Last Time

- Uniform integrability
- Convergence of expectations
- Independence of events, σ-fields, random variables

Today's lecture: Sections 2.1, 2.2

CE for Finite RV's

- Suppose X and Y are RV's that take finitely many values
- Conditional probability that $X = x$ given that $Y = y$:

$$\mathbb{P}(X = x|Y = y) \doteq \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}$$

- Conditional expectation of X given that $Y = y$:

$$\mathbb{E}(X|Y = y) \doteq \sum_x x\mathbb{P}(X = x|Y = y)$$

- Let $f(y)$ denote $\mathbb{E}(X|Y = y)$ and define the **conditional expectation of X given Y** by the RV $f(Y) \equiv \mathbb{E}(X|Y)$
- That is, if $Y(\omega) = y$ then $\mathbb{E}(X|Y)(\omega) = \mathbb{E}(X|Y = y)$

Interpretation

- $E(X|Y)$ is best guess of the value of X given information about Y
- $E(X|Y)$ should be based on the information in Y
- $E(X|Y)$ should be close to X in some sense

Motivation for CE in the L^2 case

- Let X and Y be RV's on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$
- Define $\mathcal{H}_Y \doteq L^2(\Omega, \sigma(Y), \mathbb{P})$, i.e.

$$\mathcal{H}_Y = \{g(Y) : g \text{ is measurable and } \mathbb{E}(|g(Y)|^2) < \infty\}$$

- Goal: choose the RV $\mathbb{E}(X|Y)$ to be the *minimum mean-squared error estimate* of X based on the information in Y
- That is, want the RV $\mathbb{E}(X|Y)$ to satisfy:
  - $\mathbb{E}(X|Y) \in \mathcal{H}_Y$
  - $\mathbb{E}(X|Y)$ minimizes, over $W \in \mathcal{H}_Y$,

$$\mathbb{E}|(X - W)^2|$$

Existence of CE in the L^2 case

Proposition 2.1.2 shows that

- There exists a RV $W^* \in \mathcal{H}_Y$ for which the minimum MSE is attained, i.e.

$$\mathbb{E}|(X - W^*)^2| = \inf_{W \in \mathcal{H}_Y} \mathbb{E}|(X - W)^2|$$

- The minimizer W^* is unique (in the a.s. sense)
- Furthermore, the optimality of W^* is equivalent to the orthogonality property:

$$\mathbb{E}[(X - W^*)V] = 0, \text{ for all } V \in \mathcal{H}_Y$$

Hilbert Space Properties of L^2

The space $L^2(\Omega, \mathcal{F}, I\!P)$ is a **Hilbert space**

- Linear vector space
- Norm: $||X||_2 = (I\!E|X|^2)^{1/2}$
- Inner product: $(X, Y) = I\!E(XY)$
- Schwarz inequality: $(I\!E(XY))^2 \leq I\!E|X|^2 I\!E|Y|^2$
- Complete with respect to L^2-norm

Existence of C.E. in L^2 case is due to the existence of orthogonal projection in Hilbert spaces

Definition of CE in the L^2 case

- For $X \in L^2(\Omega, \mathcal{F}, I\!P)$ the **conditional expectation of X given Y**, denoted $I\!E(X|Y)$, is the (a.s.) unique RV which satisfies:
  - *Measurability*: $I\!E(X|Y) \in L^2(\Omega, \sigma(Y), I\!P)$
  - *Orthogonality*: For all $V \in L^2(\Omega, \sigma(Y), I\!P)$,

$$I\!E[(X - I\!E(X|Y))V] = 0$$

- The orthogonality condition is equivalent to

$$I\!E|(X - I\!E(X|Y))^2| = \inf_{W \in L^2(\Omega, \sigma(Y), I\!P)} I\!E|(X - W)^2|$$

- Note that if Y and $\tilde{Y}$ are such that $\sigma(Y) = \sigma(\tilde{Y})$ then $I\!E(X|Y) = I\!E(X|\tilde{Y})$. Therefore, $I\!E(X|Y) \equiv I\!E(X|\sigma(Y))$

Example (part of Exercise 2.1.4)

- Let $\Omega = \{a, b, c, d\}$, $\mathcal{F} = 2^{\Omega}$, with probability measure $I\!P$ given by $I\!P(\{a\}) = 1/2$, $I\!P(\{b\}) = 1/4$, $I\!P(\{c\}) = 1/6$, $I\!P(\{d\}) = 1/12$.
- Let $A = \{a, d\}$ and $B = \{b, c, d\}$
- Find $I\!E(I_A | I_B)$