Section 2: Random variables

ARE 210

September 5, 2017

- Introduction (10 min)
- A few tips (10 min)
- Practice questions (30 min)

The section notes are available on the section Github at github.com/johnloeser/are210 in the "section2" folder.

1 Definitions

- A **probability space** is a triple $(\Omega, \mathbf{F}, P)$, where Ω is the **sample space**, $\mathbf{F}$ is a **σ-algebra** over Ω, and P is a **probability measure**. $(\Omega, \mathbf{F})$ is a **measurable space**.
- $\{A_i\}_{i=1}^N$ are **mutually independent** if $\forall\, I \subseteq \{1, \ldots, N\}$, $P(\cap_{i\in I} A_i) = \prod_{i\in I} P(A_i)$
- A **random variable** $X : (\Omega, \mathbf{F}) \to (E, \mathbf{E})$, where X is a measurable function, and $(\Omega, \mathbf{F})$ and $(E, \mathbf{E})$ are measurable spaces
  - Note: the mapping from $\mathbf{F}$ to $\mathbf{E}$ is induced by the mapping from Ω to E. The mapping from P to P_X is induced by the mapping from $\mathbf{F}$ to $\mathbf{E}$.
  - **stochastic process** $\equiv E =$ the space of functions mapping a set T to $\mathbf{R}^k$
  - **scalar random variable** $\equiv E = \mathbf{R}$
  - **discrete random variable** $\equiv E \cong \mathbf{N}$ (E countable)
- The **CDF** (cumulative distribution function) of a random variable $X : (\Omega, \mathbf{F}) \to (\mathbf{R}^k, \mathbf{B}^k)$ is $F_X(x) = P_X[X_1 \leq x_1, \ldots, X_k \leq x_k]$
  - $F : \mathbf{R^k} \to [0, 1]$ is a CDF if and only if 1) $\lim_{x\downarrow x_0} F(x) = F(x_0)$ (right continuity), 2) F is non-decreasing, 3) $\inf_{x\in\mathbf{R}^k} F(x) = 0$, $\sup_{x\in\mathbf{R}^k} F(x) = 1$
  - Assume X is absolutely continuous. Then $\exists$ a **density** $f_X(x)$ such that $F_X(x) \equiv \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_k} f_X(\widetilde{x}) d\widetilde{x}_k \ldots d\widetilde{x}_1$.
    - $\Rightarrow P_X[X \in A] = \int_A f_X(x)dx$
- $X : (\Omega, \mathbf{F}) \to (E, \mathbf{E})$ is a **continuous random variable** if it has no atoms (is diffuse)
  - $\{\omega\} \in \mathbf{E}$ (where $\omega \in E$) is an **atom** if $P_X(\{\omega\}) > 0$
  - Let D be the set of atoms. X is discrete if and only if $P_X(E \setminus D) = 0$

- X is continuous if and only if, for any countable set C, $P(X \in C) = 0$
- X is a **mixed random variable** if it is neither continuous nor discrete

- Random variables $X_1, \dots, X_n$ are **mutually independent** if $P(\cap_{i=1}^n X_i \in B_i) = \prod_{i=1}^n P(X_i \in B_i)$ for all measurable sets B_i for $i \in 1, \dots, n$.
  - Let $X_1, \dots, X_n$ be a vector of real valued random variables with CDF F and marginal CDFs F_i. $\{X_i\}_{i=1}^n$ are mutually independent if and only if $F(x_1, \dots, x_n) = \prod_{i=1}^n F_i(x_i)$.
- Let $E \subset \mathbf{R}^k$, $g : E \to \mathbf{R}^k$ one-to-one, continuously differentiable, $\left|\frac{\partial g(y)}{\partial y}\right| \neq 0$ in nbd of $g^{-1}(y)$ for $y \in \mathbf{R}^k$, $Y = g(X)$. Then, the density $f_Y(y) = f_X(g^{-1}(y)) \left|\frac{\partial g^{-1}(y)}{\partial y}\right|$
- Let F be a CDF, the **quantile function** is $Q(p) \equiv \inf\{x \in E : F(x) \geq p\}$.

2 Some helpful tips

- It can be helpful to think of a random variable X as a mapping just from one sample space Ω to another sample space E. However, this mapping induces a mapping from the σ-algebra $\mathbf{F}$ to the σ-algebra $\mathbf{E}$ ($A \in \mathbf{F} \Rightarrow X(A) \equiv \{X(\omega)|\omega \in A\} \in \mathbf{E}$) and from the probability measure P to the probability measure P_X ($\forall\ A \in \mathbf{F},\ P(A) = P_X(X(A))$).

3 Practice questions

1) **Probability of complement**: Show that $P(A^c) = 1 - P(A)$.

2) **Conditional Bayes' Theorem**: Show that $P(A|B, C) = \frac{P(A,B|C)}{P(B|C)}$

3) **Mixed random variables**: Let (X_1, X_2) be random variables with joint distribution F. How would you calculate the marginal distribution of X_1?

4) **Independence and densities**: Let X_1, X_2 be independent continuous random variables. Show that their joint density is equal to the product of their marginal densities, and assume for simplicity that each has a continuous marginal density.

5) **Distribution practice**: Let X_1, X_2 be standard normal random variables, each with density $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.
a) What is the joint distribution F of X_1 and X_2? What is the joint density f?

b) What is the distribution F_Y of $Y = X_1 + X_2$? What is the density f_Y?