Liam Hardiman *October 23, 2019*

271A - Homework 2

1. Use Jensen's inequality to show that for $p \geq 1$.

$$\|\mathbb{E}[X|\mathcal{G}]\|_{L^p} \leq \|X\|_{L^p}.$$

Proof. Let's look at the p-norm to the p-th power.

$$\begin{aligned}
\|\mathbb{E}[X|\mathcal{G}]\|_{L^p}^p &= \int |\mathbb{E}[X|\mathcal{G}]|^p \, d\mathbb{P} \\
&\leq \int \mathbb{E}[|X|^p|\mathcal{G}] \, d\mathbb{P} \quad \text{(by Jensen's inequality)} \\
&= \int |X|^p \, d\mathbb{P} \quad \text{(by definition of conditional expectation)} \\
&= \|X\|_{L^p}^p.
\end{aligned}$$

Taking the p-th root of both sides establishes the claim. □

2. Let $(X_n : n \in \mathbb{N})$ be a sequence of independent random variables, each exponentially distributed:

$$\mathbb{P}[X_n > x] = e^{-x}, \quad x \geq 0.$$

(a) A random variable τ has the lack of memory property if

$$\mathbb{P}[\tau > a + b \mid \tau > a] = \mathbb{P}[\tau > b].$$

Show that a random variable has the memoryless property if and only if it is exponentially distributed.

Proof. Suppose τ is exponentially distributed, i.e.

$$\mathbb{P}[\tau > x] = \begin{cases} e^{-\lambda x}, & \text{if } x \geq 0, \\ 1, & \text{if } x < 0. \end{cases}$$

By the definition of conditional probability we have

$$\mathbb{P}[\tau > a + b \mid \tau > a] = \frac{\mathbb{P}[(\tau > a + b) \wedge (\tau > a)]}{\mathbb{P}[\tau > a]}$$

Now if $b \geq 0$, $\mathbb{P}[(\tau > a+b) \wedge (\tau > a)] = \mathbb{P}[\tau > a+b]$. This gives

$$\begin{aligned}\mathbb{P}[\tau > a+b \mid \tau > a] &= \frac{\mathbb{P}[\tau > a+b]}{\mathbb{P}[\tau > a]} \\ &= \frac{e^{-\lambda(a+b)}}{e^{-\lambda a}} \\ &= e^{-\lambda b} \\ &= \mathbb{P}[\tau > b].\end{aligned}$$

On the other hand, if $b < 0$, $\mathbb{P}[(\tau > a+b) \wedge (t > a)] = \mathbb{P}[\tau > a]$, which gives

$$\begin{aligned}\mathbb{P}[\tau > a+b \mid \tau > a] &= \frac{\mathbb{P}[\tau > a]}{\mathbb{P}[\tau > a]} \\ &= 1 \\ &= \mathbb{P}[\tau > b].\end{aligned}$$

Conversely, suppose that τ is memoryless. If b is positive then we have

$$\mathbb{P}[\tau > a+b] = \mathbb{P}[\tau > a] \cdot \mathbb{P}[\tau > b].$$

If we let $F(x) = \mathbb{P}[\tau > x]$, then F satisfies the exponential property:

$$F(a+b) = F(a)F(b).$$

Setting $a = b$, we have $F(2a) = F(a)^2$. Inductively, we obtain $F(na) = F(a)^n$ for any positive integer n. Taking the n-th root of both sides gives $F(a/n) = F(a)^{1/n}$. Combining these gives $F(\frac{m}{n}a) = F(a)^{m/n}$. For any rational $\frac{m}{n}$.

Since any real number is a limit of rational numbers, we obtain $F(ra) = F(a)^r$ for any real r by continuity. Since this holds for any $a \geq 0$, we can set $a = 1$ to obtain $F(r) = F(1)^r$, so F, the distribution of τ, is exponential. $\square$

(b) Compute the expectation and variance of X_n. Let $Y = X_n + X_{n+1}$. Find the correlation coefficient between Y and X_n. Find $\mathbb{E}[Y|X_{n+1}]$.

Solution. Let's compute the first two moments.

$$\begin{aligned}\mathbb{E}[X_n] &= \int_0^\infty e^{-x}\, dx \\ &= 1.\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}[X_n^2] &= 2\int_0^\infty xe^{-x}\ dx \\
&= 2.
\end{aligned}
$$

We then have $\mathbb{E}[X_n] = 1$ and $\mathrm{Var}[X_n] = \mathbb{E}[X_n^2] - \mathbb{E}[X_n]^2 = 1$.

Let $Y = X_n + X_{n+1}$. We'll need the mean and variance of Y to compute the correlation coefficient.

$$
\mathbb{E}[Y] = \mathbb{E}[X_n] + \mathbb{E}[X_{n+1}] = 2.
$$

Since X_n and X_{n+1} are independent, we also have

$$
\mathrm{Var}[Y] = \mathrm{Var}[X_n] + \mathrm{Var}[X_{n+1}] = 2.
$$

Now let's compute the correlation coefficient.

$$
\begin{aligned}
\rho_{Y,X_{n+1}} &= \frac{\mathbb{E}[(Y-\mu_Y)(X_{n+1}-\mu_{n+1})]}{\sigma_Y\sigma_{n+1}} \\
&= \frac{\mathbb{E}[YX_{n+1}] - \mu_{n+1}\mathbb{E}[Y] - \mu_Y\mathbb{E}[X_{n+1}] + \mu_Y\mu_{n+1}}{\sqrt{2}} \\
&= \frac{\mathbb{E}[X_nX_{n+1}] + \mathbb{E}[X_{n+1}^2] - 2 - 2 + 2}{\sqrt{2}} \\
&= \frac{1}{\sqrt{2}}.
\end{aligned}
$$

Now for the conditional expectation. Since conditional expectation is linear and X_n and X_{n+1} are independent,

$$
\begin{aligned}
\mathbb{E}[Y|X_{n+1}] &= \mathbb{E}[X_n|X_{n+1}] + \mathbb{E}[X_{n+1}|X_{n+1}] \\
&= \mathbb{E}[X_n] + X_{n+1} \\
&= 1 + X_{n+1}.
\end{aligned}
$$

□

(c) Show that

$$
\mathbb{P}[X_n > \alpha\log n \text{ for infinitely many } n] = \begin{cases} 0 & \text{for } \alpha > 1, \\ 1 & \text{else} \end{cases}.
$$

Proof. Let E_n be the event $E_n = \{X_n > \alpha\log n\}$. Let's sum these events

$$
\begin{aligned}
\sum_{n=1}^\infty \mathbb{P}[E_n] &= \sum_{n=1}^\infty e^{-\alpha\log n} \\
&= \sum_{n=1}^\infty \frac{1}{n^\alpha}.
\end{aligned}
$$

This is summable if and only if $\alpha > 1$. By Borel-Cantelli, we have

$$\mathbb{P}[\limsup E_n] = \mathbb{P}[X_n > \alpha \log n \text{ for infinitely many } n] = 0.$$

Since the X_n's are given to be independent, if the above sum diverges, $\mathbb{P}[\limsup E_n] = 1$ by the converse of Borel-Cantelli. $\square$

3. Let B_t be a standard Brownian motion.

(a) Find a matrix A so that in distribution $y = [B_{t_1}, B_{t_2}, \ldots, B_{t_n}]^T = A[Z_1, Z_2, \ldots, Z_n]^T$, where the Z_j's are iid standard normal random variables.

Solution. Call $x = [Z_1, \ldots, Z_n]^T$. The i-th component of the product Ax has distribution $(Ax)_i \sim \mathcal{N}(0, a_{i,1}^2 + \cdots + a_{i,n}^2)$. Since $y_i = B_{t_i} \sim \mathcal{N}(0, t_i)$ and we want $y = Ax$, we must then have

$$a_{i,1}^2 + \cdots + a_{i,n}^2 = t_i \tag{1}$$

Without loss of generality and for notational convenience, assume that $t_1 \leq t_2 \leq \cdots \leq t_n$. We know that $\text{cov}(B_{t_i}, B_{t_j}) = t_i \wedge t_j = t_{i \wedge j}$, so, we know that the matrix $\text{cov}(y)$ has entries $\text{cov}(y)_{i,j} = t_{i\wedge j}$. We also know that

$$\text{cov}(Ax) = A \cdot \text{cov}(x) \cdot A^T = AA^T.$$

Since we want $y = Ax$, we must then have $(AA^T)_{i,j} = t_{i\wedge j}$. Note that this implies (1).

In summary, we've reduced the problem to finding A subject to $(AA^T)_{i,j} = t_{i\wedge j}$. The matrix $t_{i\wedge j}$ is clearly symmetric and it's positive definite by our reordering of the t_i's and Sylvester's criterion. If we define the matrix T to have entries $(T)_{i,j} = t_{i\wedge j}$, we can diagonalize T:

$$T = UDU^T,$$

where U is orthogonal and D is diagonal with nonnegative entries. If we set $A = UD^{1/2}$, then $AA^T = T$ as desired. $\square$

(b) Show that $e^{-at}B_{e^{2at}}$ is a Gaussian process. Find its covariance function.

Solution. This immediately follows from part (a). For any $t_1, \ldots, t_n$ there is a matrix A so that

$$[B_{e^{2at_1}}, \ldots, B_{e^{2at_n}}]^T = A[Z_1, \ldots, Z_n]^T,$$

where the Z_i's are iid standard normal random variables. Multiply both sides by the diagonal matrix whose i-th entry is e^{-at_i} and we're done.

Let's compute the covariance function.

$$\begin{aligned}\mathrm{cov}(s,t) &= \mathrm{cov}(e^{-as}B_{e^{2as}}, e^{-at}B_{e^{2at}})\\ &= e^{-a(s+t)}\mathbb{E}[B_{e^{2as}}B_{e^{2at}}]\\ &= e^{-a(s+t)}(e^{2as}\wedge e^{2at}).\end{aligned}$$

□

4. Consider the random vector $x = [X_1, \ldots, X_n] \in \mathbb{R}^n$.

   (a) Prove Chebyshev's inequality: for $p > 0$ we have

   $$\mathbb{P}\left[\sum_i |X_i|^p \geq \lambda^p\right] \leq \lambda^{-p}\mathbb{E}\left[\sum_i |X_i|^p\right].$$

   Proof. Let's integrate.

   $$\begin{aligned}\lambda^{-p}\mathbb{E}\left[\sum_i |X_i|^p\right] &\geq \lambda^{-p}\mathbb{E}\left[\sum_i |X_i|^p \cdot \mathbb{1}_{\sum_i |X_i|^p \geq \lambda^p}\right]\\ &\geq \lambda^{-p}\cdot\lambda^p\cdot\mathbb{P}\left[\sum_i |X_i|^p \geq \lambda^p\right].\end{aligned}$$

   □

   (b) Suppose there exists $k > 0$ so that

   $$M = \mathbb{E}[\exp(k(\sum_i |X_i|^p)^{1/p})] < \infty.$$

   Prove that $\mathbb{P}[\sum_i |X_i|^p \geq \lambda^p] \leq Me^{-k\lambda}$ for all $\lambda \geq 0$.

   Proof. The function $t \mapsto e^{kt^{1/p}}$ is nondecreasing, so if we let $E_\lambda = \{\sum_i |X_i|^p \geq \lambda^p\}$, then

   $$0 \leq e^{k\lambda}\mathbb{1}_{E_\lambda} \leq \exp(k(\sum_i |X_i|^p)^{1/p})]\mathbb{1}_{E_\lambda}.$$

   Now we just take the expectation.

   $$e^{k\lambda}\mathbb{P}[E_\lambda] \leq M.$$

   □

5. Let $\Omega = \{1,2,3,4,5\}$ and let $\mathcal{U}$ be the collection

   $$\mathcal{U} = \{\{1,2,3\},\{3,4,5\}\},$$

   of subsets of Ω.

(a) Find $\sigma(\mathcal{U})$, the σ-algebra generated by $\mathcal{U}$.

Solution. Taking complements gives us $\{4,5\}$ and $\{1,2\}$. Intersecting gives $\{3\}$. We can take a union to get $\{1,2,4,5\}$. We can't get anything new from unions, intersections, or complements here, so we conclude that

$$\sigma(\mathcal{U}) = \{\{1,2,3\},\{3,4,5\},\{3\},\{4,5\},\{1,2\},\{1,2,4,5\},\Omega,\emptyset\}.$$

□

(b) Define a random variable by $X(1) = X(2) = 0$, $X(3) = 10$, $X(4) = X(5) = 1$. Is X measurable wrt $\sigma(\mathcal{U})$?

Solution. X is $\sigma(\mathcal{U})$ measurable.

$$X^{-1}(-\infty,\alpha] = \begin{cases} \emptyset, & \alpha < 0 \\ \{1,2\}, & \alpha \in [0,1) \\ \{1,2,4,5\}, & \alpha \in [1,10) \\ \Omega, & \alpha \geq 10 \end{cases}.$$

Each of these preimages is measurable, so X is measurable. □

(c) Define another random variable by $Y(1) = 0$, $Y(2) = Y(3) = Y(4) = Y(5) = 1$. Find the σ-algebra generated by Y.

Solution. $\sigma(Y)$ is the σ-algebra generated by sets of the form $Y^{-1}(-\infty,\alpha]$ for $\alpha \in \mathbb{R}$. As α ranges of $\mathbb{R}$, we pick up the sets $\emptyset$, $\{1\}$, and Ω. Taking complements gives

$$\sigma(Y) = \{\emptyset,\{1\},\{2,3,4,5\},\Omega\}.$$

□