Section 1C. Elements of Probability

Statistics for Data Science

Victor M. Preciado, PhD MIT EECS
Dept of Electrical & Systems Engineering
University of Pennsylvania
preciado@seas.upenn.edu

Elements of probability

Basic elements:

- *Sample space* Ω: The set of all possible outcomes
- *Set of events* $\mathcal{F}$: A set of subsets of Ω
- *Probability measure*: A function $\Pr : \mathcal{F} \to \mathbb{R}$ satisfying
  1. For all $A \in \mathcal{F}$, $\Pr(A) \geq 0$
  2. $\Pr(\Omega) = 1$
  3. If A_1 and A_2 are *disjoint* events, then $\Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2)$

Example. Tossing a six-sided dice:

- The sample space is $\Omega = \{1, 2, \dots, 6\}$
- A possible choice for $\mathcal{F}$ is the set of all subsets of Ω
- The probability that the outcome $\omega \in \Omega$ is in a set $\mathcal{A} \subseteq \Omega$ is given by $\Pr(\omega \in \mathcal{A}) = |\mathcal{A}|/6$, where $|\mathcal{A}|$ is the cardinality of the set $\mathcal{A}$

Elements of Probability (cont.)

A few properties of probability measures:

- $\Pr(A \cap B) \leq \min(\Pr(A), \Pr(B))$
- $\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$
- If $A_1, \ldots, A_k$ are a *partition* of Ω, then $\sum_{i=1}^{k} \Pr(A_i) = 1$

Conditional probability and independence:

- The *conditional probability* of any event A given an event B is defined as,

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

  In plain words, $\Pr(A|B)$ represents the probability of event A after observing the occurrence of event B.

- Two events are called *independent* if and only if

$$\Pr(A \cap B) = \Pr(A)\Pr(B) \text{ or, equivalently, } \Pr(A|B) = \Pr(A)$$

Random Variables

A ***random variable*** (r.v.) is a function $X : \Omega \to \mathbb{R}$.

- For an outcome $\omega \in \Omega$, we denote r.v.'s using upper case letters $X(\omega)$ or simply X
- We denote the realization that the random variable may take using lower case x

Example: Toss three different (fair) coins at once

- The sample space is $\Omega = \{HHH, HHT, HTH, \ldots, TTT\}$
- Define the r.v. $H(\omega)$ as the number of heads in the random outcome ω