Homework 2: Probability

Isaac Martin

Last compiled September 30, 2022

§ *Chapter 1*

EXERCISE 2.1 (Stronger separation). Let $(S, \mathcal{S}, \mu)$ be a measure space and let $f, g \in \mathcal{L}^0(S, \mathcal{S})$ satisfy $\mu(\{x \in S : f(x) < g(x)\}) > 0$. Prove or construct a counterexample for the following statement:

"There exist constants $a, b \in \mathbb{R}$ such that $\mu(\{x \in S : f(x) \leq a < b \leq g(x)\}) > 0$."

Proof: We prove that this statement is true. Let $A = \{x \in S \mid f(x) < g(x)\}$ denote the set in questions, set $I = \{(a, b) \in \mathbb{Q}^2 \mid a < b\}$ and for each pair of rational numbers $(a, b) \in I$, define

$$B_{a,b} = \{x \in S \mid f(x) \leq a < b \leq g(x)\}.$$

Note that $B_{a,b}$ is countable for each $(a, b) \in I$ since both f and g are measurable functions and $B_{a,b} = f^{-1}((-\infty, a]) \cap g^{-1}([b, \infty))$. I claim that

$$A = \bigcup_{(a,b) \in I} B_{a,b} =: B$$

Note first that because B is the countable union of measurable sets, it too is measurable. Suppose that $x \in A$. Then $f(x) < g(x)$ and there exists some $a \in \mathbb{Q}$ such that $f(x) < a < g(x)$ because $\mathbb{Q}$ is dense in $\mathbb{R}$. Similarly, there exists some $b \in \mathbb{Q}$ such that $f(x) < a < b < g(x)$, and hence $x \in B_{a,b} \subseteq B$. Now suppose that $x \in B$. Then there exist some pair $(a, b) \in I$ such that $x \in B_{a,b} \implies f(x) \leq a < b \leq g(x)$. In particular, this means $f(x) < g(x)$ so $x \in A$. We therefore have that $A = B$.

Now suppose that $\mu(B_{a,b}) = 0$ for each pair of rational numbers $(a, b) \in I$. We would then have that

$$\mu(A) = \mu(B) = \mu\left(\bigcup_{(a,b) \in I} B_{a,b}\right) \leq \sum_{(a,b) \in I} \mu(B_{a,b}) = \infty \cdot 0 = 0.$$

Since $\mu(A) > 0$, it must be the case that $\mu(B_{a,b}) > 0$ for some $(a, b) \in I$.This proves the claim. $\square$

EXERCISE 2.2 (A uniform distribution on a circle.) Let S^1 be the unit circle and let $f : [0, 1) \longrightarrow S^1$ be the "winding map"

$$f(x) = (cos(2\pi x), \sin(2\pi x)), \ x \in [0, 1).$$

(1) Show that the map f is $(\mathcal{B}([0, 1)), \mathcal{S}^1)$-measurable, where $\mathcal{S}^1$ denotes the Borel σ-algebra on S^1 (with topology inherited from $\mathbb{R}^2$).

(2) For $\alpha \in (0, 2\pi)$, let R_α denote the (counter-clockwise) rotation of $\mathbb{R}^2$ with center $(0, 0)$ and angle α> Show that $R_\alpha(A) = \{R_\alpha(x) : x \in A\}$ is in $\mathcal{S}^1$ if and only if $A \in \mathcal{S}^1$.

(3) Let μ^1 be the pushforward of the Lebesgue measure λ by the map f. Show that μ^1 is rotation-invariant, i.e. that $\mu^1(A) = \mu^1(R_\alpha(A))$. *Note:* The measure μ^1 is called the **uniform measure** (or the **uniform distribution** on S^1).

Proof:

(1): If this were a topology class, we'd simply state that "it is clear that f is continuous," as it is a continuous map in each component. Instead, we will prove that it is continuous, and hence Borel measurable. We take for granted the continuity of $\sin$ and $\cos$ as functions on $\mathbb{R}$.

Suppose $x, a \in [0, 1)$, and consider $\|f(x) - f(a)\|^2$. With the help of trig identities, we have the following:

$$\begin{aligned}
\|f(x) - f(a)\|^2 &= |(\cos(2\pi x) - \cos(2\pi a))^2 + (\sin(2\pi x) - \sin(2\pi a))^2| \\
&= |\cos^2(2\pi x) - 2\cos(2\pi x)\cos(2\pi a) + \cos^2(2\pi a) + \sin^2(2\pi x) \\
&\quad - 2\sin(2\pi x)\sin(2\pi a) + \sin^2(2\pi a)| \\
&= |2 - \cos(2\pi x - 2\pi a) - \cos(2\pi x + 2\pi a) - \cos(2\pi x - 2\pi a) + \cos(2\pi x + 2\pi a)| \\
&= 2 - 2\cos(2\pi x - 2\pi a).
\end{aligned}$$

Note that we may drop the absolute value in the final equality since $2\cos(2\pi x - 2\pi a) \leq 2$ for all $x, a \in [0, 1)$. Thus, as x approaches a in $[0, 1)$, we have that

$$\lim_{x \to a} \|f(x) - f(a)\| = \lim_{x \to a} \big(2 - 2\cos(2\pi x - 2\pi a)\big) = 2 - 2\cos(0) = 0,$$

and hence f is continuous and therefore Borel measurable.

(2): I claim that R_α is a homeomorphism on $\mathbb{R}^2$, from which it will follow immediately that it induces a bijection on $\mathcal{S}^1$. First, notice that rotation any point $x \in \mathbb{R}^2$ first by $\alpha \in (0, 2\pi)$ and then by $2\pi - \alpha$ gives back x, i.e. $R_{2\pi-\alpha} \circ R_\alpha = \mathrm{id}_{\mathbb{R}^2}$. To see this more rigorously, we can realize $R_\alpha : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ as the $\mathbb{R}$-linear map given by left multiplication by

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

in which case the composition R_α with $R_{2\pi-\alpha}$ is the matrix product

$$\begin{aligned}
\begin{pmatrix} \cos(2\pi - \alpha) & -\sin(2\pi - \alpha) \\ \sin(2\pi - \alpha) & \cos(2\pi - \alpha) \end{pmatrix} \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} &= \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \\
&= \begin{pmatrix} \cos^2(\alpha) + \sin^2(\alpha) & -\sin(\alpha)\cos(\alpha) + \sin(\alpha)\cos(\alpha) \\ -\sin(\alpha)\cos(\alpha) + \sin(\alpha)\cos(\alpha) & \sin^2(\alpha) + \cos^2(\alpha) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}$$

We have a similar result for the composition $R_\alpha \circ R_{2\pi-\alpha}$. Since linear maps are continuous on $\mathbb{R}^2$ (this is a fact from undergraduate analysis that I feel doesn't warrant proof) R_α is a continuous map with continuous inverse, and is hence a homeomorphism.

Finally, note that R_α fixes S^1, which was implicitly assumed by the problem statement.

Now suppose that $A \subseteq S^1$ is an open set. This means there must be some open set $U \subseteq \mathbb{R}^2$ such that $A = S^1 \cap U$. Since R_α is a homeomorphism on $\mathbb{R}^2$, $R_\alpha(U) = R^{-1}_{2\pi-\alpha}(U)$, which is open by the continuity of $R_{2\pi-\alpha}$. Since R_α fixes S^1,

$$R_\alpha(A) = R_\alpha(U \cap S^1) = R_\alpha(U) \cap S^1 = R^{-1}_{2\pi-\alpha}(U) \cap S^1,$$

which is open in the subspace topology on S^1. Likewise, if $R_\alpha(A)$ is open, then $R^{-1}_{2\pi-\alpha}(R_\alpha(A)) = A$ is open.

The Borel algebra on S^1 is generated by open sets, and since the maps $A \mapsto R_\alpha(A)$ and $R_\alpha(A) \mapsto A$ send open sets to open (and hence measurable) sets, by Proposition 1.10 in the notes we conclude that R_α induces a bisection on $\mathcal{S}^1$.

(3): Fix $\alpha \in (0, 2\pi)$ and define a new measure μ^1_α on S^1 by setting $\mu^1_\alpha(A) = \mu^1(R_\alpha(A))$. Note that this is actually the pullback measure $R_{2\pi-\alpha,*}\mu^1$, since by part (2) $R_\alpha(A) = R^{-1}_{2\pi-\alpha}(A)$, so μ^1_α is indeed a measure on S^1. Let $\mathcal{P}$ denote the set of all open arcs of S^1, or equivalently the collection of all open connected subsets of S^1. We prove that $\mu^1(A) = \mu^1_\alpha(A)$ for all $A \in \mathcal{P}$.

Let $A \subseteq \mathcal{P}$ be an arc in S^1 and suppose that $(1,0) \notin A$. Then $f^{-1}(A) = (a,b) \subseteq [0,1)$ for some $a, b \in \mathbb{R}$, and hence

$$\mu_1(A) = \lambda(f^{-1}(A)) = \lambda((a,b)) = b - a.$$

Now consider μ^1_α. First note that for $x \in [0,1)$ we have

$$\begin{aligned} R_\alpha(f(x)) &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(2\pi x) \\ \sin(2\pi x) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha)\cos(2\pi x) - \sin(\alpha)\sin(2\pi x) \\ \sin(\alpha)\cos(2\pi x) + \cos(\alpha)\sin(2\pi x) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha + 2\pi x) \\ \sin(\alpha + 2\pi x) \end{pmatrix} \\ &= \begin{cases} f\left(\frac{\alpha}{2\pi} + x\right) & \text{if } \frac{\alpha}{2\pi} + x < 1 \\ f\left(\frac{\alpha}{2\pi} + x - 1\right) & \text{otherwise} \end{cases}. \end{aligned}$$

This means that if $(1,0) \notin R_\alpha(A)$, then

$$\mu^1_\alpha(A) = \lambda\left(\left(\frac{\alpha}{2\pi} + a, \frac{\alpha}{2\pi} + b\right)\right) = b - a.$$

If $(1,0) \in R_\alpha(A)$, then

$$\mu^1_\alpha(A) = \lambda\left(\left(\frac{\alpha}{2\pi} + a, 1\right) \cup [0, \frac{\alpha}{2\pi} + b - 1]\right) = b - a.$$

In either case, $\mu^1(A) = \mu^1_\alpha(A)$.

If $(1,0) \in A$, then $A \setminus \{(1,0)\}$ is the disjoint union $B \cup C$ of two arcs B and C, neither of which contains $(1,0)$. By the first case, $\mu^1(B) = \mu^1_\alpha(B)$ and $\mu^1(C) = \mu^1_\alpha(C)$. Furthermore, $\mu^1(\{(1,0)\}) = \lambda(f^{-1}(\{(1,0)\})) = 0 = \mu^1_\alpha(\{(1,0)\})$ so by additivity,

$$\mu^1(A) = \mu^1(B) + \mu^1(C) + \mu^1(\{(1,0)\}) = \mu^1_\alpha(B) + \mu^1_\alpha(C) + \mu^1_\alpha(\{(1,0)\}) = \mu^1_\alpha(A).$$

We conclude that $\mu^1(A) = \mu^1_\alpha(A)$ for all $A \in \mathcal{P}$.

Notice that the intersection of open arcs is still an open arc, hence $\mathcal{P}$ is a π-system. Furthermore, $\Lambda = \{A \in \mathcal{S} \mid \mu^1(A) = \mu^1_\alpha(A)\}$ is a λ-system. Since $\mathcal{P} \subseteq \Lambda$, by Dynkin's $\pi - \lambda$ Theorem, $\sigma(\mathcal{P}) \subseteq \mathcal{L}$. However, the set of all open arcs is a basis for the subspace topology on S^1 inherited from $\mathbb{R}^2$, hence $\mathcal{S} \subseteq \Lambda$. Hence $\mu^1 = \mu^1_\alpha$, and because α was chosen arbitrarily, we conclude that the pushforward of Lebesgue measure on S^1 is rotation invariant. $\square$

EXERCISE 2.3 (A change-of-variable formula). Let $(S, \mathcal{S}, \mu)$ and $(T, \mathcal{T}, \nu)$ be two measurable spaces, and let $F : S \longrightarrow T$ be a measurable function with the property that $\nu = F_* \mu$ (i.e., ν is the push-forward of μ through F). Show that for every $f \in \mathcal{L}^0_+(T, \mathcal{T})$ or $\mathcal{L}^1(T, \mathcal{T})$, we have

$$\int f \, d\nu = \int (f \circ F) \, d\mu.$$

Proof: The following is a procedure roughly matching the standard Lebesgue yoga.

First, we notice that if f is a simple function given by $f(x) = \alpha_i$ for $x \in A_i \in \mathcal{S}$ with $1 \leq i \leq n$, then $f \circ F$ is a simple function defined by $f \circ F(x) = \alpha_i$ for $x \in F^{-1}(A_i)$. Then

$$\int f \, d\nu = \sum_{i=1}^{n} \alpha_i \nu(A_i) = \sum_{i=1}^{n} \alpha_i \mu(F^{-1}(A_i)) = \int (f \circ F) d\mu,$$

so the desired equality holds for simple functions.

Now suppose that $f \in \mathcal{L}^0_+(T, \mathcal{T})$. By the simple approximation theorem (3.10 in the notes), we may find an increasing sequence of nonnegative simple functions f_n which uniformly approach f. Likewise, $f_n \circ F$ is an increasing sequence of simply functions (by what we proved above) which approaches $f \circ F$ from below. By monotone convergence, we then immediately get

$$\int f \, d\nu = \lim_{n \to \infty} \int f_n \, d\mu = \lim_{n \to \infty} \int f \circ F \, d\mu = \int f \circ F \, d\mu.$$

Finally, let $f \in \mathcal{L}^1(T, \mathcal{T})$ be an arbitrary Lebesgue integrable function and let f^+, f^- denote the typical $\mathcal{L}^0_+(T, \mathcal{T})$ functions representing the positive and negative portions of f. I argue that $(f \circ F)^+ = f^+ \circ F$ and $(f \circ F)^- = f^- \circ F$. Indeed, there is almost nothing to check:

$$(f \circ F)^+(x) = \max\{0, f(F(x))\} = f^+(F(x)),$$

and we have something similar for f^-. Since f^+, f^- are both $\mathcal{L}^0_+(T, \mathcal{T})$, by what we have already shown we have

$$\int f \, d\nu = \int f^+ \, d\nu \; + \; \int f^- \, d\nu = \int (f \circ F)^+ \, d\mu \; + \; \int (f \circ F)^- \, d\mu = \int (f \circ F) \, d\mu,$$

which concludes the proof. $\square$

EXERCISE 2.4 (An integrability criterion). Let $(S, \mathcal{S}, \mu)$ be a finite measure space, and let $f \in \mathcal{L}^0_+$. Show that

$$\int f d\mu < \infty \text{ if and only if } \sum_{n \in \mathbb{N}} \mu(\{f \geq n\}) < \infty$$

where, as usual, $\{f \geq n\} = \{x \in S \; : \; f(x) \geq n\}$. *Hint:* Approximate f from below and from above by a piecewise constant function.

Proof: First, some setup. Define $A_n = \{f \geq n\} \subseteq S$ for $n \in \mathbb{N}$. Note that this is a decreasing sequence, $A_n \supseteq A_{n+1}$, and that because $f \in \mathcal{L}^0_+$ we have $S = A_0$. Now define $B_n = A_n \setminus A_{n+1} = A_n \cap (A^c_{n+1})$; we'll think of B_n as the "outer shell" of A_n. Since each A_n is measurable, so is B_n. Furthermore, for each $x \in S$, if we set $k = \lfloor f(x) \rfloor$ to be the ceiling of $f(x)$, then $k \leq f(x) < k + 1$ and hence $x \in A_n$ but $x \notin A_{k+1}$. This means $x \in B_k$, and so $\{B_n\}_{n \in \mathbb{N}}$ forms a pairwise disjoint cover of S, i.e. a partition.

We'll prove both implications via contrapositive. Suppose first that $\sum_{n\in\mathbb{N}} \mu(A_n) = \infty$. Define a sequence of simple functions $g_n : S \longrightarrow \mathbb{R}$ with $B_0, ..., B_n$ as their level sets:

$$g_n(x) = \begin{cases} k & x \in B_k \text{ where } 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}.$$

This is well defined: $g_n(x)$ doesn't have contradictory definitions since $B_i \cap B_j = \emptyset$ whenever $i \neq j$ so g_n, and g_n is defined on all of S since $\{B_n\}_{n\in\mathbb{N}}$ covers S. For $x \in B_k$ and $n \geq k$, we have by definition that $f(x) \geq k = g_n(x)$, hence

$$\int f \, d\mu \geq \int g_n \, d\mu = \int g_n \, d\mu = \sum_{k=0}^{n} k\mu(B_k).$$

The above equality follows immediately from the definition of an integral of a simple function. We may take limits as this inequality doesn't depend on n, which gives us

$$\begin{aligned} \int f \, d\mu \geq \lim_{n\to\infty} \int g_n \, d\mu &= \sum_{k=1}^{\infty} k\mu(B_k) \\ &= \sum_{k=0}^{\infty} k(\mu(A_k) \setminus \mu(A_{k+1})) \\ &= \sum_{k=0}^{\infty} k\mu(A_k) - (k-1)\mu(A_k) \\ &= \sum_{k=0}^{\infty} \mu(A_k) = \sum_{k\in\mathbb{N}} \mu(\{f \geq n\}) \; - \; \mu(S). \end{aligned}$$

Since $\mu(S)$ is finite and $\sum_{k\in\mathbb{N}} \mu(\{f \geq n\})$ is infinite, we get that $\int g_n \, d\mu \longrightarrow \infty$ and hence $\int f \, d\mu = \infty$ as well.

Now suppose $\int f \, d\mu = \infty$. Using the same A_k and B_k as before, we shift the $g_n : S \longrightarrow \mathbb{R}$ we used previously up by one:

$$g_n(x) = \begin{cases} k+1 & x \in B_k \text{ where } 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}.$$

Now for $x \in B_k$ we get $f(x) < k+1 = g_n(x)$. However, it is not the case that $\int f \, d\mu \leq \int g_n d\mu$, as we'd like, since g_n is zero outside of $B_0 \cup ... \cup B_n$. To fix this, define $f_n : S \longrightarrow \mathbb{R}$ by

$$f_n(x) = f(x) \cdot 1_{B_0 \cup ... \cup B_n}.$$

Then $f_n \in \mathcal{L}^0_+$ for each $n \in \mathbb{N}$, $f_0(x) \leq f_1(x) \leq f_2(x) \leq ...$ and $\lim_{n\to\infty} f_n(x) \longrightarrow f(x)$ for all $x \in S$, so f_n satisfies the properties of the monotone convergence theorem and gives us

$$\lim_n \int f_n \, d\mu = \int f \, d\mu.$$

More importantly, f_n is less than g_n on $B_0 \cup ... \cup B_n$ and is zero everywhere else, giving us

$$\int f_n \, d\mu \leq \int g_n \, d\mu.$$

Since this is true of all n we can take limits to get

$$\begin{aligned}
\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu \leq \lim_{n\to\infty} \int g_n \, d\mu &= \sum_{k=0}^{\infty} (k+1)\mu(B_k) \\
&= \sum_{k\in\mathbb{N}} (k+1)(\mu(A_k) - \mu(A_{k+1})) \\
&= \sum_{k\in\mathbb{N}} (k+1)\mu(A_k) - (k)\mu(A_k) \\
&= \sum_{k\in\mathbb{N}} \mu(\{f \geq k\}).
\end{aligned}$$

Since $\infty = \int f d\mu \leq \sum_{k\in\mathbb{N}} \mu(\{f \geq k\})$, we conclude that $\sum_{k\in\mathbb{N}} \mu(\{f \geq k\}) = \infty$, proving the second implication of the problem.

Note: what we've really proven here is that $\sum_{n\in\mathbb{N}} \mu(\{f \geq n\}) \; - \; \mu(S) \leq \int f \, d\mu \leq \sum_{n\in\mathbb{N}} \mu(\{f \geq n\})$. From this inequality, it is clear that $\int f \, d\mu = \infty \implies \sum_{n\in\mathbb{N}} \mu(\{f \geq n\}) = \infty$, and that the reverse implication is true when $\mu(S) < \infty$. $\square$

EXERCISE 2.5

Proof: $\sim$ no time – that's all folks$\sim$

$\square$