Probability 1
CEU Budapest, fall semester 2018
Imre Péter Tóth
Homework sheet 2 – solutions

2.1 (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 the old "probability 1" definition, using its distribution,

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

2.2 The *ternary* number $0.a_1a_2a_3\ldots$ is the analogue of the usual decimal fraction, but writing numbers in base 3. That is, for any sequence $a_1, a_2, a_3, \ldots$ 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 $X = 0.a_1a_2a_3\ldots$ (ternary). In this way, X is a "uniformly" chosen random point of the famous *middle-third Cantor set* C defined as

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

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 $\mathbb{R}$.

2.3 **(homework)** In the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $\Omega = [0, 1]$, let $\mathcal{F}$ be the Borel σ-algebra and let $\mathbb{P}$ be the Lebesgue measure (restricted to $\mathcal{F}$). Let the random variable $X : \Omega \to \mathbb{R}$ be defined as

$$X(\omega) := \begin{cases} \ln \omega, & \text{if } \omega \neq 0 \\ 0, & \text{if } \omega = 0 \end{cases} .$$

(a) Show that X is measurable as a function $X : \Omega \to \mathbb{R}$ when Ω is equipped with the Borel σ-algebra $\mathcal{F}$ and $\mathbb{R}$ is also equipped with its Borel σ-algebra $\mathcal{B}$. *(Remark: This exercise is only for those interested in every mathematical detail. It is not at all as important as it may seem. You are also welcome to just believe that X is measurable.)*

(b) Let μ be the distribution of X, which means that μ is the measure on $(\mathbb{R}, \mathcal{B})$ defined by

$$\mu(A) := \mathbb{P}(\{\omega \in \Omega \,|\, X(\omega) \in A\}) \quad \text{for all } A \in \mathcal{B}.$$

(In other words, μ is the push-forward of the measure $\mathbb{P}$ to $\mathbb{R}$ by X.)

"Describe" the measure μ by calculating $F(x) := \mu((-\infty, x])$ for every $x \in \mathbb{R}$. Also calculate $\mu([a, b])$ for every interval $[a, b] \subset \mathbb{R}$ (with $a \leq b$).

(This function $F : \mathbb{R} \to [0, 1]$ is called the (cumulative) distribution function of the measure μ, or also the (cumulative) distribution function of the random variable X.)

Solution:

(a) We will use a lemma.

Lemma 1. *If Ω_1, Ω_2 are nonempty sets, $f : \Omega_1 \to \Omega_2$ is any function and $\mathcal{F} \subset 2^{\Omega_1}$ is a σ-algebra over Ω_1, then*

$$\mathcal{G} := \{B \subset \Omega_2 \,|\, f^{-1}(B) \in \mathcal{F}\}$$

is a σ-algebra over Ω_2.

Proof. We check the definition.

- $f^{-1}(\emptyset) = \emptyset \in \mathcal{F}$, so $\emptyset \in \mathcal{G}$
- If $B \in \mathcal{G}$ then $f^{-1}(B) \in \mathcal{F}$. Since $\mathcal{F}$ is a σ-algebra, this means

$$f^{-1}(\Omega_2 \setminus B) = f^{-1}(\Omega_2) \setminus f^{-1}(B) = \Omega_1 \setminus f^{-1}(B) \in \mathcal{F}$$

  as well, so $\Omega_2 \setminus B \in \mathcal{G}$.
- If $B_1, B_2, \dots \in \mathcal{G}$ then $f^{-1}(B_i) \in \mathcal{F}$ for all i. Since $\mathcal{F}$ is a σ-algebra, this means

$$f^{-1}\left(\cup_{i=1}^{\infty} B_i\right) = \cup_{i=1}^{\infty} f^{-1}(B_i) \in \mathcal{F}$$

  as well, so $\cup_{i=1}^{\infty} B_i \in \mathcal{G}$.

□

This has a trivial corollary:

Corollary 1. *Let Ω_1, Ω_2 be nonempty sets, $f : \Omega_1 \to \Omega_2$ a function and $\mathcal{F} \subset 2^{\Omega_1}$ a σ-algebra over Ω_1. Let $H \subset \Omega_2$ be a family of sets. If $f^{-1}(B) \in \mathcal{F}$ holds for every $B \in H$, then it also holds for every $B \in \sigma(H)$, where $\sigma(H)$ is the σ-algebra generated by H.*

Proof. Let $\mathcal{G} := \{B \subset \Omega_2 \,|\, f^{-1}(B) \in \mathcal{F}\}$. Then $\mathcal{G}$ is a σ-algebra by the lemma, which contains H by assumption, so it also contains $\sigma(H)$. □

We apply the corollary with $\Omega_1 = \Omega$, $\Omega_2 = \mathbb{R}$, $\mathcal{F} = \mathcal{F}$, $f = X$ and $H = \{B \subset \mathbb{R} \,|\, B \text{ is open}\}$. Then $\sigma(H) = \mathcal{B}$, so to get the measurability of X, it is enough to check that inverse images of open sets are (Borel) measurable. This is obvious, since X is continuous except at the single point 0, so the inverse image of an open set is also an open set, plus possibly a point.

(Remark: The same argument works with $H := \{\text{intervals}\}$, since then $\sigma(H) = \mathcal{B}$ as well.)

(b) since $Leb(\{0\}) = 0$, the single point $\omega = 0$ where X is defined separately, plays no role and can be ignored. Using the definitions,

$$F(x) = \mu((-\infty, x]) = \mathbb{P}(X \in (-\infty, x]) = Leb(\{\omega \in (0,1] \,|\, \ln \omega \leq x\})$$
$$= Leb(\{\omega \in (0,1] \,|\, \omega \leq e^x\}) = \begin{cases} Leb((0,1]) = 1 & \text{if } x \geq 0 \text{ (so } e^x \geq 1) \\ Leb((0,e^x]) = e^x & \text{if } x < 0 \text{ (so } e^x < 1) \end{cases}.$$

Summary:

$$F(x) = \begin{cases} e^x & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}.$$

This $F(x)$ is continuous, so $\mu(\{a\}) = 0$ for every $a \in \mathbb{R}$, so for all $a \leq b$

$$\mu([a,b]) = \mu((a,b]) = F(b) - F(a) = \begin{cases} e^b - e^a & \text{if } b < 0 \\ 1 - e^a & \text{if } a < 0 \leq b \\ 0 & \text{if } 0 \leq a \end{cases}.$$

2.4 Let χ be the counting measure on $\mathbb{N}$. Calculate $\int_{\mathbb{N}} f \,\mathrm{d}\chi$ if $f : \mathbb{N} \to \mathbb{R}$ is given by

a.) $f(k) := \frac{1}{2^k}$

b.) $f(k) := \frac{1}{k}$

c.) $f(k) := \frac{(-1)^k}{k}$

2.5 Let χ be the counting measure on $\mathbb{R}$ and μ be Lebesgue measure on $\mathbb{R}$.

a.) Show that μ is absolutely contuinuous w.r.t. χ: $\mu \ll \chi$.

b.) Show that μ does not have a density f w.r.t. χ: there is no such f that $\mu(B) = \int_B f \,\mathrm{d}\chi$ would hold for every (Borel) $B \subset \mathbb{R}$.

c.) What's wrong with the Ranod-Nikodym theorem?

2.6 Let χ be the counting measure on $\mathbb{N}$ and let the measure μ be absolutely continuous with respect to χ, with density $f(k) := q^k p$, where $p \in (0,1)$ and $q = 1 - p$. Define $X : \mathbb{N} \to \mathbb{R}$ as $X(k) := k$.

a.) Calculate $\int_{\mathbb{N}} X \,\mathrm{d}\mu$.

b.) Calculate $\int_{\mathbb{N}} X^2 \,\mathrm{d}\mu$.

2.7 **(homework)** Let μ be a measure on $\mathbb{R}$ which has density $f(x) := x^2$ with respect to Lebesgue measure. Let ν be a measure on $\mathbb{R}$ which has density $g(x) := \sqrt{x}$ with respect to μ. Calculate $\nu([1,3])$.

Solution: By the definition of the density ind its usage in integrals:

$$\nu([1,3]) = \int_{[1,3]} g \,\mathrm{d}\mu = \int_{[1,3]} gf \,\mathrm{d}Leb = \int_1^3 \sqrt{x} x^2 \,\mathrm{d}x = \left[\frac{2}{7} x^{\frac{7}{2}}\right]_1^3 = 2\frac{27\sqrt{3} - 1}{3}.$$

2.8 **(homework)** Let the random variable X have density

$$f(x) = \begin{cases} 2e^{-2x} \text{ if } x > 0 \\ 0 \text{ if not} \end{cases},$$

with respect to Lebesgue measure on $\mathbb{R}$.

a.) Show that this f is indeed the density (w.r.t. Lebesgue) of a probability distribution.

b.) Let $Y := X^2$. Show that Y is also absolutely continuous w.r.t. Lebesgue measure and find its density.

Solution:

a.) We need to check three things:

- f is measurable: don't worry about this. (Actually, f is piecewise continuous, so the argument of Exercise 3a works.)
- $f \geq 0$ almost everywhere (actually everywhere).
- $\int_R f \,\mathrm{d}Leb = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_0^{\infty} 2e^{-2x}\,\mathrm{d}x = [-e^{-2x}]_0^{\infty} = 1.$

b.) We will find the density by differentiating the distribution function $F_Y(y) := \mathbb{P}(Y \leq y)$. Clearly $X \geq 0$ almost surely, so $Y \geq 0$ almost surely, so $F_Y(y) = 0$ for $y < 0$. For $y \geq 0$, from the definitions

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq \sqrt{y}) = \int_{(-\infty,\sqrt{y}]} f \,\mathrm{d}Leb = \int_0^{\sqrt{y}} f(x)\,\mathrm{d}x.$$

This would be easy to calculate explicitly, but we don't need it: let $f_Y(y) := F'_Y(y)$ wherever F_Y is differentiable: $f_Y(y) = 0$ for $y < 0$, and for $y > 0$

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} \int_0^{\sqrt{y}} f(x)\,\mathrm{d}x = f(\sqrt{y})\frac{1}{2\sqrt{y}} = e^{2\sqrt{y}}\frac{1}{2\sqrt{y}}.$$

Summary: the function

$$f_Y(y) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{\sqrt{y}} e^{2\sqrt{y}} & \text{if } x > 0 \end{cases}$$

has the property that $\mathbb{P}(Y \in (a,b]) = F_Y(b) - F_Y(a) = \int_{(a,b]} f_Y \,\mathrm{d}Leb$ for every $a < b$, so it is indeed the density of Y with respect to Leb. The existence of the density in turn implies that Y is absolutely continuous.

2.9 *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. Help: What is the probability that the elevator stops on the first floor?)*

2.10 We take a huge bag. 1 minute before midnight we put 10 balls (numbered $1 \dots 10$) into the bag. Then we draw a ball from the bag at random, and throw it away. $\frac{1}{2}$ minute before midnight we put another 10 balls (numbered $11 \dots 20$) into the bag. Then we draw a ball from the bag at random, and throw it away. $\frac{1}{4}$ minute before midnight we put another 10 balls (numbered $21 \dots 30$) into the bag. Then we draw a ball from the bag at random, and throw it away. And so on, infinitely many times: $\frac{1}{2^n}$ minute before midnight we put 10 balls (numbered $(10n+1) \dots (10n+10)$) into the bag. Then we draw a ball from the bag at random, and throw it away.

a.) What is the probability that ball number 1 will be in the bag at midnight? *(Hint: we will see later that* $\lim_{N\to\infty} \prod_{n=0}^{N} \left(1 - \frac{1}{9n+10}\right) = 0$.*)*

b.) What is the probability that ball number 11 will be in the bag at midnight?

c.) Show that, at midnight, with probability 1, the bag will be empty. (What?!?!)

2.11 *The Fatou lemma* is the following

Theorem 1. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and f_1, f_2, ... a sequence of measureabale functions $f_n : \Omega \to \mathbb{R}$, which are nonneagtive, e.g. $f_n(x) \geq 0$ for every $n = 1, 2, \ldots$ and every $x \in \Omega$. Then*

$$\int_\Omega \liminf_{n\to\infty} f_n(x)\,\mathrm{d}\mu(x) \leq \liminf_{n\to\infty} \int_\Omega f_n(x)\,\mathrm{d}\mu(x)$$

(and both sides make sense).

Show that the inequality in the opposite direction is in general false, by choosing $\Omega = \mathbb{R}$, μ as the Lebesgue measure on $\mathbb{R}$, and constructing a sequence of nonnegative $f_n : \mathbb{R} \to \mathbb{R}$ for which $f_n(x) \xrightarrow{n\to\infty} 0$ for every $x \in \mathbb{R}$, but $\int_\mathbb{R} f_n(x)\,\mathrm{d}x \geq 1$ for all n.

2.12 **(homework)** *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? (For the Fatou lemma, see the lecture notes or Exercise 11.)

(a)

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

(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

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

Solution: Please **draw the graph of these functions!** You will see that

(a) $f_n(x) \to f(x) := 0$ for every $x \in [0,1]$, but $\int_0^1 f_n(x)dx = 1$ for every n (except for $n = 1$), so

$$\lim_{n\to\infty} \left(\int_0^1 f_n(x)dx\right) = 1 \neq 0 = \int_0^1 f(x)dx.$$

The convergence is not monotone, so the monotone convergence theorem says nothing. A natural choice for the common dominating function would be $G(x) := \frac{1}{x}$ (see the graphs you have drawn), but this is not integrable, because $\int_0^1 \frac{1}{x}dx = \infty$. Indeed, there is no integrable dominating function, so the dominated convergence theorem says nothing as well. However, the functions are non-negative, so the conditions of the Fatou lemma hold. Of course, the statement also holds:

$$\int_0^1 \liminf_{n\to\infty} f_n(x)dx = 0 \leq 1 = \liminf_{n\to\infty} \int_0^1 f_n(x)dx.$$

(b) For every fixed $x \in [0,1]$ as n grows, $g_n(x)$ will be 0 most of the time, but it will also be 1 once in a while (infinitely many times). So $g_n(x)$ is not convergent (as $n \to \infty$) for any x: no limiting g exists. Thus the monotone and dominated convergence theorems say nothing (there is no $\int_0^1 g(x)dx$ to converge to). However, the conditions of the Fatou lemma hold. Of course, the statement also holds: $\liminf_{n\to\infty} g_n(x) = 0$ for every x, while $\int\limits_0^1 g_n(x)dx = \frac{1}{2^k}$ for $2^k \le n \le 2^{k+1}$, so $\int\limits_0^1 g_n(x)dx \xrightarrow{n\to\infty} 0$. The statement of the lemma now reads

$$\int\limits_0^1 \liminf_{n\to\infty} g_n(x)dx = 0 \le 0 = \liminf_{n\to\infty} \int\limits_0^1 g_n(x)dx.$$

2.13 *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?