2 Probability Theory

Definitions

- Sample space Ω
- Outcome (point or element) $\omega \in \Omega$
- Event $A \subseteq \Omega$
- σ-algebra $\mathcal{A}$
  1. $\varnothing \in \mathcal{A}$
  2. $A_1, A_2, \ldots, \in \mathcal{A} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$
  3. $A \in \mathcal{A} \implies \neg A \in \mathcal{A}$
- Probability Distribution $\mathbb{P}$
  1. $\mathbb{P}[A] \geq 0 \quad \forall A$
  2. $\mathbb{P}[\Omega] = 1$
  3. $\mathbb{P}\left[\bigsqcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]$
- Probability space $(\Omega, \mathcal{A}, \mathbb{P})$

Properties

- $\mathbb{P}[\varnothing] = 0$
- $B = \Omega \cap B = (A \cup \neg A) \cap B = (A \cap B) \cup (\neg A \cap B)$
- $\mathbb{P}[\neg A] = 1 - \mathbb{P}[A]$
- $\mathbb{P}[B] = \mathbb{P}[A \cap B] + \mathbb{P}[\neg A \cap B]$
- $\mathbb{P}[\Omega] = 1 \qquad \mathbb{P}[\varnothing] = 0$
- $\neg(\bigcup_n A_n) = \bigcap_n \neg A_n \quad \neg(\bigcap_n A_n) = \bigcup_n \neg A_n$ DeMorgan
- $\mathbb{P}[\bigcup_n A_n] = 1 - \mathbb{P}[\bigcap_n \neg A_n]$
- $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

  $\implies \mathbb{P}[A \cup B] \leq \mathbb{P}[A] + \mathbb{P}[B]$
- $\mathbb{P}[A \cup B] = \mathbb{P}[A \cap \neg B] + \mathbb{P}[\neg A \cap B] + \mathbb{P}[A \cap B]$
- $\mathbb{P}[A \cap \neg B] = \mathbb{P}[A] - \mathbb{P}[A \cap B]$

Continuity of Probabilities

- $A_1 \subset A_2 \subset \ldots \implies \lim_{n\to\infty} \mathbb{P}[A_n] = \mathbb{P}[A] \quad$ where $A = \bigcup_{i=1}^{\infty} A_i$
- $A_1 \supset A_2 \supset \ldots \implies \lim_{n\to\infty} \mathbb{P}[A_n] = \mathbb{P}[A] \quad$ where $A = \bigcap_{i=1}^{\infty} A_i$

Independence $\perp\!\!\!\perp$

$$A \perp\!\!\!\perp B \iff \mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$$

Conditional Probability

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} \qquad \mathbb{P}[B] > 0$$

Law of Total Probability

$$\mathbb{P}[B] = \sum_{i=1}^{n} \mathbb{P}[B|A_i]\,\mathbb{P}[A_i] \qquad \Omega = \bigsqcup_{i=1}^{n} A_i$$

Bayes' Theorem

$$\mathbb{P}[A_i \mid B] = \frac{\mathbb{P}[B \mid A_i]\,\mathbb{P}[A_i]}{\sum_{j=1}^{n} \mathbb{P}[B \mid A_j]\,\mathbb{P}[A_j]} \qquad \Omega = \bigsqcup_{i=1}^{n} A_i$$

Inclusion-Exclusion Principle

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{r=1}^{n} (-1)^{r-1} \sum_{i \leq i_1 < \cdots < i_r \leq n} \left|\bigcap_{j=1}^{r} A_{i_j}\right|$$

3 Random Variables

Random Variable (RV)

$$X : \Omega \to \mathbb{R}$$

Probability Mass Function (PMF)

$$f_X(x) = \mathbb{P}[X = x] = \mathbb{P}[\{\omega \in \Omega : X(\omega) = x\}]$$

Probability Density Function (PDF)

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f(x)\,dx$$

Cumulative Distribution Function (CDF)

$$F_X : \mathbb{R} \to [0,1] \qquad F_X(x) = \mathbb{P}[X \leq x]$$

1. Nondecreasing: $x_1 < x_2 \implies F(x_1) \leq F(x_2)$
2. Normalized: $\lim_{x\to-\infty} = 0$ and $\lim_{x\to\infty} = 1$
3. Right-Continuous: $\lim_{y \downarrow x} F(y) = F(x)$

$$\mathbb{P}[a \leq Y \leq b \mid X = x] = \int_a^b f_{Y|X}(y \mid x)dy \qquad a \leq b$$

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)}$$

Independence

1. $\mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[X \leq x]\,\mathbb{P}[Y \leq y]$
2. $f_{X,Y}(x,y) = f_X(x) f_Y(y)$