Probability Theory II (G6106)

Spring 2016
http://stat.columbia.edu/∼porbanz/G6106S16.html

Peter Orbanz
porbanz@stat.columbia.edu
Florian Stebegg
florian@stat.columbia.edu

Homework 7

Due: 6 April 2016

Problem 1 (Random variables contract under conditioning)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and let $X : \Omega \to \mathbb{R}$ be a random variable in $L_2(\mathbb{P})$ (that is, whose $L_2(\mathbb{P})$-norm is $\|X\|_2 < \infty$).

Question: Show that $\|\mathbb{E}[X|\mathcal{C}]\|_2 \leq \|X\|_2$.

Remark: This is in fact true for every L_p-norm with $p > 1$, and is related to the well-known phenomenon that conditioning reduces variance (as e.g. in the Rao-Blackwell theorem), and more loosely to the fact that conditioning reduces entropy.

Problem 2 (Pairs of random variables)

Consider the probability space $([0,1], \mathcal{B}([0,1]), \lambda)$, where λ is Lebesgue measure.

Question: Give an example of real-valued random variables X, X' and Y on $[0,1]$ such that X and X' are identically distributed, but (X', Y) and (X, Y) are not.

Hint: This is every bit as easy as it seems—no horseshoes.

Problem 3 (Bayes' theorem)

You will have encountered Bayes' theorem before. In this problem, we ask you to prove the formal version of this result, using the existence theorem for conditional densities.

Question (a): Read Section 3.7 of the class notes; you will need Theorem 3.26.

The formal statement of the theorem is as follows:

Theorem 1 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $\Theta : \Omega \to \mathbf{T}$ a random variable taking values in a Borel space $\mathbf{T}$, with law Q. Let $X_i : \Omega \to \mathbf{X}$, for $i \in \mathbb{N}$, be random variables with values in a Borel space $\mathbf{X}$, which are conditionally iid, that is: All X_i have identical conditional distribution*

$$\mathbf{p}(\bullet, \theta) =_{\text{a.s.}} \mathbb{P}[X \in \bullet | \Theta = \theta] \ , \tag{1}$$

where $\mathbf{p} : \mathbf{T} \to \mathbf{PM}(\mathbf{X})$ is a probability kernel, and the joint conditional distribution factorizes as

$$\mathbb{P}[X_1 \in A_1, \ldots, X_n \in A_n | \Theta = \theta] =_{\text{a.s.}} \prod_{i=1}^{n} \mathbf{p}(A_i, \theta) \ . \tag{2}$$

Require that there exists a σ-finite measure μ on $\mathbf{X}$ such that the absolute continuity relation

$$\mathbf{p}(\bullet, \theta) \ll \mu \qquad \textit{for all } \theta \in \mathbf{T} \tag{3}$$

is satisfied. Then conditional distribution of Θ given $X_1, \ldots, X_n$ is given by

$$Q[d\theta | X_1 = x_1, \ldots, X_n = x_n] = \frac{\prod_{i=1}^{n} p(x_i|\theta)}{p(x_1, \ldots, x_n)} Q(d\theta) \ , \tag{4}$$

where $p(x|\theta)$ is the conditional density of $\mathbf{p}$ guaranteed by Theorem 3.26, and

$$p(x_1, \ldots, x_n) := \int_{\mathbf{T}} \prod_{i=1}^{n} p(x_i|\theta) Q(d\theta) \; . \tag{5}$$

Moreover, $\mathbb{P}\{p(X_1, \ldots, X_n) \in \{0, \infty\}\} = 0$.

Question (b): Prove Bayes' theorem.

Remark: Bayes' theorem is often stated in terms of densities as $p(\theta|x) = \frac{p(x|\theta)}{p(x)} p(\theta)$ (for $n = 1$), which is perfectly safe if, for example, X and Θ both take values in Euclidean space and have smooth distributions. In general, we have to be a bit more careful: Equation (4) above is a representation of the conditional law $\mathcal{L}(\Theta|X_1, \ldots, X_n)$ by a density with respect to $\mathcal{L}(\Theta)$. To ensure that such a density exists, we have to verify absolute continuity of $\mathcal{L}(\Theta|X_1, \ldots, X_n)$ with respect to $\mathcal{L}(\Theta)$. The theorem shows that whether this absolute continuity is satisfied depends only on the conditional distribution of X, via (3).

Problem 4 (Rejection Sampling)

Let P and Q be two probability measures on a countable discrete space $\mathbf{X}$. Suppose there is a constant $c > 0$ such that

$$f(x) := \frac{Q(\{x\})}{P(\{x\})} \leq c \qquad \text{for all } x \in \mathbf{X} \text{ with } P(\{x\}) > 0 \; . \tag{6}$$

Let $X_1, X_2, \ldots$ be i.i.d. random variables with law P, and $U_1, U_2, \ldots$ i.i.d. uniform variables in $[0, 1]$. Now define an integer random variable N as the smallest value of n such that

$$U_n \leq \frac{f(X_n)}{c} \; . \tag{7}$$

Question: Show that the random variable X_N has law Q.