Existence of conditional expectation

Nate Eldredge

October 1, 2010

These notes will describe some proofs of the existence of conditional expectation, which we omitted in class.

Theorem 0.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, X an integrable random variable, and $\mathcal{G} \subset \mathcal{F}$ a σ-field. There exists a random variable Y, denoted $Y = E[X|\mathcal{G}]$, such that:*

1. *$Y \in \mathcal{G}$; and*
2. *For all events $A \in \mathcal{G}$, $E[X; A] = E[Y; A]$.*

We proved in class that Y, if it exists, is unique.

1 Using the Radon-Nikodym theorem

See Durrett section 5.1. Note that his $\mathcal{F}$ is our $\mathcal{G}$ (and his $\mathcal{F}_0$ is our $\mathcal{F}$); that is, we are applying the Radon-Nikodym theorem in the measure space $(\Omega, \mathcal{G}, P)$.

A proof of the Radon-Nikodym theorem in full generality appears in Durrett's appendix A.4.

Durrett has a much more general version of the Radon-Nikodym theorem than is needed. For a more direct approach, see `http://www.math.sc.edu/~schep/Radon-update.pdf`. To get conditional expectation from it, take $X = \Omega$, $\mathcal{B} = \mathcal{G}$, $\mu = P$, and $\nu(A) = E[X; A]$ (this produces a measure on $\mathcal{G}$, which is easily checked with dominated convergence).

2 Using Hilbert space techniques

Let's first reduce to the bounded nonnegative case.

Lemma 2.1. *If Theorem 0.1 holds for bounded nonnegative random variables, then it holds for all integrable random variables.*

Proof. Suppose X is integrable and $X \geq 0$. Taking something like $X_n = X \wedge n$ gives a sequence of bounded random variables increasing to X; by assumption $E[X_n|\mathcal{G}]$ exists for each n. The conditional monotone convergence theorem (Durrett 5.1.2 (c); its proof did not rely on the existence of conditional expectation, only the uniqueness) then implies that $E[X|\mathcal{G}] = \lim E[X_n|\mathcal{G}]$ (i.e. the limit on the right side satisfies properties 1 and 2).

If X is merely integrable, then $E[X^+|\mathcal{G}]$ and $E[X^-|\mathcal{G}]$ exist by the previous step, and it's easy to check that their difference is $E[X|\mathcal{G}]$ (satisfies properties 1 and 2). $\square$

Recall the following theorem asserting the existence of orthogonal projections in Hilbert space.

Theorem 2.2. *Let H be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, $K \subset H$ a closed subspace, and $x \in H$. Then there exists $y \in K$ such that for any $z \in K$, $\langle x - y, z \rangle = 0$.*

We apply this theorem with $H = L^2(\Omega, \mathcal{F}, P)$, where the inner product is $\langle X, Y \rangle = E[XY]$. Let $K = L^2(\Omega, \mathcal{G}, P)$ be the subspace of all L^2 $\mathcal{G}$-measurable random variables. To see that K is closed, i.e. that an L^2 limit of $\mathcal{G}$-measurable random variables is still $\mathcal{G}$-measurable, suppose $Y_n \in \mathcal{G}$, $Y_n \to Y$ in L^2. By Chebyshev's inequality, $Y_n \to Y$ i.p.; by a theorem proved in class, there is a subsequence with $Y_{n_k} \to Y$ a.e. Thus Y is (a.e. equal to) a limit of $\mathcal{G}$-measurable random variables, which must be $\mathcal{G}$-measurable.

Now if $X \in L^2$ (i.e. $E(X^2) < \infty$), we can use the above theorem: there exists $Y \in K$ so that $\langle X - Y, Z \rangle = E[(X - Y)Z] = 0$ for all $Z \in K$. Y is certainly $\mathcal{G}$-measurable by definition of K, so that's property 1. For $A \in \mathcal{G}$, take $Z = 1_A$; then $E[(X - Y)1_A] = 0$, so that $E[X; A] = E[Y; A]$. That's property 2. Thus $Y = E[X|\mathcal{G}]$.

Bounded random variables are certainly in L^2, so by Lemma 2.1 we are done.