



UNSW
A U S T R A L I A



UNIVERSITY OF NEW SOUTH WALES

SCHOOL OF MATHEMATICS AND STATISTICS

Assignment 2

Measure Theory

Author:
Edward McDonald

Student Number:
3375335

Question 1

For this question, let μ and ν be probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Lemma 1. *The function $x \mapsto \nu(B - x)$ is $\mathcal{B}(\mathbb{R}^d)$ measurable for any $B \in \mathcal{B}(\mathbb{R}^d)$.*

Proof. Let $B \in \mathcal{B}(\mathbb{R}^d)$ and $s(x) = \nu(B - x)$. Then,

$$\begin{aligned} s(x) &= \int_{\mathbb{R}^d} \chi_{B-x} d\nu \\ &= \int_{\mathbb{R}^d} \chi_B(y+x) d\nu(y). \end{aligned}$$

The function $(x, y) \mapsto \chi_B(y+x)$ is a composition of a continuous function, $(x, y) \mapsto y+x$ and a measurable function $x \mapsto \chi_B(x)$. Hence the function $(x, y) \mapsto \chi_B(y+x)$ is $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^d)$ measurable, and so by Tonelli's theorem s is $\mathcal{B}(\mathbb{R}^d)$ measurable. \square

Lemma 2. *The convolution measure $\mu \star \nu(B) = \int_{\mathbb{R}^d} \nu(B - x) d\mu(x)$ is well defined and finite.*

Proof. Since $\mu \star \nu(B)$ is defined as an integral of a positive measurable function, the integral exists. Since $\nu(B - x) \leq 1$, we have $\mu \star \nu(B) \leq 1$. \square

Theorem 1. *If there exists some bounded F such that $\mu \star \nu(F) = 1$, then there are bounded sets G and H such that $\mu(G) = 1$ and $\nu(H) = 1$. Similar results hold where “bounded” is replaced with “finite” or “countable”.*

Proof. Suppose that there exists a bounded set $F \in \mathcal{B}(\mathbb{R}^d)$ such that $\mu \star \nu(F) = 1$, but for any bounded set B , $\mu(B), \nu(B) < 1$. Then,

$$\begin{aligned} \mu \star \nu(F) &= \int \nu(F - x) d\mu(x) \\ &< \int 1 d\mu \\ &= 1. \end{aligned}$$

since $F - x$ is bounded for any x . This is a contradiction. Hence we must have that ν attains the value 1 on some bounded set.

By symmetry, μ must also take the value 1 on some bounded set.

An identical argument holds if “bounded” is replaced by “finite” or “countable”. \square

Question 2

For this question, μ and ν are σ -finite positive measures on a measurable space (Ω, \mathcal{F}) .

Theorem 2. *The following are equivalent:*

1. μ and ν have exactly the same null sets.
2. $\mu \ll \nu$ and $\nu \ll \mu$.
3. There is an \mathcal{F} -measurable function g with $0 < g < +\infty$ such that $\nu(A) = \int_A g \, d\mu$ for all $A \in \mathcal{F}$.

Proof. First we prove (1) \Rightarrow (2).

Assume that μ and ν have exactly the same null sets. Then if $\mu(A) = 0$, then $\nu(A) = 0$. That is $\nu \ll \mu$. Similarly, $\mu \ll \nu$.

Now we prove (2) \Rightarrow (3).

Assume that $\nu \ll \mu$.

By the Radon-Nikodym Theorem, there is an \mathcal{F} -measurable function g such that $\nu(A) = \lambda(A) + \int_A g \, d\mu$, where λ is a measure mutually singular to μ . Suppose that $\mu(A) = 0$, then $\nu(A) = 0$ by assumption, hence $\lambda(A) = 0$. Thus $\lambda \ll \mu$ and $\lambda \perp \mu$, so $\lambda = 0$.

Suppose that $g = \infty$ only on a null set. Then we can modify g to $g = 0$ on this set since g is determined only up to μ -almost everywhere equivalence.

Suppose there is some set A with $\mu(A) > 0$ and $g(A) = \{\infty\}$.

Since ν is σ -finite, there exists B with $\nu(B) < \infty$ and $\nu(A \cap B) > 0$.

Hence $g(A \cap B) = \{\infty\}$, and we cannot have $\mu(A \cap B) = 0$ because then $\nu(A \cap B) = 0$.

But then $\nu(A \cap B) = \infty$, but this is a contradiction. Hence $g < \infty$ μ -almost everywhere.

Now suppose there is some set C with $\mu(C) > 0$ and $g(C) = \{0\}$. Hence $\nu(C) = 0$, but since $\mu \ll \nu$ this is a contradiction.

Now we prove that (3) \Rightarrow (1).

Suppose that $\mu(A) = 0$. Then clearly since $\nu(A) = \int_A g \, d\mu(A)$, we have $\nu(A) = 0$.

Now suppose that $\nu(A) = 0$. Now,

$$\nu(A) \geq \frac{1}{n} \mu(A \cap g^{-1}([1/n, \infty)))$$

Hence $\mu(A \cap g^{-1}([1/n, \infty))) = 0$. But since $g > 0$, we have

$$\mu(A) = \lim_{n \rightarrow \infty} \mu(A \cap g^{-1}([1/n, \infty))).$$

Hence $\mu(A) = 0$. Thus μ and ν have the same null sets so (1) is proved. \square

Theorem 3. Suppose that μ is a σ -finite positive measure on (Ω, \mathcal{F}) . Then there is a finite positive measure ν on (Ω, \mathcal{F}) such that $\nu \ll \mu$ and $\mu \ll \nu$.

Proof. Suppose that $\{A_k\}_{k=1}^{\infty}$ is a disjoint sequence of sets with $A = \bigcup_k A_k$ and $0 < \mu(A_k) < \infty$. This can be chosen since μ is σ -finite.

Now define,

$$\nu(A) = \sum_{k=1}^{\infty} \frac{1}{2^k \mu(A_k)} \mu(A \cap A_k)$$

for $A \in \mathcal{F}$. This sum converges since each term is bounded by $1/2^k$, so the sum is bounded by a geometric series.

We wish to show that ν is a probability measure on (Ω, \mathcal{F}) and that $\mu \ll \nu$ and $\nu \ll \mu$.

Note that $\nu(\Omega) = 1$ and $\nu(\emptyset) = 0$.

Suppose that $\{B_k\}_{k=1}^{\infty}$ is a disjoint sequence of sets in \mathcal{F} . Then

$$\nu\left(\bigcup_k B_k\right) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{2^j \mu(A_j)} \mu(B_k \cap A_j)$$

Then since this is a sum of positive numbers, we can change the order of summation by Tonelli's theorem,

$$\nu\left(\bigcup_k B_k\right) = \sum_{k=1}^{\infty} \nu(B_k).$$

Hence ν is countably additive, so is a measure on (Ω, \mathcal{F}) .

Now we wish to show that $\nu \ll \mu$. Suppose that $\mu(A) = 0$. Then clearly $\mu(A_k \cap A) = 0$ for all k , so $\nu(A) = 0$. Thus $\nu \ll \mu$.

Now to show that $\mu \ll \nu$, let $\nu(A) = 0$, then $\mu(A_k \cap A) = 0$ for all k . Hence,

$$\mu(A) = \sum_{k=1}^{\infty} \mu(A_k \cap A) = 0.$$

Thus $\mu \ll \nu$. □

Question 3

For this question, X is a d -dimensional random vector with law μ and characteristic function $\hat{\mu}(u)$.

Lemma 3. *The characteristic function of cX is $\hat{\mu}(cu)$, for any $c \in \mathbb{R}$.*

Proof. This is a simple computation. By definition, the characteristic function of cX is

$$\mathbf{E}(e^{i\langle u, cX \rangle})$$

But since $\langle u, cX \rangle = \langle cu, X \rangle$, this is simply $\hat{\mu}(cu)$. □

Theorem 4. *If X has moments up to order n , then $\hat{\mu}$ is differentiable at 0 up to n th order, and $\frac{\partial^\alpha}{\partial u^\alpha} \hat{\mu}(u)|_{u=0} = i^{|\alpha|} \mathbf{E}(X^\alpha)$ for multi-indices α , with $|\alpha| \leq n$.*

Proof. Suppose that X has moments up to order n . Then we can say that,

$$\frac{\partial}{\partial u_j} \mathbf{E}(e^{i\langle u, X \rangle}) = i \mathbf{E}(X_j e^{i\langle u, X \rangle})$$

by the Dominated convergence theorem, since $\mathbf{E}(|X_j|) < \infty$. Hence, by induction,

$$\frac{\partial^\alpha}{\partial u^\alpha} \mathbf{E}(e^{i\langle u, X \rangle}) = i^{|\alpha|} \mathbf{E}(X^\alpha e^{i\langle u, X \rangle}).$$

for all multi-indices α with $|\alpha| \leq n$ by the Dominated convergence theorem. This shows that the derivative exists.

So we simply evaluate this at zero to obtain the required result. □

Now we let X be a random variable with Lebesgue density

$$f(x) = \frac{C}{(1+x^2)\log(e+x^2)}$$

for some normalising constant $C > 0$. Let $\hat{\mu}$ be the characteristic function of X .

Lemma 4. $\mathbf{E}(X)$ is not defined. That is, X does not have moments up to order 1.

Proof. We have

$$\frac{C}{(1+x^2)\log(e+x^2)} \leq \frac{C}{1+x^2}.$$

But the integral

$$\int_{[0,\infty)} \frac{x}{1+x^2} d\lambda(x)$$

where λ is Lebesgue measure is infinite, since $x/(1+x^2)$ has antiderivative $(1/2)\log(1+x^2)$, which is an unbounded function. Hence,

$$\int_{[0,\infty)} \frac{Cx}{(1+x^2)\log(e+x^2)} d\lambda(x)$$

is infinite, so $\mathbf{E}(X)$ is not defined. □

Lemma 5. $\hat{\mu}$ is differentiable at 0.

Proof. □

Question 4

For this question, μ is the binomial distribution $\text{Bin}(n, p)$, and ν is the Poisson distribution with mean $\lambda > 0$.

Lemma 6. The characteristic function of a Bernoulli random variable with probability p is

$$1 - p + pe^{iu}$$

Proof. This is a simple computation, if X takes the value 1 with probability p and 0 with probability $1 - p$, then

$$\mathbf{E}(e^{iuX}) = 1 - p + pe^{iu}.$$

□

Lemma 7. $\hat{\mu}(u) = (1 - p + pe^{iu})^n$.

Proof. A sum of n independent Bernoulli random variables with probability p is $\text{Bin}(n, p)$ distributed. So,

$$\mu = \text{Bern}(p)^{\star n}$$

Hence,

$$\hat{\mu}(u) = (1 - p + pe^{iu})^n.$$

□

Lemma 8. $\hat{\nu}(u) = \exp(\lambda(e^{iu} - 1))$.

Proof. We can compute,

$$\hat{\nu}(u) = \sum_{k=0}^{\infty} \frac{e^{iuk} e^{-\lambda} \lambda^k}{k!}$$

So we write this as,

$$\begin{aligned} \hat{\nu}(u) &= \sum_{k=0}^{\infty} e^{-\lambda} \frac{(e^{iu}\lambda)^k}{k!} \\ &= e^{-\lambda} \exp(\lambda e^{iu}) \\ &= \exp(\lambda(e^{iu} - 1)). \end{aligned}$$

□

Lemma 9. Suppose that $q_n \rightarrow \lambda \in \mathbb{C}$ is a sequence of complex numbers. Then

$$\lim_{n \rightarrow \infty} \left(1 + \frac{q_n}{n}\right)^n = e^\lambda$$

Proof. Fix n large enough such that $|q_n|/n < 1/2$.

Since q_n is a convergent sequence, it is bounded. Let M be large enough such that $|q_n| < M$ for all n .

Re-write $\left(1 + \frac{q_n}{n}\right)^n$ as $\exp(n \text{Log}(1 + \frac{q_n}{n}))$.

The branch of the logarithm taken here is complex differentiable in the set $\mathbb{C} \setminus (-\infty, 0]$. Since $|q_n|/n < 1$, the above is valid.

So it is sufficient to show that,

$$\lim_{n \rightarrow \infty} n \text{Log} \left(1 + \frac{q_n}{n}\right) = \lambda$$

The $z \mapsto \text{Log}(1+z)$ function is complex differentiable in the unit disc, and has a power series representation

$$\text{Log}(1+z) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{z^k}{k}$$

which converges uniformly on compact subsets of the open unit disc $\{z \in \mathbb{C} : |z| < 1\}$.

Now, since $|q_n|/n < 1$, we have

$$n \text{Log}\left(1 + \frac{q_n}{n}\right) = q_n + \sum_{k=2}^{\infty} (-1)^{k-1} \frac{q_n^k}{kn^{k-1}}$$

Now we consider the tail of the left hand side, let

$$L_n := \sum_{k=2}^{\infty} (-1)^{k-1} \frac{q_n^k}{kn^{k-1}}$$

By the triangle inequality,

$$|L_n| \leq \sum_{k=2}^{\infty} \frac{M^k}{kn^{k-1}}$$

Thus,

$$\begin{aligned} |L_n| &\leq M \sum_{k=1}^{\infty} \left(\frac{M}{n}\right)^k \\ &= M \frac{M/n}{(1 - M/n)} \end{aligned}$$

Hence, $L_n \rightarrow 0$ as $n \rightarrow \infty$. Thus, the limit

$$\lim_{n \rightarrow \infty} n \text{Log}\left(1 + \frac{q_n}{n}\right)$$

exists, and equals $\lim_{n \rightarrow \infty} q_n = \lambda$.

Hence, the limit

$$\lim_{n \rightarrow \infty} \left(1 + \frac{q_n}{n}\right)^n$$

exists, and equals e^λ . □

Now we let $\{p_n\}_{n=1}^\infty$ be a monotone decreasing sequence, such that $np_n \rightarrow \lambda$. We let $\mu_n = \text{Bin}(n, p_n)$.

Theorem 5. *There is weak convergence, $\mu_n \rightarrow \nu$.*

Proof. By Lévy's continuity theorem, it is sufficient to show pointwise convergence of characteristic functions, $\hat{\mu}_n(u) \rightarrow \hat{\nu}(u)$ for all u . That is, we must show

$$\lim_{n \rightarrow \infty} (1 - p_n + p_n e^{iu})^n = \exp(\lambda(e^{iu} - 1)).$$

Rewrite $\hat{\mu}_n(u)$ as

$$\left(1 + \frac{np_n(e^{iu} - 1)}{n}\right)^n$$

Now by lemma 9, we see

$$\lim_{n \rightarrow \infty} \hat{\mu}_n(u) = \exp(\lambda(e^{iu} - 1)).$$

Thus the result follows. □

Theorem 6. *For $k \in \mathbb{N}$, we have the pointwise convergence*

$$\mu_n(\{k\}) \rightarrow \nu(\{k\}).$$

Proof. Let $F_n(x) = \mu_n(-\infty, x]$ be the cumulative distribution function of μ_n , and $G(x) = \nu(-\infty, x]$

Weak convergence implies that $F_n(x) \rightarrow G(x)$ at all points of continuity x . μ_n and ν are discrete, so the points of continuity are $\mathbb{R} \setminus \mathbb{N}$.

Now let $k \in \mathbb{N}$. Then

$$\begin{aligned} \lim_{n \rightarrow \infty} \mu_n\{k\} &= \lim_{n \rightarrow \infty} F_n(k + 1/2) - F_n(k - 1/2) \\ &= G(k + 1/2) - G(k - 1/2) \\ &= \nu\{k\}. \end{aligned}$$

□

Question 5

Now we let $(\Omega, \mathcal{F}, \mathbf{P}) = ([0, 1], \mathcal{B}([0, 1]), \lambda)$ where λ is Lebesgue measure.

For $\omega \in [0, 1]$, we let $d_n(\omega)$ be the n binary digit of ω , where we take the binary expansion containing infinitely many ones. Let

$$B_n = \{\omega \in [0, 1] : d_n(\omega) = 0\}.$$

Lemma 10. $\mathbf{P}(B_n) = 1/2$.

Proof. See that,

$$B_n = \bigcup_{k=0}^{2^{n-1}-1} k/2^{n-1} + [0, 2^{-n}]$$

So $B_n \in \mathcal{B}([0, 1])$.

See that $B_n^c = \{\omega \in [0, 1] : d_n(\omega) = 1\}$.

So that $B_n^c = B_n + 2^{-n}$. Since Lebesgue measure is translation invariant, we have $\mathbf{P}(B_n^c) = \mathbf{P}(B_n)$. Since $\mathbf{P}(\Omega) = \mathbf{P}(B_n) + \mathbf{P}(B_n^c) = 1$, we have $\mathbf{P}(B_n) = 1/2$. \square

Lemma 11. *The events $\{B_n\}_{n=1}^\infty$ form an infinite sequence of independent events.*

Proof. Suppose we have some finite subset, $\{B_{n(k)}\}_{k=1}^m$. Then let

$$B = \bigcap_{k=1}^m B_{n(k)}$$

See that,

$$\bigcap_{k=2}^m B_{n(k)} = B_{n(1)} \cap \bigcap_{k=2}^m B_{n(k)} \cup B_{n(1)}^c \cap \bigcap_{k=2}^m B_{n(k)}.$$

But $B_{n(1)}^c = B_{n(1)} + 2^{-n(1)}$.

Hence,

$$\mathbf{P}\left(\bigcap_{k=2}^m B_{n(k)}\right) = 2\mathbf{P}\left(\bigcap_{k=1}^m B_{n(k)}\right).$$

So by induction,

$$\mathbf{P}\left(\bigcap_{k=2}^m B_{n(k)}\right) = 2^{-m} = \prod_{k=1}^m \mathbf{P}(B_{n(k)}).$$

Hence the sequence $\{B_n\}_{n=1}^\infty$ is independent. \square

Remark 1. *Precisely the same arguments would work if we replaced binary digits with decimal digits.*

Theorem 7. *Given any finite sequence of digits, the probability that a randomly sampled number in $[0, 1]$ contained that sequence infinitely many times is 1.*

Proof. Suppose that the sequence has length L . Let E_n be the event that the sequence occurs in the nL th position. Then the E_n form an independent sequence of independent events, each with probability 10^{-L} . Then since

$$\sum_{n=1}^{\infty} \mathbf{P}(E_n) = \infty.$$

Hence, by the Borel-Cantelli lemma, the probability that infinitely many of the E_n occur is 1. \square

Question 6

Lemma 12. *Let s be a continuous solution to the functional equation*

$$4s(2x) = 3s(x) + s(-x).$$

for x in a neighbourhood of 0. Then s is constant.

Proof. For any constant c , if s is a solution to the functional equation then so is $s + c$. So we may assume without loss of generality that $s(0) = 0$.

Suppose $s(x) > \varepsilon > 0$. Then

$$3s(x/2) + s(-x/2) > 4\varepsilon.$$

Hence since the average of $(s(x/2), s(x/2), s(x/2), s(-x/2))$ is larger than ε , at least one of these numbers must exceed ε . Hence at least one of $s(x/2), s(-x/2)$ exceeds ε .

Thus we have a sequence of numbers approaching 0 which all exceeds ε . But this contradicts $s(0) = 0$ and continuity. \square

Lemma 13. *Suppose that f is a solution of the functional equation*

$$f(2x) = f(x)^3 f(-x)$$

for x in a neighbourhood of 0, with $f(0) = 1$, and f is assumed to be twice continuously differentiable in a neighbourhood of the origin.

Then $f(x) = \exp(Ax^2 + Bx)$ for parameters A and B .

Proof. Since $f(0) = 1$, and f is continuous, then we may restrict x sufficiently small such that $f(x) > 0$. Hence $L(x) := \log(f(x))$ is well defined, and

$$L(2x) = 3L(x) + L(-x).$$

Differentiating twice, we have

$$4L''(2x) = 3L''(x) + L''(-x)$$

But by lemma 12, we have $L''(x)$ is constant.

Thus, L is a quadratic, so $f(x) = \exp(Ax^2 + Bx + C)$ for constants A, B and C . We see $C = 0$ since $f(0) = 1$. \square

Theorem 8. *Suppose that X and Y are independent identically distributed random variables, with finite variances.*

Also assume that $X + Y$ and $X - Y$ are independent.

Then X and Y are Gaussian.

Proof. Let $\varphi(u)$ be the characteristic function of X and Y . Then the characteristic function of $X + Y$ is $\varphi(u)^2$, and the characteristic function of $X - Y$ is $\varphi(u)\varphi(-u)$. The characteristic function of $2X$ is $\varphi(2u)$. Since $2X = X + Y + X - Y$, and $X + Y$ and $X - Y$ are independent, we have

$$\varphi(2u) = \varphi(u)^3\varphi(-u)$$

and $\varphi(0) = 1$. Since X and Y have finite variances, φ is twice continuously differentiable in a neighbourhood of 0. Thus, by lemma 13 we have $\varphi(u) = \exp(Au^2 + Bu)$ for some parameters A and B .

This is the characteristic function of a Gaussian random variable. \square