ECE 5630: Information Theory for Data Transmission, Security and Machine Learning 01/28/20

Lecture 3: Random Variables & CDFs

Lecturer: Prof. Ziv Goldfeld

Scriber: Chris Schwarze, *Net ID*: cs977

Assistant Editor: Kia Khezeli

3.1 Random Variables

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we would often like to map elements of the sample set Ω to the real numbers. Such mappings corresponds to measurements performed with respect to our random experiment (as captured by the probability space). Formally, mappings from Ω to $\mathbb{R}^d$ need to be measurable functions, which is formulated in the definition of random variables, as given next.

Definition 3.1 (Random Variable) *A function $X : \Omega \to \mathbb{R}^d$ is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ if for any $B \in \mathcal{B}(\mathbb{R}^d)$, the preimage $X^{-1}(B) := \{\omega \in \Omega \ : \ X(\omega) \in B\}$ satisfies $X^{-1}(B) \in \mathcal{F}$.*

Remark 3.1 (Notation) *Random variables are typically denoted by capital letters, e.g., X. We often use $X \in B$ as a shorthand for $X^{-1}(B)$.*

Definition 3.2 (Induced Probability Law) *Let X be a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The probability measure on $\mathbb{R}^d$ induced by X, also called the* law *or the* distribution *of X, is*

$$\mathbb{P}_X(B) := \mathbb{P}\left(X^{-1}(B)\right) = \mathbb{P}(X \in B), \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Random variables are inherently tied to a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Indeed, the mere definition of a random variable (Definition 3.1) requires that $X^{-1}(B)$ is $\mathcal{F}$-measurable, for all $B \in \mathcal{B}(\mathbb{R}^d)$. Similarly, the probability law $\mathbb{P}_X$ on $\mathbb{R}^d$ is defined by X and the original probability measure $\mathbb{P}$ on Ω. When $(\Omega, \mathcal{F}, \mathbb{P})$ and the random variable X are given, one can define the image probability space (of X) without ambiguity.

Remark 3.2 (Image Probability Space) *Let $X : \Omega \to \mathbb{R}^d$ be a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, X induces a mapping from the original probability space to the probability space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_X)$.*

Example 3.1 (Rolling A Die) *Consider the experiment of rolling a six-sided fair die. This is modelled by the probability space $(\Omega, 2^\Omega, \mathbb{P})$, where $\Omega = \{1, \ldots, 6\}$ and $\mathbb{P}$ is the uniform measure. Consider a game in which each outcome of the roll corresponds to a monetary payout. We model the monetary payout by random variable X. More precisely, for each possible outcome $\omega \in \Omega$, the money gained (or lost) is $X(\omega) \in \mathbb{R}$. We can address different questions regarding the amount of money gained (e.g., "what is the probability of gaining money?") using the probability law of X. For example, the answer to the above question is $\mathbb{P}(X > 0) = \mathbb{P}\big(X \in (0, \infty)\big)$.*

3.2 Cumulative Distribution Functions

The cumulative distribution function (CDF) of a random variable is specified by its probability law. Interestingly, the converse is also true, that is, any CDF gives rise to a unique probability law. We start by defining CDFs.

Definition 3.3 (Cumulative Distribution Function) *Let $X : \Omega \to \mathbb{R}^d$ be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. The CDF $F_X : \mathbb{R}^d \to [0, 1]$ of X is defined as*

$$F_X(a_1, \ldots, a_d) := \mathbb{P}_X\left((-\infty, a_1] \times \cdots \times (-\infty, a_d]\right), \quad (a_1, \ldots, a_d) \in \mathbb{R}^d.$$

The following are some of the important properties of CDFs.

Proposition 3.1 (Properties of CDFs) *Let F_X be the CDF of a real-valued random variable $X : \Omega \to \mathbb{R}$. Then,*

1. $\lim_{t\to\infty} F_X(t) = 1$.
2. $\lim_{t\to-\infty} F_X(t) = 0$.
3. *F_X is monotonically non-decreasing and right continuous.*
4. $\mathbb{P}_X([a, b]) = F_X(b) - F_X(a)$ *for all* $a, b \in \mathbb{R}$.

From Proposition 3.1 and Carathéodory's extension theorem it follows that the CDF F_X uniquely defines $\mathbb{P}_X$ the probability measure induced by X.

Exercise 3.1 *Prove proposition 3.1.*

Definition 3.4 (Support) *Let X be a random variable with law $\mathbb{P}_X$. We denote the support of $\mathbb{P}_X$ by* $\mathrm{supp}(\mathbb{P}_X)$ *and define it as the smallest Borel set $B \in \mathcal{B}(\mathbb{R}^d)$ such that $\mathbb{P}_X(B) = 1$.*

Remark 3.3 (Null Set) *Define $\Omega \setminus \mathrm{supp}(\mathbb{P}_X) = \Omega \cap (\mathrm{supp}(\mathbb{P}_X))^c$. Then, for all $A \subset (\Omega \setminus \mathrm{supp}(\mathbb{P}_X))$, it holds that $\mathbb{P}(X \in A) = 0$.*

Definition 3.5 (Discrete Random Variables) *Let X be a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, we say X is a discrete random variable if and only if* $\mathrm{supp}(\mathbb{P}_X)$ *is countable.*

Definition 3.6 (Probability Mass Function) *If X is a discrete random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, then its probability mass function (PMF) $p_X : \mathrm{supp}(\mathbb{P}_X) \to [0, 1]$ is defined as*

$$p_X(x) := \mathbb{P}_X(\{x\}) = \lim_{\epsilon\to 0}(F_X(x) - F_X(x - \epsilon))$$

Proposition 3.2 *Recall that a PMF $p : \Omega \to [0, 1]$, gives rise to a probability space $(\Omega, \mathcal{F}, \mathbb{P}_p)$, where $\mathbb{P}_p(A) = \sum_{\omega\in A} p(\omega)$, for all $A \in \mathcal{F}$. Show that $\mathbb{P}_X(B \cap \mathrm{supp}(\mathbb{P}_X)) = \mathbb{P}_{p_X}(B \cap \mathrm{supp}(\mathbb{P}_X))$, holds for all $B \in \mathcal{B}(\mathbb{R}^d)$ (see Exercise 5(a) of the 1st homework sheet).*

Definition 3.7 (Continuous Random Variables) *Let X be a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, we say X is a continuous random variable if and only if* $\mathrm{supp}(\mathbb{P}_X)$ *is uncountable and F_X is differentiable on the interior of* $\mathrm{supp}(\mathbb{P}_X)$.

Definition 3.8 (Probability Density Function) *If X is a continuous random variable over $(\Omega, \mathcal{F}, \mathbb{P})$, then its probability density function (PDF) $f_X : \mathrm{supp}(\mathbb{P}_X) \to \mathbb{R}_{\geq 0}$ is defined as $f_X(t) := \frac{dF_X(t)}{dt}$.*

Proposition 3.3 *As in the case of discrete random variables, it holds that $\mathbb{P}_{f_X} = \mathbb{P}_X$ (see homework exercise 5(b)). Here P_f is the probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ induced by the PDF $f : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$, i.e., $\mathbb{P}_f(B) = \int_B f(x)\mathrm{d}x$, for all $B \in \mathcal{B}(\mathbb{R}^d)$.*