Probability 1
CEU Budapest, fall semester 2013
Imre Péter Tóth
Homework sheet 2 – solutions

2.1 **(homework)** Exercise 3 of "Homework sheet 1", delayed from last week (unless already done)

Solution: See the solutions of "Homework sheet 1".

2.2 **(homework)** Exercise 4 of "Homework sheet 1", delayed from last week (unless already done)

Solution: See the solutions of "Homework sheet 1".

2.3 *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.

2.4 *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.)*

2.5 Calculate the characteristic function of

(a) The Bernoulli distribution $B(p)$ (see Homework sheet 1)

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

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

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

Solution:

$$\psi_{Poi(\lambda)}(t) = \sum_{k=0}^{\infty} e^{itk} e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} e^{itk} \eta(\{k\}) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

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

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

Solution:

$$\phi_{Exp(\lambda)}(t) = \int_{\mathbb{R}} e^{itx} f_\lambda(x) \,\mathrm{dLeb}(x) = \int_0^\infty e^{itx} \lambda e^{-\lambda x} \,\mathrm{d}x = \lambda \left[\frac{e^{(it-\lambda)x}}{it-\lambda} \right]_0^\infty = \frac{\lambda}{\lambda - it}.$$

2.6 **(homework)** Calculate the characteristic function of the normal distribution $\mathcal{N}(m, \sigma^2)$. (Remember the definition from the old times: $\mathcal{N}(m, \sigma^2)$ is the distribution on $\mathbb{R}$ with density (w.r.t. Lebesgue measure)

$$f_{m,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

You can save yourself some paperwork if you only do the calculation for $\mathcal{N}(0, 1)$ and reduce the general case to this using the relation between different normal distributions. You can and should use the fact that

$$\int_{-\infty}^\infty f_{m,\sigma^2}(x) \,\mathrm{d}x = 1$$

for every m and σ.

Solution: First we reduce the general case to the case of the standard normal distribution using the fact (known from old times, easy to check from the formulas) that if $X \sim \mathcal{N}(0, 1)$ and $Y = m + \sigma X$, then $Y \sim \mathcal{N}(m, \sigma^2)$. As a result, the characteristic function for the normal distribution with expectation m and variance σ^2 is

$$\psi_{\mathcal{N}(m,\sigma^2)}(t) = \mathbb{E}(e^{itY}) = \mathbb{E}(e^{itm+it\sigma X}) = e^{itm}\mathbb{E}(e^{i(t\sigma)X}) = e^{itm}\psi_{\mathcal{N}(0,1)}(\sigma t), \tag{1}$$

where $\psi_{\mathcal{N}(0,1)}(t) := \mathbb{E}(e^{itX})$ is the characteristic function of the standard normal distribution.

Now we go on to calculate

$$\begin{aligned} \psi_{\mathcal{N}(0,1)}(t) &= \mathbb{E}(e^{itX}) = \int_{-\infty}^\infty e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \,\mathrm{d}x = \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2-2itx}{2}} \,\mathrm{d}x = \\ &= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-it)^2-(it)^2}{2}} \,\mathrm{d}x = e^{-\frac{t^2}{2}} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-it)^2}{2}} \,\mathrm{d}x. \end{aligned}$$

We use the substitution $y := x - it$ to get

$$\psi_{\mathcal{N}(0,1)}(t) = e^{-\frac{t^2}{2}} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \,\mathrm{d}x = e^{-\frac{t^2}{2}}.$$

In the last step we used that the standard normal density function (just like every probability density function) integrates to 1. Writing this back to (1), we get the final result

$$\psi_{\mathcal{N}(m,\sigma^2)}(t) = e^{itm} e^{-\frac{(\sigma t)^2}{2}}.$$

Remark: The substitution $y = x - it$ is not completely trivial to make rigorous. In fact, with this substitution, while x runs over the real line, y will run over a line in the complex plane, namely the line γ of complex numbers with imaginary part $-it$, so leaving the boundaries as $-\infty$ and ∞ after the substitution is cheating. To make the argument precise, one has to show that the integral on γ is equal to the integral on the real line. This is a typical application of a standard, but strong tool of complex analysis, called the *residue theorem*. I will not go into that here, and I don't expect the students to do so either.

2.7 *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 everywehere. (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

Theorem 3 (differentiability of the characteristic function) *Let X be a real valued random variable, $\psi(t) = \mathbb{E}(e^{itX})$ its characteristic function and $n \in \mathbb{N}$. If the n-th moment of X exists and is finite (i.e. $\mathbb{E}(|X|^n) < \infty$), then ψ is n times continuously differentiable and*

$$\psi^{(k)}(0) = i^k \mathbb{E}(X^k), \quad k = 0, 1, 2, \ldots, n.$$