Stat 310A/Math 230A Theory of Probability

Homework 4 Solutions

Andrea Montanari

Due on October 21, 2015

Exercises on independent random variables and product measures

Exercise [1.3.65]

1. Let $Z_n = \sum_{k=0}^n Y_k$ for $Y_k \geq 0$. Since $Z_n \uparrow Z_\infty$, it follows by monotone convergence and linearity of the expectation that as $n \to \infty$
$$\sum_{k=0}^n \mathbf{E} Y_k = \mathbf{E} Z_n \uparrow \mathbf{E} Z_\infty = \mathbf{E}(\sum_{k=0}^\infty Y_k)$$
In particular, since A_k are disjoint sets, for $Y_k = X_+ I_{A_k} \geq 0$ (with $X_+ = \max(X, 0)$), we have that $Z_\infty = X_+ I_A$. Consequently,
$$\sum_{k=0}^\infty \mathbf{E}(X_+ I_{A_k}) = \mathbf{E} X_+ I_A \,.$$
Similarly, for $X_- = -\min(X, 0)$,
$$\sum_{k=0}^\infty \mathbf{E}(X_- I_{A_k}) = \mathbf{E} X_- I_A \,.$$
Since $\mathbf{E}|X| < \infty$ it follows that $\mathbf{E} X_+ I_A + \mathbf{E} X_- I_A = \mathbf{E}|X| I_A < \infty$ as well. It thus follows by linearity of the expectation that for $n \to \infty$,
$$\sum_{k=0}^n \mathbf{E}(X I_{A_k}) = \sum_{k=0}^n \mathbf{E}(X_+ I_{A_k}) - \sum_{k=0}^n \mathbf{E}(X_- I_{A_k}) \to \mathbf{E} X_+ I_A - \mathbf{E} X_- I_A = \mathbf{E} X I_A$$
(recall that $X = X_+ - X_-$). By Jensen's inequality, $|\mathbf{E}(X I_{A_k})| \leq \mathbf{E}|X| I_{A_k}$ and so by the same argument as before,
$$\sum_{k=0}^\infty |\mathbf{E}(X I_{A_k})| \leq \sum_{k=0}^\infty \mathbf{E}|X| I_{A_k} = \mathbf{E}|X| I_A < \infty \,,$$
namely, $\mathbf{E} X I_{A_n}$ is absolutely summable.

2. $\mathbf{Q}(A) \geq 0$ and $\mathbf{Q}(\Omega) = 1$, with countable additivity of $\mathbf{Q}$ shown in part (a).

3. Per our assumption $\mathbf{E} X = \mathbf{E} Y$. If $\mathbf{E} X = \mathbf{E} Y = 0$ then both $X = 0$ a.s. and $Y = 0$ a.s. so we are done. Otherwise, the probability measures $\mathbf{Q}_X(A) = \mathbf{E} X I_A / \mathbf{E} X$ and $\mathbf{Q}_Y(A) = \mathbf{E} Y I_A / \mathbf{E} Y$ of part (b) agree on the π-system $\mathcal{A}$, hence they must agree on $\mathcal{F} = \sigma(\mathcal{A})$. Considering the events $A_n = \{\omega : X(\omega) - Y(\omega) \geq 1/n\}$ we thus have that for every n,
$$0 = (\mathbf{E} X)[\mathbf{Q}_X(A_n) - \mathbf{Q}_Y(A_n)] = \mathbf{E}[(X - Y) I_{A_n}] \geq n^{-1} \mathbf{P}(A_n) \,,$$
which means that $\mathbf{P}(A_n) = \mathbf{P}(X - Y \geq 1/n) = 0$. It follows that $\mathbf{P}(A) = 0$ for $A := \bigcup_n A_n = \{\omega : X(\omega) - Y(\omega) > 0\}$. Reversing the roles of X and Y the same argument shows that also $\mathbf{P}(X < Y) = 0$, so $X \overset{a.s.}{=} Y$.

Exercise [1.4.15]

(a). It is not hard to check that each of the three pairs of random variables, namely (Z_0, Z_1), (Z_0, Z_2) and (Z_1, Z_2) take all values $\omega \in \Omega$ with equal probability (of 1/9). This of course implies that Z_0, Z_1 and Z_2 are pairwise independent. However, easy to check that also $(Z_0 + Z_1 + Z_2) \mod 3 = 0$, so as stated Z_0 and Z_1 determine the value of Z_2.
(b). Let X_1, X_2, X_3, X_4 be mutually independent and take values 1 and -1 with probability $\frac{1}{2}$ each. Let $Y_1 = X_1X_2$, $Y_2 = X_2X_3$, $Y_3 = X_3X_4$, $Y_4 = X_4X_1$. It is easy to see that $\mathbf{P}(Y_i = 1) = \mathbf{P}(Y_i = -1) = \frac{1}{2}$. Since $Y_1Y_2Y_3Y_4 = 1$, $\mathbf{P}(Y_1 = Y_2 = Y_3 = 1, Y_4 = -1) = 0$ so the four random variable are not mutually independent. To check that any three are mutually independent, it suffices by symmetry to check only for Y_1, Y_2, Y_3. Let $i_1, i_2, i_3 \in \{-1, 1\}$, noting that since X_i are mutually independent, with $(X_1, \dots, X_4)$ taking each value in $\{-1, 1\}^4$ with probability $\frac{1}{16}$, we have that

$$\begin{aligned}\mathbf{P}(Y_k = i_k, k = 1, 2, 3) &= \sum_{x_2 \in \{-1,1\}} \mathbf{P}(X_2 = x_2)\mathbf{P}(X_1 = i_1x_2, X_3 = i_2x_2, X_4 = i_3i_2x_2) \\ &= \frac{1}{8} = \mathbf{P}(Y_1 = i_1)\mathbf{P}(Y_2 = i_2)\mathbf{P}(Y_3 = i_3).\end{aligned}$$

Since this applies for any $i_1, i_2, i_3 \in \{-1, 1\}$ we established the independence of Y_1, Y_2, and Y_3 (for example, add over $i_1 \le y_1$, $i_2 \le y_2$, $i_3 \le y_3$ and apply Corollary 1.4.12).

Exercise [1.4.18]

We will use the following fact repeatedly: the probability that X is divisible by j is

$$\mathbf{P}(D_j) = \sum_{k=1}^{\infty} (kj)^{-s}/\zeta(s) = j^{-s}.$$

1. Let p_i be an enumeration of distinct primes. By definition, it suffices to show that for any finite sub-collection $\{p_j\}_{j=1}^n$ and any $n < \infty$,

   $$\mathbf{P}(\bigcap_{j=1}^{n} D_{p_j}) = \prod_{j=1}^{n} \mathbf{P}(D_{p_j}).$$

   Indeed, all the primes p_j in the sub-collection divide k if and only if k is divisible by their product. So the LHS is just $(\prod_{i=1}^n p_j)^{-s}$ which by the fact we derived before, equals the RHS.

2. Euler's formula is the statement that $\{X = 1\}$ if and only if X is not divisible by any prime number. Indeed, as the latter event is $\bigcap_p D_p^c$, by continuity from below of $\mathbf{P}(\cdot)$ we have that

   $$\frac{1}{\zeta(s)} = \mathbf{P}(X = 1) = \lim_{n \to \infty} \mathbf{P}(\bigcap_{j=1}^{n} D_{p_j}^c).$$

   In part (a) we verified the mutual independence of the collections $\{D_p\}$, p prime, each of which is trivially a π-system. This implies the mutual independence of $\sigma(D_p)$, p prime (see Corollary 1.4.7), hence that of the events D_p^c. With $\mathbf{P}(D_p) = p^{-s}$, we get that $\mathbf{P}(\bigcap_{j=1}^n D_{p_j}^c) = \prod_{j=1}^n (1 - p_j^{-s})$ leading to Euler's formula when taking $n \to \infty$.

3. The event that no perfect square other than 1 divides X, is precisely the event that p^2 does not divide X for every prime p, which is $\bigcap_p D_{p^2}^c$. Similarly to part (a), it is not hard to verify that $\{D_{p^2}\}$, p prime, are mutually independent, hence so are $\{D_{p^2}^c\}$, p prime. As in part (b), this leads to the probability of the event of interest being $\prod_p (1 - p^{-2s})$, which by Euler's formula is $1/\zeta(2s)$.

4. Let $c = \mathbf{P}(G = 1)$ denote the probability that the i.i.d. variables X and Y have no common factors. Applying the elementary conditioning formula $\mathbf{P}(A|B) = \mathbf{P}(A \cap B)/\mathbf{P}(B)$, it follows from the definition of the law of X, that the conditional law of X/k given that X is a multiple of k, is the same as the original law of X. Therefore, given that X and Y are both multiples of k, an event whose probability is $\mathbf{P}(D_k)^2 = k^{-2s}$, the probability that X/k and Y/k have no common factors is precisely c. Consequently, by the same elementary formula, we deduce that $\mathbf{P}(G = k) = ck^{-2s}$ for $k = 1, 2, \ldots$. Since $\sum_k \mathbf{P}(G = k) = 1$ it follows that $c = 1/\zeta(2s)$, as stated.

Exercises on L_p spaces

1. Obviously $X \simeq X$ and $X \simeq Y$ implies $Y \simeq X$ (here and below $\simeq$ denotes the equivalence relation). To prove transitivity, consider X, Y, Z random variables on $(\Omega, \mathcal{F}, \mathbf{P})$ and let $\Omega_1 = \{\omega : X(\omega) \neq Y(\omega)\}$, $\Omega_2 = \{\omega : Y(\omega) \neq Z(\omega)\}$. If $X \simeq Y$ and $Y \simeq Z$ then $\mathbf{P}(\Omega_1) = \mathbf{P}(\Omega_2) = 0$ and since $\{\omega : X(\omega) \neq Z(\omega)\} \subseteq \Omega_1 \cup \Omega_2$, we have $X \sim Z$.

2. First of all $L_p(\Omega, \mathcal{F}, \mathbf{P})$ is a linear space. Indeed if $X \simeq X'$ and $Y \simeq Y'$, then $(aX + bY) \simeq (aX' + bY')$, and $\mathbb{E}|X|^p, \mathbb{E}|Y|^p \leq \infty$ implies $\mathbb{E}|aX + bY|^p \leq \infty$.

To check that $|| \cdot ||_p$ is indeed a norm, recall the following elementary facts: (i) $\mathbb{E}|X|^p \geq 0$, with $\mathbb{E}|X|^p = 0$ if and only if $X = 0$ almost everywhere (i.e. $X \simeq 0$); (ii) $\mathbb{E}|aX|^p = a^p \mathbb{E}|X|^p$ for a non-negative (linearity of expectation); (iii) $||X + Y||_p \leq ||X||_p + ||Y||_p$ by Minkowski inequality.

3. Notice that the definition of $||X||_\infty$ only depends on te equivalence class of X. Conditions (i) and (ii) follow as above. For (iii) -triangular inequality- let $\Omega_1 = \{\omega : |X(\omega)| \leq ||X||_\infty\}$, $\Omega_2 = \{\omega : |Y(\omega)| \leq ||Y||_\infty\}$. By definition $\mathbf{P}(\Omega_1) = \mathbf{P}(\Omega_2) = 1$, whence $\mathbf{P}(\Omega_1 \cap \Omega_2) = 1$. Further, for $\omega \in \Omega_1 \cap \Omega_2$, $|X(\omega) + Y(\omega)| \leq |X(\omega)| + |Y(\omega)| \leq ||X||_\infty + ||Y||_\infty$. This implies $||X + Y||_\infty \leq ||X||_\infty + ||Y||_\infty$ (and -in passing- that $L_\infty(\Omega, \mathcal{F}, \mathbf{P})$ is indeed a vector space).

4. We can fix a representative of the equivalence class such that $|X(\omega)| \leq ||X||_\infty$ for any $\omega \in \Omega$. By monotonicity of the integral, we have $||X||_p \leq ||X||_\infty$ for any $p > 0$. Without loss of generality, we can assume $||X||_\infty > 0$. Fix $\varepsilon > 0$ and let $\Omega_\varepsilon = \{\omega : |X(\omega)| \geq (1 - \varepsilon)||X||_\infty\}$. By definition $\mathbb{P}(\Omega_\varepsilon) > 0$. Again by monotonicity of the integral

$$||X||_p \geq \mathbb{E}\{(1 - \varepsilon)^p ||X||_\infty^p \, I_{\Omega_\varepsilon}\}^{1/p} = (1 - \varepsilon)||X||_\infty \, \mathbb{P}(\Omega_\varepsilon)^{1/p} \, . \tag{1}$$

The right hand side converges to $(1 - \varepsilon)||X||_\infty$ as $p \to 0$. The thesis follows since ε is arbitrary.

5. For $n \geq 1$, let $X_n \equiv |X| \, I_{|X| \leq n}$. Then, for $p \leq q$,

$$0 \leq \mathbb{E}\{|X|^p\} - \mathbb{E}\{X_n^p\} \leq \mathbb{E}\{|X|^q\} - \mathbb{E}\{X_n^q\} \leq \delta_n \tag{2}$$

for some sequence $\delta_n \downarrow 0$ by monotone convergence. Further, by bounded convergence $\lim_{p \to 0} \mathbb{E}\{X_n^p\} = S(X)$, whence $|\limsup_{p \to 0} \mathbb{E}\{|X|^p\} - S(X)| \leq \delta_n$ and $|\liminf_{p \to 0} \mathbb{E}\{|X|^p\} - S(X)| \leq \delta_n$. The claim follows by letting $n \to \infty$.

6. For any random variable X, we can construct a sequence $\{X_n\} \subseteq \mathrm{SF}$ such that $X_n(\omega) \to X(\omega)$ for any $\omega \in \Omega$. Further, if $X \in L_\infty(\Omega, \mathcal{F}, \mathbf{P})$ (and choosing a representative of the equivalence class that is itself bounded), the standard construction yields a sequence such that $|X_n(\omega) - X(\omega)| \leq 1/n$ for all $\omega \in \Omega$. This proves the claim for $p = \infty$. It is therefore sufficient to prove that $L_\infty(\Omega, \mathcal{F}, \mathbf{P})$ is dense in $L_p(\Omega, \mathcal{F}, \mathbf{P})$. For $X \in L_p(\Omega, \mathcal{F}, \mathbf{P})$ and $M > 0$, let $X_M = \mathrm{sign}(X) \max(|X|, M)$. By monotone convergence $\lim_{M \to \infty} \mathbb{E}\{|X - X_M|^p\} = 0$, which finishes the proof.

Completeness

We give a proof for $p < \infty$ (the case $p = \infty$ is simpler) which uses the following fact. fact A normed space is complete if, for any sequence $\{Z_n\}$ such that $\sum_{n=1}^{\infty} ||Z_n|| < \infty$, the sequence of sums $W_n \equiv \sum_{i=1}^{n} Z_i$ converges (in the topology induced by the norm $||\cdot||$). fact Consider next a sequence of random variables $Z_n \in L_p(\Omega, \mathcal{F}, \mathbb{P})$ as in this statement and let $M \equiv \sum_{n=1}^{\infty} ||Z_n||_p < \infty$. By triangular inequality $||(\sum_{n=1}^{m} |Z_n|)||_p \leq M$, and monotone convergence implies $||(\sum_{n=1}^{\infty} |Z_n|)||_p \leq M$. This implies that $\sum_{n=1}^{\infty} |Z_n(\omega)|$ is finite for almost every ω and therefore $W_n(\omega) = \sum_{i=1}^{\infty} Z_i(\omega)$ converges absolutely for almost every ω. Call W the limit, which is obviously measurable and in $L_p(\Omega, \mathcal{F}, \mathbf{P})$ (with $||W||_p \leq M$).

For any ω

$$|W(\omega) - W_n(\omega)| \leq \sum_{i=n+1}^{\infty} |Z_i(\omega)| \,. \tag{3}$$

The right hand side converges to 0 monotonically as $n \to \infty$, and therefore the thesis follows by monotone convergence.

Proof of the Fact

Let $\{X_k\}_{k\in\mathbb{N}}$ be a Cauchy sequence. It is clearly sufficient to exhibit a subsequence $\{X_{k(n)}\}_{n\in\mathbb{N}}$ that converges. In order to achieve this goal, let $k(n)$ be the smallest integer such that $||X_i - X_j|| \leq 2^{-n}$ for all $i, j \geq k(n)$. Define $Z_n \equiv X_{k(n+1)} - X_{k(n)}$. Clearly $||Z_n|| \leq 2^{-n}$ and therefore $\sum_{n=1} ||Z_n|| < \infty$. By hypothesis $W_n = \sum_{i=1}^{n} Z_i$ converges and therefore the subsequence $X_{k(n)} = X_k(1) + W_{n-1}$ converges as well.