Chapter 1

Review of Probability

Beginning of Jan.10, 2022

Some preliminaries first:

- Throughout this course, we will use Ω to denote the **universal set**.
- A **probability law** on ω is a function $\mathbb{P} : \Omega \to [0,1]$ satisfying the following axioms:

  (1) (Nonnegativity) $\mathbb{P}(A) \geqslant 0$ for all $A \subset X$[1].

  (2) (Countable additivity) For $\{A_i\}_{i\geqslant 1}$ with $A_i \cap A_j = \varnothing$ whenever $i \neq j$, $\mathbb{P}(\bigcup_{i\geqslant 1} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

  (3) (Normalization) $\mathbb{P}(\Omega) = 1$.

- The following are direct consequences of the definition of a probability law:

  (1) If $A \subset B$ then $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.

  (2) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

  (3) (Union bound) $\mathbb{P}(A \cup B) \leqslant \mathbb{P}(A) + \mathbb{P}(B)$ and more generally $\mathbb{P}(\bigcup_{k=1}^{\infty} A_k) \leqslant \sum_{k=1}^{\infty} \mathbb{P}(A_k)$.

- Random variable definitions:

  (1) A **random variable** is a function $X : \Omega \to \mathbb{R}$ (or some different codomains). A **random vector** X is a function $X : \Omega \to \mathbb{R}^n$.

  (2) A **discrete random variable** is a random variable with finite or countable range.

  (3) A **probability density function** (PDF) is a function $f : \mathbb{R} \to [0, \infty)$ such that

  $$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1 \qquad \text{and } \int_a^b f(x)\,\mathrm{d}x \text{ exists for all } -\infty \leqslant a \leqslant b \leqslant \infty.$$

  (4) A random variable X is **continuous** if there exists a PDF f with

  $$\mathbb{P}(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x \qquad \text{for all } -\infty \leqslant a \leqslant b \leqslant \infty.$$

  If so we say f is the PDF of X.

[1] For technical reasons we avoid measure theories and assume all $A \subset X$ are measurable.

(5) Let X be a random variable. We define the **cumulative distribution function** (CDF) to be $F : \mathbb{R} \to [0,1]$ by

$$F(x) := \mathbb{P}(X \leqslant x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

- Examples of some distributions:

(1) Bernoulli: let $0 < p < 1$ and define $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$ and $\mathbb{P} \equiv 0$ otherwise. "Flip one coin. Count the number of heads."

(2) Binomial: let $n \in \mathbb{N}$ and $0 < p < 1$. For $k \in \{0, \ldots, n\}$, define $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ and define $\mathbb{P} \equiv 0$ otherwise. Can be thought of the sum of n independent Bernoulli with parameter p. "Flip n coins. Count the number of heads."

(3) Geometric: let $0 < p < 1$ and define $\mathbb{P}(X = k) = (1-p)^{k-1} p$ for $k \in \mathbb{N}$ and 0 otherwise. "Flip a coin until heads shows up. Count the number of flips."

(4) Normal / Gaussian with mean μ and variance σ^2: the PDF is given by

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

(5) Poisson with parameter $\lambda > 0$:

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad \text{for } k \in \mathbb{N}.$$

"Limit of binomial random variables subject to $\lim p_n = 0$ and $\lim np_n = \lambda$."

Definition: (1.17) Independent Sets

Let $\{A_i\}_{i \in I} \subset \Omega$ equipped with probability law Ω. We say $\{A_i\}$ are **independent** if, for all $S \subset I$ we have

$$\mathbb{P}(\bigcap_{i \in S} A_i) = \prod_{i \in S} \mathbb{P}(A_i).$$

Remark. This is *stronger* than pairwise independence, which only says $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ for $i \neq j$. An example can be found here.