Probability and Random Processes

Lecture 7

- Conditional probability and expectation
- Decomposition of measures

Conditional Probability

- A probability space $(\Omega, \mathcal{A}, P)$
- An event $F \in \mathcal{A}$ with $P(F) > 0$; the σ-algebra generated by F, $\mathcal{G} = \sigma(\{F\}) = \{\emptyset, F, F^c, \Omega\}$
- Elementary conditional probability of $E \in \mathcal{A}$ given F

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

- The conditional probability of $E \in \mathcal{A}$ conditioned on $\mathcal{G}$ = "the probability of E knowing which events in $\mathcal{G}$ occurred" = "probability of E knowing whether F or F^c occurred"

$$P(E|\mathcal{G}) = P(E|F)\chi_F(\omega) + P(E|F^c)\chi_{F^c}(\omega)$$

a function : $\Omega \to \mathbb{R}$

- Note that $P(E|\mathcal{G})$
  - is a random variable on $(\Omega, \mathcal{A}, P)$;
  - is $\mathcal{G}$-measurable;

  and that

$$P(G \cap E) = \int_G P(E|\mathcal{G}) dP, \;\; G \in \mathcal{G}$$

- A basis for generalizing $P(E|\mathcal{G})$ to conditioning on arbitrary σ-algebras

- Given $(\Omega, \mathcal{A}, P)$, $E \in \mathcal{A}$ and $\mathcal{G} \subset \mathcal{A}$, there exists a nonnegative $\mathcal{G}$-measurable function $P(E|\mathcal{G})$ such that

$$P(G \cap E) = \int_G P(E|\mathcal{G}) dP, \;\; G \in \mathcal{G}$$

  Also, $P(E|\mathcal{G})$ is unique P-a.e.
- Proof: Define $\mu_E(G) = P(G \cap E)$ for any $G \in \mathcal{G}$, then $\mu_E \ll P$ and

$$P(E|\mathcal{G}) = \frac{d\mu_E}{dP}$$

- The function $P(E|\mathcal{G})$ is called the conditional probability of E given $\mathcal{G}$
  - "the probability of E knowing which events in $\mathcal{G}$ occurred"

- Again, for fixed $\mathcal{G}$ and E, the entity $P(E|\mathcal{G})$ is a *function* $f(\omega) = P(E|\mathcal{G})(\omega)$ on Ω
- Alternatively, by instead fixing $\mathcal{G}$ and ω we get a set function

$$m(E) = P(E|\mathcal{G})(\omega), \quad E \in \mathcal{A}$$

- If $m(E)$ is a probability measure on $(\Omega, \mathcal{A})$ then $P(E|\mathcal{G})$ is said to be regular
  - $P(E|\mathcal{G})$ is in general not necessarily regular...
- If the space $(\Omega, \mathcal{A})$ is standard (more about this later in the course), then $m(E)$ is a probability measure

Conditioning on a Random Variable

- Given $(\Omega, \mathcal{A}, P)$ and a random variable X, let $\sigma(X) =$ smallest $\mathcal{F} \subset \mathcal{A}$ such that X is (still) measurable w.r.t. $\mathcal{F}$ = the σ-algebra generated by X,
  - $\sigma(X)$ is exactly the class of events for which you can get to know whether they occured or not by observing X
- The conditional probability of $E \in \mathcal{A}$ given X is defined as

$$P(E|X) = P(E|\sigma(X))$$

Signed Measure

- Given a measurable space $(\Omega, \mathcal{A})$, a signed measure ν on $\mathcal{A}$ is an extended real-valued function such that
  - $\nu(\emptyset) = 0$
  - for a sequence $\{A_i\}$ of pairwise disjoint sets in $\mathcal{A}$

$$\nu\left(\bigcup_i A_i\right) = \sum_i \nu(A_i)$$

  (i.e., simply a measure that doesn't need to be positive)

Radon–Nikodym for Signed Measures

- If μ is a σ-finite measure and ν a finite signed measure on $(\Omega, \mathcal{A})$, and also $\nu \ll \mu$, then there is an integrable real-valued $\mathcal{A}$-measurable function f on Ω such that

$$\nu(A) = \int_A f d\mu$$

  for any $A \in \mathcal{A}$. Furthermore, f is unique μ-a.e.
- The function f is the Radon–Nikodym derivative of ν w.r.t. μ, notation $f = \frac{d\nu}{d\mu}$

Conditional Expectation

- Given $(\Omega, \mathcal{A}, P)$, a random variable Y (with $E[|Y|] < \infty$) and $\mathcal{G} \subset \mathcal{A}$, there exists a $\mathcal{G}$-measurable function $E[Y|\mathcal{G}]$ such that

$$\int_G Y dP = \int_G E[Y|\mathcal{G}] dP, \;\; G \in \mathcal{G}$$

  Also, the function $E[Y|\mathcal{G}]$ is unique P-a.e.
- Proof: Define $\mu_Y(G) = \int_G Y dP$ for any $G \in \mathcal{G}$, then $\mu_Y \ll P$ and

$$E[Y|\mathcal{G}] = \frac{d\mu_Y}{dP}$$

- The function $E[Y|\mathcal{G}]$ is called the conditional expectation of Y given $\mathcal{G}$
  - "the expectation of Y knowing which events in $\mathcal{G}$ occurred"

- Note that with $\mathcal{G} = \{\emptyset, \Omega\}$ and $G = \Omega$, we get

$$E[Y] = \int_\Omega Y dP = \int_\Omega E[Y|\mathcal{G}] dP \Rightarrow E[Y|\{\emptyset, \Omega\}] = E[Y] \;\; P\text{-a.e.}$$

  (noting that the definition is verified also for $G = \emptyset$)
- If $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{A}$ then

$$E[E[Y|\mathcal{G}_1]|\mathcal{G}_2] = E[E[Y|\mathcal{G}_2]|\mathcal{G}_1] = E[Y|\mathcal{G}_1] \;\; P\text{-a.e.}$$

  so in particular, for any $\mathcal{G} \subset \mathcal{A}$,

$$E[E[Y|\mathcal{G}]] = E[E[Y|\mathcal{G}]|\{\emptyset, \Omega\}] = E[Y|\{\emptyset, \Omega\}] = E[Y] \;\; P\text{-a.e.}$$

- If Z is (already) $\mathcal{G}$-measurable, then

$$E[ZY|\mathcal{G}] = ZE[Y|\mathcal{G}] \;\; P\text{-a.e.}$$

Conditional Expectation vs. Probability

- The entity $E[Y|\mathcal{G}]$ is a function $g(\omega) = E[Y|\mathcal{G}](\omega)$
- If $(\Omega, \mathcal{A})$ is standard, then $P(E|\mathcal{G})$ is regular $\Rightarrow m(E) = P(E|\mathcal{G})(\omega)$ is a probability measure on $(\Omega, \mathcal{A})$ for fixed ω and $\mathcal{G}$. Furthermore, in this case

$$E[Y|\mathcal{G}] = \int Y(u)dm(u) = \int Y(u)dP(u|\mathcal{G})$$

- This interpretation for conditional expectation does not hold in general (for non-standard $(\Omega, \mathcal{A})$)

Projections and Atoms

Conditional expectation as a projection

- Given $(\Omega, \mathcal{A}, P)$ assume $\mathcal{G} \subset \mathcal{A}$ and let $\mathcal{M} = \{\,\mathcal{G}$-measurable functions$\,\}$
- For an $\mathcal{A}$-measurable Y, let $g(\omega) = E[Y|\mathcal{G}](\omega)$, then

$$E[(Y-g)^2] \leq E[(Y-g')^2] \text{ for all } g' \in \mathcal{M}$$

- If Y is already in $\mathcal{M}$, then $g(\omega) = Y(\omega)$ P-a.e.

Conditioning on a random variable

- For two random variables X and Y, $E[Y|X] = E[Y|\sigma(X)]$
- If $E|Y| < \infty$ then there is a Borel-measurable $f : \mathbb{R} \to \mathbb{R}$ such that $E[Y|X] = f(X(\omega))$ P-a.e.
- Thus $E[Y|X](\omega)$ is constant on the sets $\{\omega : X(\omega) = x\}$

Atoms

- $A \in \mathcal{A}$ is an atom of $\mathcal{A}$ if the only set in $\mathcal{A}$ which is a proper subset of A is $\emptyset$
- If there is a countable $\{A_i\}$ such that $\mathcal{A} = \sigma(\{A_i\})$ then $\mathcal{A}$ is separable
- $(\Omega, \mathcal{A})$ standard $\Rightarrow \mathcal{A}$ separable [more about "standard" later]
- $\mathcal{A}$ separable $\Rightarrow$ every $A \in \mathcal{A}$ is a union of atoms
- If f is $\mathcal{A}$-measurable, then f is constant on the atoms of $\mathcal{A}$
- If $\mathcal{A}$ is separable and $\mathcal{G} \subset \mathcal{A}$, then the atoms of $\mathcal{G}$ are bigger
- If G is an atom of $\mathcal{G} \subset \mathcal{A}$ and $P(G) > 0$, then

$$E[Y|\mathcal{G}](\omega) = \frac{1}{P(G)} \int_G Y dP, \quad \text{for } \omega \in G$$

"smoothing over atoms"

Mutually Singular Measures

- Given $(\Omega, \mathcal{A})$, two measures μ_1 and μ_2 are mutually singular, notation $\mu_1 \perp \mu_2$, if there is a set $E \in \mathcal{A}$ such that $\mu_1(E^c) = 0$ and $\mu_2(E) = 0$.
- Lebesgue decomposition: Given a σ-finite measure space $(\Omega, \mathcal{A}, \mu)$ and an additional σ-finite measure ν on $\mathcal{A}$, there exist measures ν_1 and ν_2 on $\mathcal{A}$ such that $\nu_1 \ll \mu$, $\nu_2 \perp \mu$ and $\nu = \nu_1 + \nu_2$. This representation is unique.

Continuous and Discrete Measures

- For a measure space $(\Omega, \mathcal{A}, \mu)$ such that $\{x\} \in \mathcal{A}$ for all $x \in \Omega$:
  - $x \in \Omega$ is an atom of μ if $\mu(\{x\}) > 0$
  - μ is continuous if it has no atoms
  - μ is discrete if there is a countable $K \subset \Omega$ such that $\mu(K^c) = 0$
- Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space and ν an additional σ-finite measure on $\mathcal{A}$. Assume that $\{x\} \in \mathcal{A}$ for all $x \in \Omega$. Then there exist measures ν_{ac}, ν_{sc} and ν_{d} such that
  - $\nu_{\text{ac}} \ll \mu$, $\nu_{\text{sc}} \perp \mu$ an $\nu_{\text{d}} \perp \mu$
  - ν_{sc} is continuous and ν_{d} is discrete
  - $\nu = \nu_{\text{ac}} + \nu_{\text{sc}} + \nu_{\text{d}}$, uniquely

Decomposition on the Real Line

- Let ν be a finite measure on $(\mathbb{R}, \mathcal{B})$, then ν can be decomposed uniquely as $\nu = \nu_{\text{ac}} + \nu_{\text{sc}} + \nu_{\text{d}}$ where
  - ν_{ac} is absolutely continuous w.r.t. Lebesgue measure
  - ν_{sc} is continuous and singular w.r.t Lebesgue measure
  - ν_{d} is discrete
- Furthermore, if F_ν is the distribution function of ν, then

$$\nu(\{x\}) = F_\nu(x) - \lim_{x' \to x^-} F_\nu(x')$$

  That is, if there are atoms, they are the points of discontinuity of F_ν