Stat 310B/Math 230B Theory of Probability

Homework 1 Solutions

Andrea Montanari

Exercise [4.1.3]

We need to show that $\mathbf{E}(XI_A) = \mathbf{E}(YI_A)$, for all $A \in \mathcal{G} = \sigma(\mathcal{P})$. Let $\mathcal{L} = \{A \in \mathcal{F} : \mathbf{E}(XI_A) = \mathbf{E}(YI_A)\}$. The assumption implies $\mathcal{P} \subseteq \mathcal{L}$. By Dynkin's $\pi - \lambda$ theorem, it suffices to show that $\mathcal{L}$ is λ-system, which we proceed to check.

First, $\Omega \in \mathcal{L}$ since $\Omega \in \mathcal{P} \subseteq \mathcal{L}$. Second, if $A \in \mathcal{L}, B \in \mathcal{L}$ and $A \subseteq B$, then taking the difference of the two integral identities we see that $B \backslash A \in \mathcal{L}$. Finally, if $A_i \in \mathcal{L}, A_i \uparrow A$ then applying dominated convergence we conclude that $A \in \mathcal{L}$.

Exercise [4.1.8]

1. It suffices to prove the claim for $n = 2$ as the general case then follows by $n - 1$ iterations. Fixing hereafter $n = 2$, $f(x_1, x_2) = f_1(x_1)f_2(x_2)$ is a finite valued element of $m\mathcal{F}_+$ (see Corollary 1.2.19), hence $\widehat{\nu} = f\mu$ is a measure on $\mathcal{F}$ (see Proposition 1.3.56), such that $\widehat{\nu} \ll \mu$. Since $\nu_k = f_k\mu_k$, if $A = A_1 \times A_2$ and $A_k \in \mathcal{F}_k$, then by Fubini's theorem and the definition of product measures,
$$\nu(A) = \nu_1(A_1)\nu_2(A_2) = \mu_1(f_1 I_{A_1})\mu_2(f_2 I_{A_2}) = \mu(f I_A) = \widehat{\nu}(A)\,.$$
In conclusion, the measures ν and $\widehat{\nu}$ coincide on the π-system $\mathcal{R} = \{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$. Further, since ν_k are σ-finite, there exist $R_\ell \in \mathcal{R}$ such that $\nu(R_\ell) < \infty$ and $R_\ell \uparrow \mathbb{S}$. In view of the remark following Proposition 1.1.39, $\nu = \widehat{\nu}$ throughout $\mathcal{F} = \sigma(\mathcal{R})$, as claimed.

2. Consider the random variables $X_k = f_k(s_k)$ in probability space $(\mathbb{S}, \mathcal{F}, \mu)$. By definition $\sigma(X_k)$ consists of all sets of the form $A_k = \{(s_1, \ldots, s_n) : f_k(s_k) \in B_k\}$ for some $B_k \in \mathcal{B}$. Therefore, $\cap_{k=1}^n A_k = A_1 \times \cdots \times A_n$ and by the construction of product measure μ we have that
$$\mu(\bigcap_{i=1}^n A_i) = \mu(A_1 \times \cdots \times A_n) = \prod_{i=1}^n \mu(A_i)$$
(see Eq. (1.4.3)). This is precisely the definition of mutual independence of X_k and the same argument applies in the probability space $(\mathbb{S}, \mathcal{F}, \nu)$.

Exercise [4.2.16]

1. Expanding both sides of the inequality, you see that it amounts to showing that
$$\mathbf{E}[\mathbf{E}(X|\mathcal{G}_1)^2] \le \mathbf{E}[\mathbf{E}(X|\mathcal{G}_2)^2]$$
Let $Y = \mathbf{E}(X|\mathcal{G}_1)$ and $Z = \mathbf{E}(X|\mathcal{G}_2)$. From Example 4.2.20 of the notes we know that both Y and Z are square integrable. Further, by the tower property $\mathbf{E}(Z|\mathcal{G}_1) = Y$ so after also taking out what is known
$$\mathbf{E}(YZ) = \mathbf{E}(\mathbf{E}(YZ|\mathcal{G}_1)) = \mathbf{E}(Y\mathbf{E}(Z|\mathcal{G}_1)) = \mathbf{E}(Y^2)\,.$$
Consequently,
$$0 \le \mathbf{E}[(Z - Y)^2] = \mathbf{E}(Z^2) - 2\mathbf{E}(ZY) + \mathbf{E}(Y^2) = \mathbf{E}(Z^2) - \mathbf{E}(Y^2)$$
which is precisely what you are asked to prove. Note that this result is also a direct consequence of Proposition 4.3.1.

2. Let $Y = \mathbf{E}(X|\mathcal{G})$. You get the stated identity by adding the equations

$$\mathrm{Var}(Y) \;=\; \mathbf{E}Y^2 - (\mathbf{E}Y)^2 \;=\; \mathbf{E}Y^2 - (\mathbf{E}X)^2$$

and

$$\mathbf{E}(\mathrm{Var}(X|\mathcal{G})) = \mathbf{E}(X - Y)^2 = \mathbf{E}X^2 - 2\mathbf{E}XY + \mathbf{E}Y^2 = \mathbf{E}X^2 - \mathbf{E}Y^2$$

Exercise [4.2.22]

1. Fix $p > 0$ and $A \in \mathcal{G}$. Then, setting $Y = |X|I_A$ and $U_x = \mathbf{P}(|X| > x|\mathcal{G})$ we have from part (a) of Lemma 1.4.31 and the definition of U_x that

$$\begin{aligned}\mathbf{E}[|X|^p I_A] &= \mathbf{E}Y^p = \int_0^\infty p y^{p-1}\mathbf{P}(Y > y)dy \\ &= \int_0^\infty p x^{p-1}\mathbf{E}[I_{\{|X|>x\}} I_A]dx = \int_0^\infty p x^{p-1}\mathbf{E}[U_x I_A]dx\,.\end{aligned}$$

Recall that $U_x I_A \in m\mathcal{G}$ for each $x \geq 0$. Without loss of generality we further assume that the non-negative function $h(x, \omega) = p x^{p-1} U_x I_A$ is measurable on the product space $\mathcal{B} \times \mathcal{G}$ (the easiest way to see this is by taking the version of U_x given by the measure of the open interval (x, ∞) under the R.C.P.D. of $|X|$ given $\mathcal{G}$, which exists by Proposition 4.4.3).

Thus, by Fubini's theorem $Z = \int_0^\infty p x^{p-1} U_x dx$ is measurable on $\mathcal{G}$ and

$$\mathbf{E}[|X|^p I_A] = \int_0^\infty \mathbf{E}[h(x, \omega)]dx = \mathbf{E}[\int_0^\infty h(x, \omega)dx] = \mathbf{E}[Z I_A]\,.$$

In particular, $\mathbf{E}Z = \mathbf{E}|X|^p$ is finite and as the preceding applies for all $A \in \mathcal{G}$, it follows by definition of conditional expectation that $Z = \mathbf{E}[|X|^p|\mathcal{G}]$.

2. Fixing $a > 0$ let $V_a = \mathbf{P}(|X| \geq a|\mathcal{G})$. By the monotonicity of the conditional expectation $U_x \geq 0$ for any $x \geq 0$ and further $U_x \geq V_a$ whenever $x \in [0, a)$. Hence, for any $a > 0$ and $A \in \mathcal{G}$ we have from our proof of part (a) that

$$\mathbf{E}[|X|^p I_A] = \int_0^\infty p x^{p-1}\mathbf{E}[U_x I_A]dx \geq \int_0^a p x^{p-1} dx \mathbf{E}[V_a I_A] = a^p \mathbf{E}[V_a I_A]\,.$$

To conclude that almost surely $V_a \leq a^{-p}\mathbf{E}[|X|^p\ |\mathcal{G}]$ consider the above inequality for $A_n = \{\omega : a^p V_a \geq n^{-1} + \mathbf{E}[|X|^p\ |\mathcal{G}]\ \}$, then take $n \to \infty$.

Exercise [4.2.21]

1. Taking out what is known,

$$\mathbf{E}(XZ) = \mathbf{E}(\mathbf{E}(XZ|\mathcal{G})) = \mathbf{E}(Z\mathbf{E}(X|\mathcal{G})) = \mathbf{E}Z^2.$$

Therefore,

$$\mathbf{E}(X - Z)^2 \;=\; \mathbf{E}X^2 - 2\mathbf{E}XZ + \mathbf{E}Z^2 \;=\; \mathbf{E}X^2 - \mathbf{E}Z^2 = 0\,,$$

from which we deduce that $Z = X$ a.s.

2. We know from (cJENSEN) that almost surely $|Z| = |\mathbf{E}(X|\mathcal{G})| \leq \mathbf{E}(|X||\mathcal{G})$. Hence, if $\mathbf{P}(|Z| < \mathbf{E}(|X||\mathcal{G})) > 0$ then $\mathbf{E}(|\mathbf{E}(X|\mathcal{G})|) < \mathbf{E}(\mathbf{E}(|X||\mathcal{G})) = \mathbf{E}(|X|)$, in contradiction with our hypothesis that $|Z| = |\mathbf{E}(X|\mathcal{G})|$ has the same law as $|X|$ (hence the same expected value). We thus conclude that

$|Z| = \mathbf{E}(|X||\mathcal{G})$ almost surely. Note that $A = \{Z \geq 0\}$ is by definition of Z in $\mathcal{G}$ and further by the preceding,

$$\mathbf{E}[XI_A] = \mathbf{E}[ZI_A] = \mathbf{E}[|Z|I_A] = \mathbf{E}[\mathbf{E}(|X|\,|\mathcal{G})I_A] = \mathbf{E}[|X|I_A]\,.$$

That is, $\mathbf{E}[(|X| - X)I_A] = 0$, namely, $X \geq 0$ for almost every $\omega \in A$. Our hypothesis that $\mathbf{E}[X|\mathcal{G}]$ has the same law as X implies that the same hypothesis holds for $Y = X - c$ and any non-random constant c. Therefore, by the preceding we get that

$$\mathbf{P}(\{X < c \leq E(X|\mathcal{G})\}) = \mathbf{P}(\{Y < 0 \leq E(Y|\mathcal{G})\}) = 0.$$

Since $\{X < \mathbf{E}(X|\mathcal{G})\} = \bigcup_{c\in\mathbf{Q}}\{X < c \leq \mathbf{E}(X|\mathcal{G})\}$, it follows that $X \geq \mathbf{E}(X|\mathcal{G})$ a.s. To complete the proof re-run the above argument for $-X$ instead of X.

Exercise [4.2.23]

For $\varepsilon \geq 0$, let $U_\varepsilon = \big[\mathbf{E}(|X|^p|\mathcal{G}) + \varepsilon\big]^{1/p}$ in $L^p(\Omega, \mathcal{G}, \mathbf{P})$ and $V_\varepsilon = \big[\mathbf{E}(|Y|^q|\mathcal{G}) + \varepsilon\big]^{1/q}$, in $L^q(\Omega, \mathcal{G}, \mathbf{P})$, that per $\varepsilon > 0$ are both uniformly bounded below away from zero. Recall that

$$\frac{x^p}{p} + \frac{y^q}{q} - xy \geq 0, \qquad \text{for all } x, y \geq 0$$

(which you verify by considering the first two derivatives in x of the function on the left side). Hence, for each ω and $\varepsilon > 0$,

$$\Big|\frac{X(\omega)Y(\omega)}{U_\varepsilon(\omega)V_\varepsilon(\omega)}\Big| \leq \frac{1}{p}\Big|\frac{X(\omega)}{U_\varepsilon(\omega)}\Big|^p + \frac{1}{q}\Big|\frac{Y(\omega)}{V_\varepsilon(\omega)}\Big|^q\,.$$

With $1/U_\varepsilon$ and $1/V_\varepsilon$ uniformly bounded, the expectation of both sides conditional upon $\mathcal{G}$ is well defined, and it follows from monotonicity of the C.E. (i.e. Corollary 4.2.6), upon taking out what is known that for any $\varepsilon > 0$ and a.e. ω,

$$\frac{\mathbf{E}(|XY|\,|\mathcal{G})}{U_\varepsilon V_\varepsilon} \leq \frac{1}{p}\frac{U_0^p}{U_\varepsilon^p} + \frac{1}{q}\frac{V_0^q}{V_\varepsilon^q} \leq \frac{1}{p} + \frac{1}{q} = 1\,.$$

Multiplying both sides by $U_\varepsilon V_\varepsilon$ and considering $\varepsilon_k \downarrow 0$ yields the stated claim that $\mathbf{E}(|XY|\,|\mathcal{G}) \leq U_0 V_0$.

Exercise [4.2.27]

(a) implies (b): (b) holds for indicator functions h_1 and h_2 by (a). By linearity of $\mathbf{E}(\cdot|\mathcal{G})$ and (cDOM), upon using the standard machine (see Definition 1.3.6), we see that (b) holds for all bounded Borel functions h_1 and h_2.

(b) implies (c): We know that $\mathcal{G} \subseteq \mathcal{H} = \sigma(\mathcal{G}, \sigma(X_2))$ and have to show that for any $A \in \mathcal{H}$,

$$\mathbf{E}(\mathbf{E}(h_1(X_1)|\mathcal{G})I_A) = \mathbf{E}(h_1(X_1)I_A).$$

We use Dynkin's π-λ theorem for the generating class $\{A = B \cap C : B \in \mathcal{G},\ C \in \sigma(X_2)\}$ of $\mathcal{H}$, which is closed under finite intersection. Then, by (b),

$$\mathbf{E}[h_1(X_1)I_B I_C] = \mathbf{E}[\mathbf{E}(h_1(X_1)I_C|\mathcal{G})I_B] = \mathbf{E}[\mathbf{E}(h_1(X_1)|\mathcal{G})\mathbf{E}(I_C|\mathcal{G})I_B]\,.$$

Thus, using the tower property and taking out the $\mathcal{G}$-measurable $I_B\mathbf{E}(h_1(X_1)|\mathcal{G})$,

$$\mathbf{E}\{\mathbf{E}(h_1(X_1)|\mathcal{G})I_A\} = \mathbf{E}\{\mathbf{E}[\mathbf{E}(h_1(X_1)|\mathcal{G})I_B I_C|\mathcal{G}]\} = \mathbf{E}(h_1(X_1)I_A).$$

(c) implies (a): By the tower property,

$$\mathbf{P}(X_1 \in B_1, X_2 \in B_2|\mathcal{G}) = \mathbf{E}[\mathbf{E}(I_{B_1}(X_1)I_{B_2}(X_2)|\mathcal{H})\,|\,\mathcal{G}].$$

As $I_{B_2}(X_2)$ is $\mathcal{H}$-measurable, using (c) for $h(x) = I_{B_1}(x)$ and taking out the $\mathcal{G}$-measurable $\mathbf{E}(I_{B_1}(X_1)|\mathcal{G})$ we also have that

$$\begin{aligned} \mathbf{E}[\mathbf{E}(I_{B_1}(X_1)I_{B_2}(X_2)|\mathcal{H}) \,|\, \mathcal{G}] &= \mathbf{E}[I_{B_2}(X_2)\mathbf{E}(I_{B_1}(X_1)|\mathcal{H}) \,|\, \mathcal{G}] \\ &= \mathbf{E}[I_{B_2}(X_2)\mathbf{E}(I_{B_1}(X_1)|\mathcal{G}) \,|\, \mathcal{G}] = \prod_{i=1,2} \mathbf{P}(X_i \in B_i|\mathcal{G}). \end{aligned}$$