Probability 1
CEU Budapest, fall semester 2013
Imre Péter Tóth
Homework sheet 1 – solutions

1.1 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*

(i) $\emptyset \in \mathcal{F}$

(ii) if $A \in \mathcal{F}$, then $A^C := \Omega \setminus A \in \mathcal{F}$ (that is, $\mathcal{F}$ is closed under complement taking)

(iii) 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.

Solution:
If $B_1, B_2, \dots \in \mathcal{F}$ then $A_i := \Omega \setminus A_i \in \mathcal{F}$ as well, for $i = 1, 2, \dots$ due to (1ii), and thus $C := (\cup_{i=1}^\infty A_i) \in \mathcal{F}$ by (1iii). Finally, $\Omega \setminus C \in \mathcal{F}$ by (1ii), but $\Omega \setminus C = \cap_{i=1}^\infty B_i$ by the basics of set algebra, so we have shown that $\mathcal{F}$ is closed under countable intersection. For finite union, notice that if $A_1, A_2, \dots, A_n \in \mathcal{F}$, then we can choose $A_{n+1} = A_{n+2} = \dots = \emptyset \in \mathcal{F}$ by (1i), to get $(\cup_{i=1}^n A_i) = (\cup_{i=1}^\infty A_i) \in \mathcal{F}$ by (1iii). So $\mathcal{F}$ is shown to be closed under finite union. Closedness under finite intersection can be seen similarly.

1.2 (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 distirbution 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.

Solution:

(a) i. The possible values are 0 and 1, their probabilities are $\mathbb{P}(\xi = 0) = 1 - p$ and $\mathbb{P}(\xi = 1) = p$.

ii. $\mu(B) = \mathbb{P}(\xi \in B) = \begin{cases} 1, & \text{if } 0 \in B \text{ and } 1 \in B, \\ 1-p, & \text{if } 0 \in B \text{ but } 1 \notin B, \\ p, & \text{if } 1 \in B \text{ but } 0 \notin B, \\ 0, & \text{if } 0 \notin B \text{ and } 1 \notin B. \end{cases}$

iii. $\mathbb{E}\xi = 0 \cdot \mathbb{P}(\xi = 0) + 1 \cdot \mathbb{P}(\xi = 1) = 0 \cdot (1-p) + 1 \cdot p = p.$

(b) i. The possible values are $0, 1, 2, \ldots, n$, their probabilities are

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n.$$

ii. If we denote the distribution of X by μ, then

$$\mathbb{E}X = \int_{\mathbb{R}} x \,\mathrm{d}\mu(x) = \sum_{k=0}^{n} k \cdot \mu(\{k\}) = \sum_{k=0}^{n} k \cdot \mathbb{P}(X = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}.$$

To calculate this sum, one of the many ways is to consider the two-variable function

$$f(u, v) := \sum_{k=0}^{n} k \binom{n}{k} u^k v^{n-k}.$$

Then what we want to know is $\mathbb{E}X = f(p, 1-p)$, but of course we are even more happy if we can calculate $f(u, v)$ for every (u, v). Now we notice that

$$f(u, v) = u \frac{\partial}{\partial u} g(u, v) \text{ where } g(u, v) = \sum_{k=0}^{n} \binom{n}{k} u^k v^{n-k}.$$

This is now easy: by the binomial theorem $g(u, v) = (u + v)^n$, so

$$f(u, v) = u \frac{\partial}{\partial u} (u + v)^n = nu(u + v)^{n-1},$$

and

$$\mathbb{E}X = f(p, 1-p) = np(p + 1 - p)^n = np.$$

iii. This is much easier:

$$\mathbb{E}X = \mathbb{E}(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} \mathbb{E}\xi_i = \sum_{i=1}^{n} p = np.$$

1.3 *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.

Solution: The standard counterexample is

$$f_n(x) := \begin{cases} 1, & \text{if } n \leq x \leq n+1, \\ 0, & \text{if not.} \end{cases}$$

The phenomenon behind the counteraxample – as often – is that exchangeability of integral and limit can fail if mass "escapes to infinity".

1.4 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 $\mathbb{R}$.

Solution:

(a) Similarly to a decimal expansion, the ternary expansion of a real number $x \in [0,1]$ is essentially unique: every x can be written in the form $x = 0.a_1a_2a_3\dots$ in only one, or possibly two ways. (There are actually two ways for some rational numbers, since e.g. $0.1022222\dot{2} = 0.1100000\dot{0}$.) However, every individual sequence $a_1, a_2, a_3, \dots$ has probability $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \dots = 0$, so every x is given weight at most twice zero which is still zero.

(b) The distribution of X cannot be absolutely continuous w.r.t. Lebesgue measure, since it gives positive measure to C ($\mathbb{P}(X \in C) = 1$), which has Lebesgue measure zero ($\mathrm{Leb}(C) = 0$). To see that the Lebesgue measure of C is indeed zero, notice that the set in the n-t level of the construction of C,

$$C_n := \left\{ \sum_{k=1}^{\infty} \frac{a_k}{3^k}\,,\ a_k \in \{0,2\} \text{ for } k = 1,2,\dots,n \text{ but } a_k \in \{0,1,2\} \text{ for } k \geq n+1 \right\},$$

has Lebesgue measure

$$\mathrm{Leb}(C_n) = \left(\frac{2}{3}\right)^n .$$

Now $C \subset C_n$ for every $n \in \mathbb{N}$, so

$$\mathrm{Leb}(C) \leq \mathrm{Leb}(C_n) = \left(\frac{2}{3}\right)^n \quad \text{for every } n,$$

which implies that $\mathrm{Leb}(C) = 0$.

(Actually, this means not only that the distribution μ of X is not absolutely continuous w.r.t. Lebesgue measure, but that the two measures are singular w.r.t each other, which means that $\mathbb{R}$ can be decomposed into two disjoint subsets (namely C and $\mathbb{R} \setminus C$,) such that one is "unseen" by one measure ($\mathrm{Leb}(C) = 0$), while the other is "unseen" by the other measure ($\mu(\mathbb{R} \setminus C) = 0$).)