Tools of Modern Probability
Imre Péter Tóth
Exercise sheet 3

3.1 Let V be an inner product space over $\mathbb{R}$ and let $f : V \to \mathbb{R}$ be a linear form. Let $E := \{y \in V \mid f(y) = 0\}$ be the null-space of f. Suppose that $f(a) = 1$, $c \in E$ and $a - c$ is orthogonal to E, meaning $(a - c)y = 0$ for every $y \in E$. Now, for any $x \in V$, find the $\lambda \in \mathbb{R}$ for which $x_1 := x - \lambda(a - c) \in E$. Use this to get the relation between $f(x)$ and $(a - c)x$.

3.2 Represent the following functions $f : V \to \mathbb{R}$ as multiplication by a fixed vector, whenever this is possible due to the Riesz representation theorem.

a.) $V = \mathbb{R}^{10}$ with the usual inner product, $f((x_1, \dots, x_{10})) := x_5$ (evaluation at 5)

b.) $V = \mathbb{R}^{10}$ with the usual inner product, $f((x_1, \dots, x_{10})) := x_6 - x_5$ (discrete derivative at 5).

c.) $V = \mathbb{R}^{10}$ with the usual inner product, $f((x_1, \dots, x_{10})) := x_6 - 2x_5 + x_4$ (discrete second derivative at 5).

d.) $V = l^2 := \{x : \mathbb{N} \to \mathbb{R} \mid \sum_{i=1}^{\infty} x^2(i) < \infty\}$, with the inner product $x \cdot y := \sum_{i=1}^{\infty} x(i)y(i)$; $f(x) := \sum_{i=1}^{100} x(i)$.

e.) $V = l^2 := \{x : \mathbb{N} \to \mathbb{R} \mid \sum_{i=1}^{\infty} x^2(i) < \infty\}$, with the inner product $x \cdot y := \sum_{i=1}^{\infty} x(i)y(i)$; $f(x) := \sum_{i=1}^{\infty} x(i)$.

f.) $V = l^2 := \{x : \mathbb{N} \to \mathbb{R} \mid \sum_{i=1}^{\infty} x^2(i) < \infty\}$, with the inner product $x \cdot y := \sum_{i=1}^{\infty} x(i)y(i)$; $f(x) := \sum_{i=1}^{\infty} x^2(i)$.

g.) $V = L^2([0,1]) := \{x : [0,1] \to \mathbb{R} \mid \int_0^1 x^2(t)\,\mathrm{d}t < \infty\}$, with the inner product $x \cdot y := \int_0^1 x(t)y(t)\,\mathrm{d}t$; $f(x) := x(\frac{1}{2})$ (evaluation at $\frac{1}{2}$).

h.) $V = L^2([0,1]) := \{x : [0,1] \to \mathbb{R} \mid \int_0^1 x^2(t)\,\mathrm{d}t < \infty\}$, with the inner product $x \cdot y := \int_0^1 x(t)y(t)\,\mathrm{d}t$; $f(x) := x'(\frac{1}{2})$ (derivative at $\frac{1}{2}$).

i.) $V = L^2([0,1]) := \{x : [0,1] \to \mathbb{R} \mid \int_0^1 x^2(t)\,\mathrm{d}t < \infty\}$, with the inner product $x \cdot y := \int_0^1 x(t)y(t)\,\mathrm{d}t$; $f(x) := \int_{0.2}^{0.7} x(t)\,\mathrm{d}t$.

j.) $V = \{x : [0,1] \to \mathbb{R} \mid \int_0^1 x^2(t)\,\mathrm{d}t < \infty,\ f \text{ is differentiable}\}$, with the inner product $x \cdot y := \int_0^1 x(t)y(t)\,\mathrm{d}t$; $f(x) := x'(\frac{1}{2})$.

k.) $V = \{x : [0,1] \to \mathbb{R} \mid \int_0^1 x^2(t)\,\mathrm{d}t < \infty,\ f \text{ is continuous}\}$, with the inner product $x \cdot y := \int_0^1 x(t)y(t)\,\mathrm{d}t$; $f(x) := x(\frac{1}{2})$.

l.) $V = \{x : [0,1] \to \mathbb{R} \mid \int_0^1 x^2(t)\,\mathrm{d}t < \infty,\ f \text{ is continuous}\}$, with the inner product $x \cdot y := \int_0^1 x(t)y(t)\,\mathrm{d}t$; $f(x) := \int_{0.2}^{0.7} x(t)\,\mathrm{d}t$.

3.3 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.

3.4 (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 $X = \xi_1 + \xi_2 + \cdots + \xi_n$, where ξ_i is the indicator of the i-th toss being heads, and using linearity of the expectation.

3.5 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 $\mathbb{R}$.

3.6 *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 $A_1, A_2, \ldots$ 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,* $A_1, A_2, \ldots$ *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.

3.7 *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.)*