# Capitolo 7

# 1 Probabilità

## 1.1 Semantica della Probabilità

- **Definizione:** La probabilità misura il grado di certezza associato a un evento.
- È utilizzata in due principali interpretazioni:
  - *Frequenza relativa:* la probabilità è il limite della frequenza relativa di un evento in una serie di prove ripetute.
  - *Probabilità bayesiana:* rappresenta il grado di credenza soggettiva di un individuo rispetto a un evento.

## 1.2 Distribuzioni di Probabilità

### 1.2.1 Caso Discreto

- Rappresenta una probabilità associata a un insieme finito di risultati.
- Esempi comuni: lancio di un dado (valori da 1 a 6) o lancio di una moneta (testa o croce).
- Somma delle probabilità:

$$\sum_{x \in X} P(x) = 1.$$

### 1.2.2 Caso Continuo

- Utilizza funzioni di densità di probabilità (PDF) per descrivere eventi su intervalli continui.
- La probabilità di un intervallo è calcolata tramite integrali:

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx.$$

- Esempi: distribuzioni Normale, Uniforme, Esponenziale.

## 1.3 Definizione Assiomatica della Probabilità

- **Assiomi di Kolmogorov:**
  - $0 \leq P(A) \leq 1$
  - $P(S) = 1$, dove $S$ è l'intero spazio campionario.
  - Per eventi disgiunti $A$ e $B$,

$$P(A \cup B) = P(A) + P(B).$$

## 1.4 Proprietà delle Distribuzioni

- **Media (valore atteso):**

$$E[X] = \sum_x x \cdot P(x) \quad \text{(discreto)},$$

oppure

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx \quad \text{(continuo)}.$$

- **Varianza:**

$$\mathrm{Var}(X) = E[(X - E[X])^2].$$

- **Momenti:** descrivono la forma della distribuzione (media, varianza, asimmetria, curtosi).

## 1.5 Probabilità Condizionata

- Formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{con } P(B) > 0.$$

- Uso: aggiorna la probabilità di un evento dato che un altro evento è noto.

## 1.6 Regola della Catena

- Per più eventi:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot \cdots \cdot P(A_n|A_1, \ldots, A_{n-1}).$$

## 1.7 Teorema di Bayes

- Formula:

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}.$$

- Permette di aggiornare credenze iniziali $P(H)$ alla luce di nuove evidenze $E$.

## 1.8 Informazione ed Entropia

- **Entropia (H):** misura dell'incertezza:

$$H(X) = -\sum_x P(x) \log_2 P(x).$$

- **Information Gain:** riduzione dell'entropia dopo aver osservato un'evidenza:

$$IG = H(\text{prima}) - H(\text{dopo}).$$

# 2 Indipendenza

## 2.1 Indipendenza Condizionata

- Due variabili $X$ e $Y$ sono indipendenti condizionatamente su $Z$ se:
$$P(X, Y|Z) = P(X|Z) \cdot P(Y|Z).$$

## 2.2 Indipendenza Incondizionata

- Due eventi $A$ e $B$ sono indipendenti se:
$$P(A \cap B) = P(A) \cdot P(B).$$

# 3 Belief Network (Reti Bayesiane)

## 3.1 Definizione

- Una rete probabilistica diretta aciclica (DAG) che rappresenta relazioni tra variabili casuali.
- Costituita da:
  - **Nodi:** rappresentano variabili casuali.
  - **Archi:** relazioni probabilistiche condizionali.

## 3.2 Costruzione

1. Identifica le variabili (osservabili e latenti).
2. Definisci le dipendenze usando probabilità condizionali.

## 3.3 Inferenza Probabilistica

- **Inferenza Esatta:** calcolo diretto delle distribuzioni posteriori.
- **Inferenza Approssimata:**
  - Rejection Sampling.
  - Importance Sampling.
  - Particle Filtering.

# 4 Modelli Probabilistici Sequenziali

## 4.1 Catene di Markov

- **Definizione:** ogni stato dipende solo dal precedente:
$$P(S_{t+1}|S_t, S_{t-1}, \ldots) = P(S_{t+1}|S_t).$$
- **Applicazioni:** modellare processi temporali.

### 4.2 Modelli di Markov Nascosti (HMM)

- Estensione delle catene di Markov con stati latenti e osservazioni parziali.
- Usati per riconoscimento vocale, sequenze di DNA.

## 5 Simulazione Stocastica

### 5.1 Sampling

- **Forward Sampling:** genera campioni partendo dalle radici di una rete bayesiana.
- **Markov Chain Monte Carlo (MCMC):** campiona da distribuzioni complesse tramite catene di Markov.
  - **Gibbs Sampling:** aggiornamenti iterativi campionando una variabile alla volta.

## Conclusioni

Questi appunti riassumono i principi fondamentali della probabilità e dell'inferenza. Le tecniche descritte (reti bayesiane, HMM, simulazioni stocastiche) trovano applicazione in molte aree, dall'AI alla biologia computazionale.