (Continuity of the characteristic function, 2)** *For any probability distribution ν on $\mathbb{R}$, its characteristic function $\psi_\nu(t) = \int_\mathbb{R} e^{itx} \,\mathrm{d}\nu(x)$ is continuous.*

c.) **Theorem 5 (Differentiability of the characteristic function, 1)** *Let X be a real valued random variable, its characteristic function $\psi_X(t) = \mathbb{E}(e^{itX})$. If X is integrable, then ψ_X is differentiable.*

d.) **Theorem 6 (Differentiability of the characteristic function, 2)** *Let ν be a probability distribution on $\mathbb{R}$, its characteristic function $\psi_\nu(t) = \int_\mathbb{R} e^{itx} \,\mathrm{d}\nu(x)$. If $\mathbb{E}\nu \in \mathbb{R}$, then ψ_ν is differentiable.*

e.) **Theorem 7 (Continuous differentiability of the characteristic function, 1)** *Let X be a real valued random variable, its characteristic function $\psi_X(t) = \mathbb{E}(e^{itX})$. If X is integrable, then ψ'_X is continuous.*

f.) **Theorem 8 (Continuous differentiability of the characteristic function, 2)** *Let ν be a probability distribution on $\mathbb{R}$, its characteristic function $\psi_\nu(t) = \int_\mathbb{R} e^{itx} \,\mathrm{d}\nu(x)$. If $\mathbb{E}\nu \in \mathbb{R}$, then ψ'_ν is continuous.*

2.15 *Exchangeability of integral and limit.* Consider the sequences of functions $f_n : [0,1] \to \mathbb{R}$ and $g_n : [0,1] \to \mathbb{R}$ concerning their pointwise limits and the limits of their integrals. Do there exist integrable functions $f : [0,1] \to \mathbb{R}$ and $g : [0,1] \to \mathbb{R}$, such that $f_n(x) \to f(x)$ and $g_n(x) \to g(x)$ for Lebesgue almost every $x \in [0,1]$? What is $\lim_{n\to\infty} \left(\int_0^1 f_n(x)dx \right)$ and $\lim_{n\to\infty} \left(\int_0^1 g_n(x)dx \right)$? Are the conditions of the dominated and monotone convergence theorems and the Fatou lemma satisfied? If yes, what do these theorems ensure about these specific examples?

(a)

$$f_n(x) = \begin{cases} n^2 x & \text{if } 0 \le x < 1/n, \\ 2n - n^2 x & \text{if } 1/n \le x \le 2/n, \\ 0 & \text{otherwise.} \end{cases}$$

(b) Write n as $n = 2^k + l$, where $k = 0, 1, 2 \ldots$ and $l = 0, 1, \ldots, 2^k - 1$ (this can be done in a unique way for every n). Now let

$$g_n(x) = \begin{cases} 1 & \text{if } \frac{l}{2^k} \le x < \frac{l+1}{2^k}, \\ 0 & \text{otherwise.} \end{cases}$$

2.16 *Exchangeability of integrals.* Consider the following function $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x) = \begin{cases} 1 & \text{if} \quad 0 < x,\, 0 < y \text{ and } 0 \le x - y \le 1, \\ -1 & \text{if} \quad 0 < x,\, 0 < y \text{ and } 0 < y - x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Calculate $\int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{+\infty} f(x,y)dx \right) dy$ and $\int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{+\infty} f(x,y)dy \right) dx$. What's the situation with the Fubini theorem?

2.17 Let λ be Lebesgue measure and χ be counting measure on $\mathbb{R}$ (with the Borel σ-algebra). Show that λ does not have a density with respect to χ. (Hint: consider 1-element sets.)

2.18 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $A \in \mathcal{F}$. Define $X : \Omega \to \mathbb{R}$ as $X(\omega) = \mathbf{1}_A(\omega)$ and let $\mu = X_*\mathbb{P}$ be the distribution of X. Show that μ is absolutely continuous w.r.t counting measure, show that it also has a density. What is the density?

2.19 Let X be a discrete random variable and let μ be its distribution. Give the density of μ w.r.t. counting measure.