Workshop 3

January 22-28, 2024

For this week's submitted solutions, you are encouraged to try exercises at the level of 4, 6, 7 or 9 here.

1 Short and basic exercises about integration

Exercise 1. Let $(X, \mathcal{F}, \mu)$ be a measure space and f an element of $m(\mathcal{F})^+$. Show that if $\mu(f) = 0$ then $\mu(\{f > 0\}) = 0$.

Exercise 2. Consider $f \in (m\mathcal{F})^+$ and $(f_n)_{n\in\mathbb{N}} \subset (m\mathcal{F})^+$ such that $f_n \uparrow f$, except for a μ-null set N (this means $\mu(N) = 0$). Show that $\mu(f_n) \uparrow \mu(f)$.

Exercise 3. Use the standard machine to prove the linearity of the integral: for any integrable functions f, g and any $\alpha, \beta \in \mathbb{R}$,

$$\mu(\alpha f + \beta g) = \alpha\mu(f) + \beta\mu(g).$$

2 Approximation by simpler functions and an implication of measurability

Exercise 4. Let $(\Omega, \mathcal{F}, P)$ be a probability triple.

1. Let X be a real-valued, bounded measurable function. Show that for all $\epsilon > 0$ there exists a finite partition $(A_i)_{i=1}^N \subset \mathcal{F}$ of Ω (which means that these sets are pairwise disjoint and their union is Ω) and real numbers $(\lambda_i)_{i=1}^N$, such that

$$\sup_{\omega\in\Omega} \left| X(\omega) - \sum_{i=1}^{N} \lambda_i I_{A_i}(\omega) \right| < \epsilon. \tag{1}$$

   Here I_A denotes the indicator function of A: $I_A(\omega) = 1$ if and only if $\omega \in A$.

2. Let $Y : \Omega \to \mathbb{R}$ be a random variable. Use the previous result to show that if X is a bounded $\sigma(Y)$-measurable function, then $X = f(Y)$ for some Borel function $f : \mathbb{R} \to \mathbb{R}$. (Hint: Apply the insights from the standard machine.)

3 Convergence of integrals

Read first the statements and proofs of Fatou's lemma and Lebesgue's dominated convergence theorem.

Exercise 5 (cf. Williams E5.1)**.** Consider the triple $([0, 1], \mathcal{B}([0, 1]), \lambda)$, where λ is the Lebesgue measure. Define $f_n = nI_{[0,1/n)}$. Prove that $f_n \to f$ almost surely (a.s.) on $[0, 1]$, but that $\lambda(f_n) = 1$ for every $n \in \mathbb{N}$. Draw a picture of $g = \sup_n |f_n|$ and show that $g \notin \mathcal{L}^1([0, 1], \mathcal{B}([0, 1]), \lambda))$.

Exercise 6. Given a probability triple $(\Omega, \mathcal{F}, P)$ and a random variable X on it, we introduce the function[1]

$$\phi(t) := P(e^{itX}) = \int_\Omega e^{itx} \mathrm{d}P(x). \tag{2}$$

Prove that ϕ is continuous on $\mathbb{R}$. (Hint: You may need that, for functions between metric spaces, continuity is equivalent to sequential continuity.)

4 Relative entropy a.k.a. Kullback-Leibler divergence

Let $(X, \mathcal{F})$ be a measurable space and μ, ν two (σ-)finite measures on $(X, \mathcal{F})$. If there exists an $f \in \mathcal{L}^1(X, \mathcal{F}, \nu)$ such that $\mu(A) = \int_A f d\nu$ for every $A \in \mathcal{F}$, we say that f is the Radon-Nikodym derivative of μ with respect to ν; this is denoted $f = \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$.

Exercise 7. Let $(\Omega, \mathcal{F}, \mu)$ be a probability triple. For every probability measure ρ on $(\Omega, \mathcal{F})$ absolutely continuous with respect to μ, we introduce the quantity

$$D(\rho|\mu) = \int_\Omega \log \frac{\mathrm{d}\rho}{\mathrm{d}\mu} \, \mathrm{d}\rho, \tag{3}$$

called *the relative entropy of ρ with respect to μ*. Here we interpret log as an extended function from $[0, \infty)$ to $[-\infty, \infty)$, such that $\log 0 = -\infty$. Remember that $0 \le \frac{\mathrm{d}\rho}{\mathrm{d}\mu} < \infty$ almost surely.

1. Show that

$$\int_\Omega \left| \log \frac{\mathrm{d}\rho}{\mathrm{d}\mu} \right| \, \mathrm{d}\rho = \int_{\{\frac{\mathrm{d}\rho}{\mathrm{d}\mu} > 0\}} \left| \log \frac{\mathrm{d}\rho}{\mathrm{d}\mu} \right| \, \mathrm{d}\rho \tag{4}$$

   and that if any of these integrals converge, then

$$D(\rho|\mu) = \int_{\{\frac{\mathrm{d}\rho}{\mathrm{d}\mu} > 0\}} \log \frac{\mathrm{d}\rho}{\mathrm{d}\mu} \, \mathrm{d}\rho. \tag{5}$$

2. (possibly more difficult) Show that if $\mu \ll \nu \ll \xi$, then

$$\frac{\mathrm{d}\mu}{\mathrm{d}\xi} = \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \frac{\mathrm{d}\nu}{\mathrm{d}\xi}. \tag{6}$$

   Two measures μ and ρ are called *equivalent* (denoted $\mu \sim \rho$) if $\mu \ll \rho$ and $\rho \ll \mu$. Use (6) to prove that if $\mu \sim \rho$, then

$$\frac{\mathrm{d}\rho}{\mathrm{d}\mu} = \left(\frac{\mathrm{d}\mu}{\mathrm{d}\rho} \right)^{-1} \tag{7}$$

3. Show that if $\rho \sim \mu$ and the integrals in (4) are finite, $D(\rho|\mu) \ge 0$. (Hint: log is concave, so...; you may use (7); be explicit with the necessary conditions of inequalities.)

[1] By definition, a function $h : \Omega \to \mathbb{C}$ is integrable (with respect to a measure μ) if its real and imaginary parts are integrable (as real-valued functions); in this case,

$$\mu(h) = \mu(\operatorname{Re} h) + i\mu(\operatorname{Im} h).$$

5 Independence again

Exercise 8. (an easy one, but good warm-up) Let X and Y be independent Bernoulli random variables, both with parameter $1/2$.

1. Compute the probability mass functions of $X+Y$ and $|X-Y|$.
2. Show that $X+Y$ and $|X-Y|$ are dependent.
3. Show that $X+Y$ and $|X-Y|$ are uncorrelated. (This means that $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.)

Exercise 9.

1. Let X be a real-valued random variable with law Λ_X and cumulative distribution function (c.d.f.) F_X, and Y a real-valued random variable with law Λ_Y and c.d.f. F_Y. We denote by $\Lambda_{X,Y}$ the law of the joint variable (X,Y) and by $F_{X,Y}$ the corresponding c.d.f.:
$$F_{X,Y} := \Lambda_{X,Y}((-\infty, x] \times (-\infty, y]). \tag{8}$$
Prove that the following statements are equivalent:
   (a) X and Y are independent.
   (b) $\Lambda_{X,Y} = \Lambda_X \otimes \Lambda_Y$.
   (c) $F_{X,Y}(x,y) = F_X(x)F_Y(y)$.
2. Show that if (X,Y) has a joint probability density function $f_{X,Y}$ (with respect to the Lebesgue measure $\lambda^{\otimes 2}$), then some of the statements above (therefore each of them) is equivalent to
   (a) $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for $\lambda^{\otimes 2}$-almost every (x,y).

Remember that $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}y$, etc.
(Hint: Uniqueness lemma!)