Solution Sheet 2

Exercise 2.1.

Let $\mathcal{X}$ be a complete separable metric space. Prove that any probability measure on $\mathcal{X}$ is tight. *Hint: A complete, totally bounded set is compact.*

Proof. Let μ be a probability measure on $\mathcal{X}$, and fix an $\varepsilon > 0$ with the intention of finding a compact $K \subset \mathcal{X}$ such that $\mu(K) > 1-\varepsilon$. By separability there exists a countable dense subset (x_i) of $\mathcal{X}$. We define $B_r(y)$ as the open ball of radius r centred at $y \in \mathcal{X}$, and for any $n \in \mathbb{N}$ consider the collection of sets $(B_{\frac{1}{n}}(x_i))$. Due to density of the (x_i) then this collection covers $\mathcal{X}$ and in particular

$$\mu\left(\bigcup_{i=1}^{\infty} B_{\frac{1}{n}}(x_i)\right) = 1.$$

Moreover the sequence

$$\mu\left(\bigcup_{i=1}^{j} B_{\frac{1}{n}}(x_i)\right)$$

is monotonically increasing and convergent to 1 as $j \to \infty$, hence there exists an n_j such that

$$\mu\left(\bigcup_{i=1}^{n_j} B_{\frac{1}{n}}(x_i)\right) > 1 - \frac{\varepsilon}{2^n}.$$

Now we set

$$A := \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{n_j} B_{\frac{1}{n}}(x_i)$$

and wish to show that the closure $\bar{A}$ is the desired K. The first task is to show that $\bar{A}$ is compact; recall that a complete, totally bounded set is compact. As $\bar{A}$ is a closed subset of a complete metric space then it is itself complete, so we are only required to show that it is totally bounded. To show that A is totally bounded we must verify that for every $\delta > 0$, there exists finitely many $y_1, \ldots, y_k \in \mathcal{X}$ such that $A \subset \bigcup_{i=1}^{k} B_\delta(y_i)$. Observe that for $x \in A$, then $x \in \bigcup_{i=1}^{n_j} B_{\frac{1}{n}}(x_i)$ for every n. We choose an n large enough such that $\frac{1}{n} < \delta$, so that for every x_i, $B_{\frac{1}{n}}(x_i) \subset B_\delta(x_i)$. Thus,

$$x \in \bigcup_{i=1}^{n_j} B_{\frac{1}{n}}(x_i) \subset \bigcup_{i=1}^{n_j} B_\delta(x_i)$$

so one can take $k = n_j$, $y_i = x_i$ in the definition of totally bounded. The closure of a totally bounded set is totally bounded, hence $\bar{A}$ is totally bounded and complete so compact. To conclude the proof it only remains to show that $\mu(\bar{A}) > 1-\varepsilon$, or equivalently that $\mu(\bar{A}^C) < \varepsilon$. Indeed,

$$\begin{aligned}
\mu(\bar{A}^C) &\leq \mu(A^C) \\
&= \mu\left(\bigcup_{n=1}^{\infty}\left(\bigcup_{i=1}^{n_j} B_{\frac{1}{n}}(x_i)\right)^C\right) \\
&\leq \sum_{n=1}^{\infty} \mu\left(\left(\bigcup_{i=1}^{n_j} B_{\frac{1}{n}}(x_i)\right)^C\right) \\
&< \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} \\
&= \varepsilon.
\end{aligned}$$

□

Exercise 2.2.

Let μ, ν be two probability measures supported on an interval $[a,b] \subset \mathbb{R}$. Show that moments separate the measures, that is if for every non-negative integer k we have that

$$\int_a^b x^k d\mu = \int_a^b x^k d\nu$$

then $\mu = \nu$.

Proof. Recall Theorem 2.2.2 of the notes, which asserts that if for every bounded and uniformly continuous function $f : [a,b] \to \mathbb{R}$ we have

$$\int_a^b f d\mu = \int_a^b f d\nu$$

then $\mu = \nu$. Every continuous function on the compact set $[a,b]$ is bounded and uniformly continuous hence it is sufficient to verify this condition for only continuous f which we now fix. By the Stone-Weierstrass Theorem there exists a sequence of polynomials (p_n) such that $p_n \to f$ uniformly on $[a,b]$. From the assumption that the moments of the measures agree, we have that

$$\int_a^b p_n d\mu = \int_a^b p_n d\nu \tag{1}$$

and now claim that $\int_a^b p_n d\mu \to \int_a^b f d\mu$ as $n \to \infty$. Indeed this follows from the Dominated Convergence Theorem with dominating function $|f|+1$ which exceeds p_n for n large enough such that $\sup_{x\in[a,b]}\|p_n(x) - f(x)\| < 1$. Of course the same is true for ν, but referring to (1) then $\int_a^b p_n d\mu \to \int_a^b f d\nu$ as $n \to \infty$ as well, so by uniqueness of limits we have proven that

$$\int_a^b f d\mu = \int_a^b f d\nu$$

as required to demonstrate that $\mu = \nu$. □

Exercise 2.3.

Let μ, ν be two probability measures on $\mathbb{R}$ such that for every non-negative integer k, we have that

$$\int_{\mathbb{R}} x^k d\mu = \int_{\mathbb{R}} x^k d\nu =: \alpha_k.$$

In addition suppose that there exists an $r > 0$ such that for all $s \in [0, r]$, the infinite sum

$$\sum_{k=1}^{\infty} \frac{\alpha_k s^k}{k!}$$

is well defined in $\mathbb{R}$. Demonstrate the following:

1. Defining $\beta_k := \int_{\mathbb{R}} |x|^k d\mu$, we have that for any $s \in (0, r)$, $\frac{\beta_k s^k}{k!} \to 0$ as $k \to \infty$;
2. You are given that for any $x \in \mathbb{R}$,

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \frac{|x|^{n+1}}{(n+1)!}. \tag{2}$$

   Denote the characteristic function of μ by ϕ, that is for any $t \in \mathbb{R}$,

$$\phi(t) := \int_{\mathbb{R}} e^{itx} d\mu.$$

   Then for any $t \in \mathbb{R}$ and $h \in \mathbb{R}$ with $|h| < r$,

$$\phi(t+h) = \sum_{k=0}^{\infty} \frac{h^k}{k!} \int_{\mathbb{R}} (ix)^k e^{itx} d\mu.$$

3. Further denote $\left(\frac{d}{dt}\right)^k \phi = \phi^{(k)}$. Then

$$\phi^{(k)}(t) = \int_{\mathbb{R}} (ix)^k e^{itx} d\mu.$$

   Hint: Use (2) with $n = 1$.

4. Let ψ denote the characteristic function of ν. Show that $\phi = \psi$ and hence $\mu = \nu$.

Proof. We prove the steps in turn:

1. The idea is to use that $\alpha_{2k} = \beta_{2k}$ and that the convergence is known for α. Therefore it is sufficient to show that $\frac{\beta_{2k-1} s^{2k-1}}{(2k-1)!} \to 0$ as $k \to \infty$. Using the simple inequality $|x|^{2k-1} \leq 1 + |x|^{2k}$,

$$\frac{\beta_{2k-1} s^{2k-1}}{(2k-1)!} \leq \frac{\left(\int_{\mathbb{R}} 1 + |x|^{2k} d\mu\right) s^{2k-1}}{(2k-1)!} = \frac{s^{2k-1}}{(2k-1)!} + \frac{\beta_{2k} s^{2k-1}}{(2k-1)!}.$$

   It is standard that

$$\frac{s^{2k-1}}{(2k-1)!} \to 0$$

as $k \to \infty$ hence the first term vanishes, and as $\beta_{2k} = \alpha_{2k}$ in the second term it is sufficient to show that for large k,

$$\frac{s^{2k-1}}{(2k-1)!} \leq \frac{r^{2k}}{2k!}.$$

This is equivalent to

$$2k\left(\frac{s}{r}\right)^{2k-1} \leq r$$

which must be true for large k as the left hand side converges to zero as $k \to \infty$.

2. For any $x \in \mathbb{R}$, we use that $|e^{itx}| = 1$ and (2) for hx to see that

$$\left| e^{itx} \left(e^{ihx} - \sum_{k=0}^{n} \frac{(ihx)^k}{k!} \right) \right| \leq \frac{|hx|^{n+1}}{(n+1)!}.$$

In particular,

$$\begin{aligned} \left| \int_{\mathbb{R}} e^{itx} \left(e^{ihx} - \sum_{k=0}^{n} \frac{(ihx)^k}{k!} \right) d\mu \right| &\leq \int_{\mathbb{R}} \left| e^{itx} \left(e^{ihx} - \sum_{k=0}^{n} \frac{(ihx)^k}{k!} \right) \right| d\mu \\ &\leq \int_{\mathbb{R}} \frac{|hx|^{n+1}}{(n+1)!} d\mu \\ &= \frac{|h|^{n+1}\beta_{n+1}}{(n+1)!}. \end{aligned}$$

By the first part of the exercise, this bound vanishes as $n \to \infty$. Therefore the limit as $n \to \infty$ of the left hand side exists and is zero, which we rearrange to give that

$$\phi(t+h) = \sum_{k=0}^{\infty} \int_{\mathbb{R}} \frac{(ihx)^k}{k!} e^{itx} d\mu = \sum_{k=0}^{\infty} \frac{h^k}{k!} \int_{\mathbb{R}} (ix)^k e^{itx} d\mu$$

as required.

3. Following the hint we will look to use the inequality

$$\left| e^{ix} - 1 - ix \right| \leq \frac{x^2}{2}. \tag{3}$$

We consider the base case, wishing to prove that

$$\phi^{(1)}(t) = \int_{\mathbb{R}} ixe^{itx} d\mu.$$

We would be successful if we can verify that

$$\frac{\phi(t+s) - \phi(t)}{s} - \int_{\mathbb{R}} ixe^{itx} d\mu \tag{4}$$

tends to zero as s does. We rewrite this expression as

$$\int_{\mathbb{R}} \frac{e^{i(t+s)x} - e^{itx}}{s} - \frac{isxe^{itx}}{s} d\mu = \int_{\mathbb{R}} e^{itx} \frac{e^{isx} - 1 - isx}{s} d\mu.$$

Due to (3) we have that $\left|e^{isx} - 1 - isx\right| \leq \frac{(sx)^2}{2}$ and in particular

$$\left|e^{itx}\frac{e^{isx} - 1 - isx}{s}\right| \leq \frac{sx^2}{2}$$

so the integrand converges to 0 with s. Clearly the function $g(x) = x^2$ dominates for small s, and is integrable as second moments of μ are assumed, so (4) is shown by the Dominated Convergence Theorem. Higher derivatives are shown similarly by induction, so we conclude here.

4. Of course one can make all of the same arguments for ψ and ν in place of ϕ and μ. Combining steps two and three we have that for any $t \in \mathbb{R}$ and $|h| < r$,

$$\phi(t + h) = \sum_{k=0}^{\infty} \frac{h^k}{k!}\phi^{(k)}(t) \tag{5}$$

and similarly for ψ. We first argue that for all $|s| < r$, $\phi(s) = \psi(s)$. For this we consider $t = 0$, and observe from step three that $\phi^{(k)}(0) = \int_{\mathbb{R}}(ix)^k d\mu$ hence is completely determined by the moment α_k. As the moments for μ and ν agree then $\phi^{(k)}(0) = \psi^{(k)}(0)$. Then from (5),

$$\phi(s) = \sum_{k=0}^{\infty} \frac{s^k}{k!}\phi^{(k)}(0) = \sum_{k=0}^{\infty} \frac{s^k}{k!}\psi^{(k)}(0) = \psi(s)$$

as desired. To extend this equality beyond $(-r, r)$ now consider $t = r - \varepsilon$ for any small $\varepsilon > 0$. As ϕ and ψ are identical on $(-r, r)$ then all derivatives agree at $r - \varepsilon$ (and similarly, $-r + \varepsilon$). Once more for any $|s| < r$, from (5) we have that

$$\phi(r - \varepsilon + s) = \sum_{k=0}^{\infty} \frac{s^k}{k!}\phi^{(k)}(r - \varepsilon) = \sum_{k=0}^{\infty} \frac{s^k}{k!}\psi^{(k)}(r - \varepsilon) = \psi(r - \varepsilon + s)$$

and similarly for $-r + \varepsilon - s$, hence ϕ and ψ must agree on $(-2r, 2r)$. Inductively, $\phi = \psi$ on the whole of $\mathbb{R}$. As characteristic functions determine the measure, the result is proven.

□

Exercise 2.4.

Answer the following:

1. Is the space of continuous functions $f : \mathcal{X} \to \mathbb{R}$ measure separating?

2. Give sufficient conditions for the space of continuous and compactly supported functions $f : \mathcal{X} \to \mathbb{R}$ to be measure separating.

Proof. We answer the questions in turn:

1. Suppose that for every continuous $f : \mathcal{X} \to \mathbb{R}$ and for any probability measures μ, ν, we have that

$$\int_{\mathcal{X}} f d\mu = \int_{\mathcal{X}} f d\nu,$$

then in particular the equality is true for all bounded and uniformly continuous $f : \mathcal{X} \to \mathbb{R}$. Thus by Theorem 2.2.2 $\mu = \nu$, so the collection of continuous functions is measure separating.

2. Our conditions are that $\mathcal{X}$ is complete, separable, and a finite dimensional normed vector space. The idea is to use Theorem 2.2.5 for M the space of continuous and compactly supported functions. Thus, it is sufficient to prove that M is an algebra of bounded continuous functions separating points. Firstly M is an algebra as the pointwise product of two compactly supported continuous functions is again continuous and of support within the union of their corresponding compact supports (which is again compact). Boundedness and continuity are immediate from continuity on a compact space. It only remains to show that M separates points, so take any $x \neq y \in \mathcal{X}$. Consider the closed ball of radius $\frac{d(x,y)}{2}$ centred at x, $\bar{B}_{\frac{d(x,y)}{2}}(x)$, and define $f : \mathcal{X} \to \mathbb{R}$ by

$$f(z) = \begin{cases} \frac{d(x,y)}{2} - d(x,z) & \text{if} \quad \mathrm{z} \in \bar{B}_{\frac{d(x,y)}{2}}(x) \\ 0 & \text{otherwise.} \end{cases}$$

Clearly f is compactly supported as closed balls are compact in finite dimensional normed vector spaces, and it is also continuous as the distance function inside $\bar{B}_{\frac{d(x,y)}{2}}(x)$ is continuous and f is null on the boundary of $\bar{B}_{\frac{d(x,y)}{2}}(x)$ so the extension to zero outside of this ball is continuous. So $f \in M$ and satisfies that $f(x) = \frac{d(x,y)}{2}$, $f(y) = 0$ as required.

□

Exercise 2.5

Let $\mathcal{X}$ be a normed vector space. Consider the dual space $\mathcal{X}^*$ of all continuous linear functionals $f : \mathcal{X} \to \mathbb{R}$. Prove that $\mathcal{X}^*$ separates points.

Proof. Take any $x \neq y \in \mathcal{X}$. We construct a continuous linear functional f on $A := \text{span}(x, y)$ separating x and y by considering the linear relationship between these points across three different cases:

1. One of x or y is zero, assume $x = 0$, so define $f(ay) = a$. Then $f(x) = 0$, $f(y) = 1$.
2. There exists a $\lambda \neq 0, 1$ such that $x = \lambda y$, so define $f(ax) = a$. Then $f(x) = 1$ and $f(y) = \lambda \neq 1$.
3. x and y are linearly independent, so define $f(ax + by) = a$. Then $f(x) = 1$, $f(y) = 0$.

In all cases we have constructed a continuous linear functional f on A such that f separates points. By the Hahn-Banach Extension Theorem there exists a continuous linear functional $\tilde{f}$ on $\mathcal{X}$ such that $\tilde{f} = f$ on the subspace A. In particular $\tilde{f} \in \mathcal{X}^*$ and separates x and y, concluding the proof.

□

Exercise 2.6

Let M be a collection of functions $f : \mathcal{X} \to \mathbb{R}$. Define M^2 to be the collection of functions $\phi : \mathcal{X}^2 \to \mathbb{R}$ given by

$$M^2 := \{\phi : \phi(x, y) = f(x)g(y) \,\text{some}\, f, g \in M\}.$$

Suppose that M separates points in $\mathcal{X}$.

1. Does M^2 separate points in $\mathcal{X}^2$?

2. Suppose that M is an algebra of functions. Construct a larger algebra H, that is $f \in M$ implies that $f \in H$, such that H^2 separates points in $\mathcal{X}^2$.

Proof. We answer the questions in turn:

1. No; take $\mathcal{X} = \mathbb{R}$ and M to contain the single function $f(x) = x$. This is point separating by definition. There is only one element ϕ of M^2, given by $\phi(x, y) = f(x)f(y) = xy$. For any $x \neq 0$ consider $(x, x), (-x, -x) \in \mathbb{R}^2$. Then $\phi(x, x) = x^2 = \phi(-x, -x)$ hence M^2 is not point separating.

2. The hint is in the notation! As in Theorem 2.2.5, $H = \{f + a : f \in M,\ a \in \mathbb{R}\}$ is an algebra of functions containing M. We now show that H^2 is point separating in $\mathcal{X}^2$, so take any $(x_1, y_1) \neq (x_2, y_2) \in \mathcal{X}^2$. At least one of $x_1 \neq x_2$ or $y_1 \neq y_2$ must be true, so suppose $x_1 \neq x_2$. As M is point separating there exists an $f \in M \subset H$ such that $f(x_1) \neq f(x_2)$. As H contains all constant functions it certainly contains g such that $g(x) = 1$ for all $x \in \mathcal{X}$. Then $f(x_1)g(y_1) = f(x_1) \neq f(x_2) = f(x_2)g(y_2)$ as required.

□