The University of Sydney
School of Mathematics and Statistics

Solutions to Tutorial 12 (Week 13)

MATH3969: Measure Theory and Fourier Analysis (Advanced) Semester 2, 2011

Web Page: http://www.maths.usyd.edu.au/u/UG/SM/MATH3969/
Lecturer: Daniel Daners

Material covered

(1) applications of measure theory to probability theory

(2) conditional expectation

Outcomes

After completing this tutorial you should

(1) have an idea on how measure theory is applied to probability theory.

(2) be familiar with conditional expectation.

Questions to complete during the tutorial

1. Let $\Omega = [0,1]$ and $P = m$ the Lebesgue measure. Then $[0,1]$ is a probability space. Give examples of two distinct random variables which have the same distribution.

Solution: Clearly $X(\omega) := \omega$ and $Y(\omega) := 1-\omega$ have the same distribution.

2. Let $X : \Omega \to \mathbb{R}$ be a random variable on the probability space $(\Omega, \mathcal{A}, P)$.

(a) Prove that

$$P[|X| \geq \alpha] \leq \frac{1}{\alpha^p}\int_{[|X|\geq\alpha]} |X|^p\,dP \leq \frac{1}{\alpha^p}\mathbb{E}[|X|^p]$$

for all $\alpha > 0$ and $1 \leq p < \infty$.

Solution: Note that for $\omega \in [|X| \geq \alpha] = \{\omega \in \Omega : |X(\omega)| \geq \alpha\}$ we have $1 \leq |X(\omega)|^p/\alpha^p$. Hence

$$\begin{aligned} P[|X| \geq \alpha] = \int_{[|X|\geq\alpha]} 1\,dP &\leq \int_{[|X|\geq\alpha]} \frac{|X|^p}{\alpha^p}\,dP \\ &= \frac{1}{\alpha^p}\int_{[|X|\geq\alpha]} |X|^p\,dP \leq \frac{1}{\alpha^p}\int_{\Omega} |X|^p\,dP = \frac{1}{\alpha^p}\mathbb{E}[|X|^p]. \end{aligned}$$

(b) Prove *Chebychev's inequality*

$$P[|X-\mu| \geq \alpha] \leq \frac{1}{\alpha^2}\operatorname{Var}(X)$$

for all $\alpha > 0$, where $\mu := \mathbb{E}[X]$ is the expectation of X.

Solution: By the previous part applied to $X - \mu$ and $p = 2$ we get

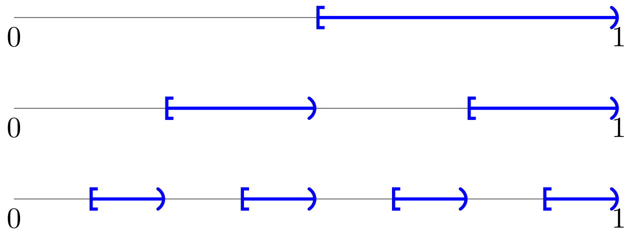
$$P[|X-\mu| \geq \alpha] \leq \frac{1}{\alpha^2}\mathbb{E}[|X-\mu|^2] = \frac{1}{\alpha^2}\operatorname{Var}(X)$$

as claimed.

3. Let Ω be the space obtained by coin tossing countably many times. Such coin tosses can be represented as infinite sequences of the form $\Omega := \{(a_1, a_2, \dots) \colon a_k = 0 \text{ or } 1\}$. A zero means head and a one means tail for instance. Such sequences can be interpreted as binary expansions of the number $\alpha = \sum_{k=1}^{\infty} \frac{a_k}{2^k}$ which is between zero and one. With that identification we can set $\Omega = [0, 1)$.

(a) Denote by $A_k := \{(a_1, \dots, a_k, \dots) \in \Omega \colon a_k = 1\}$. Using that $\Omega = [0, 1)$ sketch A_1, A_2, A_3 and then describe the sets A_k for general k. What is the probability of A_k, and how does it compare to the Lebesgue measure of A_k?

Solution: If $a_1 = 1$, then $\alpha = \frac{1}{2} + \sum_{j=1}^{\infty} \frac{a_j}{2^j} \geq \frac{1}{2}$, so $A_1 = [\frac{1}{2}, 1)$. If $k = 2$, then $\alpha = \frac{a_1}{2} + \frac{1}{4} \sum_{j=1}^{\infty} \frac{a_j}{2^j}$, so we have $A_2 = [\frac{1}{4}, \frac{1}{2}) \cup [\frac{3}{4}, 1)$ and similarly $A_3 = [\frac{1}{8}, \frac{1}{4}) \cup [\frac{3}{8} \cup \frac{1}{2}) \cup [\frac{5}{8}, \frac{3}{4}) \cup [\frac{7}{8}, 1)$. The Lebesgue measure of each of the sets is clearly $1/2$. A sketch of A_1, A_2 and A_3 is as follows:



(b) Show that every interval of the form $I_{n,j} = [j/2^n, (j+1)/2^n)$ $(j = 0, \dots, 2^{n-1})$ can be written as a finite intersection of the sets A_k and their complements.

Solution: First note that $I_{n,j} \subseteq A_k$ or $I_{n,j} \subseteq A_k^c$ for all $k = 1, \dots, n$. Hence, for $k = 1, \dots, n$ we let $B_{j,k} := A_k$ if $I_{n,j} \subseteq A_k$ and $B_{j,k} := A_k^c$ if $I_{n,j} \subseteq A_k^c$. Then $I_{n,j} = \bigcap_{k=1}^{n} B_{j,k}$.

(c) Argue why the probability measure in the above situation is the Lebesgue measure on $[0, 1)$.

Solution: Every interval $(a, b) \subseteq [0, 1)$ can be written as a disjoint union of countably many intervals of the form $[j/2^n, (j+1)/2^n)$ (dyadic decomposition of the interval) and therefore the measure is equal to $b - a$. This induces Lebesgue measure.

Extra questions for further practice

4. Let $(\Omega, \mathcal{A}, P)$ be a probability space and $\mathcal{A}_0$ a σ-algebra with $\mathcal{A}_0 \subseteq \mathcal{A}$. Then clearly $L^2(\Omega, \mathcal{A}_0, P)$ is a closed subspace of $L^2(\Omega, \mathcal{A}, P)$ and therefore, by the projection theorem in a Hilbert space, for every random variable $X \in L^2(\Omega, \mathcal{A}, P)$ there exists $X_0 \in L^2(\Omega, \mathcal{A}_0, P)$ such that $X - X_0$ is orthogonal to $L^2(\Omega, \mathcal{A}_0, P)$. Prove that $X_0 = \mathbb{E}[X|\mathcal{A}_0]$ almost everywhere.

Solution: By the properties of conditional expectation proved in lectures

$$\int_\Omega (X - X_0) Y \, dP = \int_\Omega XY \, dP - \int_\Omega X_0 Y \, dP = \int_\Omega XY \, dP - \int_\Omega XY \, dP = 0$$

For all $Y \in L^\infty(\Omega, \mathcal{A}_0, P)$. By density of $L^\infty(\Omega, \mathcal{A}_0, P)$ in $L^2(\Omega, \mathcal{A}_0, P)$ it follows that $X - X_0$ is orthogonal to $L^2(\Omega, \mathcal{A}_0, P)$ as claimed.

The following question generalises *Jensen's inequality* to conditional expectation.

5. Let $(\Omega, \mathcal{A}, P)$ be a probability space and $\mathcal{A}_0$ a σ-algebra with $\mathcal{A}_0 \subseteq \mathcal{A}$. Let $X \in L^1(\Omega, \mathcal{A}, P)$ be a random variable and $X_0 := \mathbb{E}[X|\mathcal{A}_0]$. Finally let $\varphi \colon \mathbb{R} \to \mathbb{R}$ be a convex function.

(a) Show that there exists an increasing function $m \colon \mathbb{R} \to \mathbb{R}$ such that

$$m(t) = \sup_{s<t} \frac{\varphi(s) - \varphi(t)}{s - t}$$

for all $s, t \in \mathbb{R}$.

Solution: If $\varphi\colon \mathbb{R}\to\mathbb{R}$ is convex, then according to lectures we can write

$$\varphi(s) \geq \varphi(t) - m(t)(s-t)$$

for all $s,t\in\mathbb{R}$ if we set

$$m(t) = \sup_{s<t} \frac{\varphi(s)-\varphi(t)}{s-t}.$$

Then $m\colon \mathbb{R}\to\mathbb{R}$ is an increasing function and therefore Borel measurable. Since φ is convex it is also continuous and therefore Borel measurable as well.

(b) Let $X \in L^\infty(\Omega,\mathcal{A},P)$. Prove that $|X_0| \leq |X|$ almost everywhere.

Solution: $A := \{\omega\in\Omega\colon X_0(\omega) > \|X\|_\infty\}$. By definition of conditional expectation $A\in\mathcal{A}_0$ and

$$\|X\|_\infty P(A) \leq \int_A X_0\,dP = \int_A X\,dP = 0.$$

Hence $X_0 \leq X$ almost everywhere. Similarly we show that $X_0 \geq X$ almost everywhere, so that $\|X_0\|_\infty \leq \|X\|_\infty$.

(c) Let $X \in L^\infty(\Omega,\mathcal{A},P)$. Prove that

$$\mathbb{E}[\varphi\circ X|\mathcal{A}_0] \geq \varphi\circ X_0 = \varphi\circ\mathbb{E}[X|\mathcal{A}_0]$$

almost everywhere.

Hint: Use (a) to show that $\varphi\circ X \geq \varphi\circ X_0 - (m\circ X_0)(X-X_0)$, and then integrate. Check the measurability of the functions carefully.

Solution: The inequality in (a) implies that

$$\varphi\circ X \geq \varphi\circ X_0 - (m\circ X_0)(X-X_0).$$

From the construction of conditional expectation we know that $\mathbb{E}[Y|\mathcal{A}_0]\geq 0$ whenever $Y\geq 0$. We also know that taking conditional expectation is a linear map. Hence

$$\mathbb{E}[\varphi\circ X|\mathcal{A}_0] \geq \mathbb{E}[\varphi\circ X_0|\mathcal{A}_0] - \mathbb{E}[(m\circ X_0)(X-X_0)|\mathcal{A}_0].$$

Since $\varphi\circ X_0$ is $\mathcal{A}_0$-measurable we have $\mathbb{E}[\varphi\circ X_0|\mathcal{A}_0] = \varphi\circ X_0$. Moreover, since $X_0\in L^\infty(\Omega,\mathcal{A}_0,P)$ by (b) and $m(\cdot)$ is an increasing function $m\circ X_0 \in L^\infty(\Omega,\mathcal{A}_0,P)$. By the properties of conditional expectation from lectures

$$\begin{aligned}\mathbb{E}[(m\circ X_0)(X-X_0)|\mathcal{A}_0] &= (m\circ X_0)\mathbb{E}[X-X_0|\mathcal{A}_0]\\ &= (m\circ X_0)\big(\mathbb{E}[X|\mathcal{A}_0]-\mathbb{E}[X_0|\mathcal{A}_0]\big) = (m\circ X_0)\big(X_0-X_0\big) = 0\end{aligned}$$

Putting everything together we get

$$\mathbb{E}[\varphi\circ X|\mathcal{A}_0] \geq \varphi\circ X_0 = \varphi\circ\mathbb{E}[X|\mathcal{A}_0]$$

for all $X\in L^\infty(\Omega,\mathcal{A},P)$ as claimed.

(d) If $\varphi\circ X\in L^1(\Omega,\mathcal{A},P)$, prove that $\varphi\circ\mathbb{E}[X|\mathcal{A}_0] \leq \mathbb{E}[\varphi\circ X|\mathcal{A}_0]$ almost everywhere. (This generalises Jensen's inequality.)

Hint: Approximate X by a sequence of bounded functions.

Solution: Let

$$A_n := \{\omega\in\Omega\colon |X_0(\omega)| < n\}.$$

From what we proved above we have

$$\mathbb{E}[\varphi\circ(1_{A_n}X)|\mathcal{A}_0] \geq \varphi\circ\mathbb{E}[1_{A_n}X|\mathcal{A}_0] \tag{1}$$

Since $A_n \in \mathcal{A}_0$ we have $\mathbb{E}[1_{A_n}X|\mathcal{A}_0] = 1_{A_n}\mathbb{E}[X|\mathcal{A}_0]$ and by the continuity of φ we have

$$\varphi \circ \mathbb{E}[1_{A_n}X|\mathcal{A}_0] = \varphi \circ (1_{A_n}\mathbb{E}[X|\mathcal{A}_0]) \to \varphi \circ \mathbb{E}[X|\mathcal{A}_0]$$

pointwise. Now clearly $\varphi(1_{A_n}X(\omega)) = \varphi(X(\omega))$ if $\omega \in A_n$ and $\varphi(1_{A_n}X(\omega)) = \varphi(0)$ if $\omega \in A_n^c$ and so

$$\varphi \circ (1_{A_n}X) = 1_{A_n}\varphi \circ X + 1_{A_n^c}\varphi(0).$$

Since $1_{A_n}, 1_{A_n^c} \in L^\infty(\Omega, \mathcal{A}_0, P)$ we have

$$\mathbb{E}[\varphi \circ (1_{A_n}X)|\mathcal{A}_0] = \mathbb{E}[1_{A_n}\varphi \circ X|\mathcal{A}_0] + \mathbb{E}[1_{A_n^c}\varphi(0)|\mathcal{A}_0] = 1_{A_n}\mathbb{E}[\varphi \circ X|\mathcal{A}_0] + 1_{A_n^c}\varphi(0)$$

for all $n \in \mathbb{N}$. Note that $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$ and set $A := \bigcup_{n\in\mathbb{N}}$. Hence $1_{A_n} \to 1_A$ and $1_{A_n^c} \to 1_{A^c}$ pointwise as $n \to \infty$. Since $X_0 \in L^1(\Omega, \mathcal{A}, P)$ we have $P(A) = 1$ and therefore

$$\mathbb{E}[\varphi \circ (1_{A_n}X)|\mathcal{A}_0] = 1_{A_n}\mathbb{E}[\varphi \circ X|\mathcal{A}_0] + 1_{A_n^c}\varphi(0) \to \mathbb{E}[\varphi \circ X|\mathcal{A}_0]$$

almost everywhere (with probability one). Hence we get the required inequality almost everywhere by passing to the limit in (1).

6. Let $(\Omega, \mathcal{A}, P)$ be a probability space and $\mathcal{A}_0$ a σ-algebra with $\mathcal{A}_0 \subseteq \mathcal{A}$. Let $X \in L^1(\Omega, \mathcal{A}, P)$ be a random variable. Use Question **5** to show that the linear map $X \mapsto \mathbb{E}[X|\mathcal{A}_0]$ is continuous from $L^p(\Omega, \mathcal{A}, P)$ to $L^p(\Omega, \mathcal{A}_0, P)$ if $1 \leq p < \infty$.

Solution: Since $t \to |t|^p$ is convex for $1 \leq p < \infty$ we conclude from (5) that

$$|E[X|\mathcal{A}_0]|^p \leq E[|X|^p|\mathcal{A}_0]$$

almost everywhere. Hence by definition of conditional expectation

$$\|E[X|\mathcal{A}_0]\|_p = \Big(\int_\Omega |E[X|\mathcal{A}_0]|^p \, dP\Big)^{1/p} \leq \Big(\int_\Omega E[|X|^p\mathcal{A}_0] \, dP\Big)^{1/p} = \Big(\int_\Omega |X|^p \, dP\Big)^{1/p} = \|X\|_p.$$

By the linearity of the map $X \mapsto \mathbb{E}[X|\mathcal{A}_0]$ continuity follows.

The question below is an application of the theory of Fourier transforms to probability.

7. Let $(\Omega, \mathcal{A}, P)$ be a probability space and $X\colon \Omega \to \mathbb{R}$ a random variable. Let P_X be the distribution of X, that is, $P_X[A] := P[X \in A]$ for every Borel set $A \subseteq \mathbb{R}$. Define

$$\varphi_X(t) := \int_\mathbb{R} e^{ist} \, dP_X(s)$$

Then $\varphi_X\colon \mathbb{R} \to \mathbb{C}$ is called the *characteristic function* of the random variable X. Note that up to some normalising factors this is the inverse Fourier transform of the measure P_X.

(a) Show that φ_X is continuous and that $\|\varphi_X\|_\infty \leq 1$.

Solution: The function $t \to e^{i\omega t}$ is continuous and $|e^{i\omega t}| \leq 1$. Since 1 is integrable with respect to the probability measure P_X the theorem on the continuity of integrals with parameters applies and φ_X is continuous. Moreover,

$$|\varphi_X(t)| \leq \int_\mathbb{R} |e^{ist}| \, dP_X(s) = \int_\mathbb{R} 1 \, dP_X(s) = P_X[\mathbb{R}] = 1$$

as claimed.

(b) Show that $\varphi_X(t) = \mathbb{E}[e^{itX}]$

Solution: Let $f(x) := e^{i\omega x}$. We proved in lectures that

$$\int_{\mathbb{R}} e^{ist}\, dP_X(s) = \int_{\mathbb{R}} f\, dP_X = \int_{\Omega} f \circ X\, dP = \int_{\Omega} e^{isX}\, dP(s) = \mathbb{E}[e^{itX}].$$

(c) Suppose that P_X has a density function. Show that $\varphi_X \in C_0(\mathbb{R})$.

Solution: Let g be the density of P_X. Then

$$\varphi_X(t) = \int_{\mathbb{R}} e^{ist} g(s)\, ds.$$

This is essentially the (inverse) Fourier transform of the density function g. As $g \in L^1(\mathbb{R})$ we can use the Riemann-Lebesgue lemma to conclude that $\varphi_X \in C_0(\mathbb{R})$.

(d) Suppose that $P_X = \delta_0$, where δ_0 is the Dirac measure concentrated at $t = 0$. Compute the characteristic function φ_X.

Solution: Using integration with respect to the Dirac measure we get

$$\varphi_X(t) = \int_{\mathbb{R}} e^{ist}\, d\delta_0 = e^{i0t} = 1$$

for all $t \in \mathbb{R}$. Hence the characteristic function is constant.

(e) Compute the characteristic function of a normally distributed random variable, that is, a random variable with distribution density $(2\pi\sigma)^{-1/2} e^{-|x-\mu|^2/2\sigma}$ (mean $\mu \in \mathbb{R}$ and standard deviation $\sigma > 0$).

Solution: By definition of the characteristic function

$$\varphi_X(t) = \frac{1}{\sqrt{2\pi\sigma}} \int_{\mathbb{R}} e^{ist} e^{-|s-\mu|^2/2\sigma}\, ds$$

We know that $\hat{\varphi} = \varphi$ if $\varphi(x) := e^{-\pi x^2}$. Hence we make a substitution so that

$$\pi x^2 = \frac{|s-\mu|^2}{2\sigma},$$

that is,

$$s = \mu + \sqrt{2\pi\sigma}\, x, \qquad ds = \sqrt{2\pi\sigma}\, dx.$$

Looking at the exponent we write

$$ist = i(\mu + \sqrt{2\pi\sigma}\, x)t = i\mu - 2\pi x \sqrt{\frac{\sigma}{2\pi}} t$$

Hence

$$\begin{aligned}\varphi_X(t) = \frac{1}{\sqrt{2\pi\sigma}} \int_{\mathbb{R}} e^{ist} e^{-|s-\mu|^2/2\sigma}\, ds = e^{\mu t} \int_{\mathbb{R}} \varphi(x) e^{-2\pi x \sqrt{\frac{\sigma}{2\pi}} t}\, dx \\ = e^{i\mu} \hat{\varphi}\Big(\sqrt{\frac{\sigma}{2\pi}} t\Big) = e^{i\mu} \varphi\Big(\sqrt{\frac{\sigma}{2\pi}} t\Big) = e^{i\mu} e^{-\frac{\sigma}{2} t^2}.\end{aligned}$$

In particular, if X is normally distributed with mean zero and standard deviation one we get

$$\varphi_X(t) = e^{-t^2/2}.$$