Calcolo delle Probabilità

Leonardo Ganzaroli

Indice

Introduzione

Questi appunti del corso *Calcolo delle Probabilità* sono stati creati durante la laurea Triennale di informatica all'università "La Sapienza".

Prima di procedere rivedere la parte di insiemistica e combinatoria negli appunti di *Metodi Matematici per l'informatica*.

1 Probabilità classica

1.1 Definizioni

Definizione L'insieme ambiente è l'insieme contenente tutti i possibili esiti di un esperimento.

Definizione Un evento è un sottoinsieme dell'insieme ambiente, se contiene un solo elemento è detto elementare.

Definizione L'evento complementare di x (x^C) è l'insieme dei possibili esiti non nell'evento x.

Se lancio un dado a 6 facce ottengo l'insieme ambiente $\{1, 2, 3, 4, 5, 6\}$.
L'evento ottengo un numero pari è $\{2, 4, 6\}$, il complementare è $\{1, 3, 5\}$.

Definizione Dato un evento A. La funzione indicatrice è definita come:

$$I_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{altrimenti} \end{cases}$$

Definizione Dato l'insieme ambiente S e A suo evento. La probabilità che si verifichi A è data da:

$$P(A) = \frac{|A|}{|S|}$$

Teorema 1 (Passaggio al complemento)

$$P(A) = 1 - P(A^C)$$

1.2 Def. assiomatiche

Definizione Dato S insieme ambiente ed $E \subseteq S$. La probabilità di un evento è la funzione:

$$P : P(E) \to [0, 1]$$

Risulta:

- $P(S) = 1 = P(E_1) + P(E_2) + \dots$ con E_i evento e tutti gli eventi disgiunti
- $\forall\ E \subseteq S\ \ 0 \leq P(E) \leq 1$
- $P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$ se gli eventi sono tutti disgiunti
- $P(\emptyset) = 0$
- $A \subseteq B \Rightarrow P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Definizione Due eventi si dicono indipendenti sse:

$$P(A \cap B) = P(A) * P(B)$$

Definizione Dati $A, B \subseteq S$. La probabilità che avvenga A sapendo che è avvenuto B è:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{con } P(B) \neq 0$$

Definizione In alternativa si può usare la formula di Bayes:

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}$$

Definizione Dati $A \subseteq S$. Se:

$$S = \bigcup_{i=1}^{n} B_i \mid \forall\ i, j \in [1, n]\ (i \neq j \Rightarrow B_i \cap B_j = \emptyset)$$

Si definisce la probabilità totale di A come:

$$P(A) = \sum_{i=1}^{n} [P(A \mid B_i) * P(B_i)]$$

2 Variabili aleatorie

Definizione Uno spazio di probabilità è la tripla $(S, P(S), P)$ con:

- S insieme ambiente
- $P(S)$ insieme delle parti di S
- P funzione di probabilità

Definizione Una variabile aleatoria è una funzione $X : S \to \mathbb{R}$ che dà la probabilità di un evento, se assume al più un'infinità numerabile di valori è detta discreta.

Definizione Data una v.a. X. La densità discreta di X è:

$$p(x) = P(X = x)$$

Definizione Data una v.a. X. La funzione di distribuzione di X è:

$$F(x) = P(X \leq x)$$

Definizione Data X v.a. discreta e $x_1, \ldots, x_n$ i valori assumibili da essa. Il valore atteso di X è:

$$E[X] = \sum_{i=1}^{n} x_i * p_i$$

Definizione Un gioco è detto equo se il costo per giocare coincide con il valore atteso del premio.

Definizione Data una v.a. X. La varianza di X è:

$$Var(X) = E[(X - E[X])^2]$$

Definizione Date le v.a. X, Y. La covarianza tra X e Y è:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

Se $= 0$ si dicono decorrelate.

Teorema 2

$$Var(X_1 + \ldots + X_n) = \sum_{k=1}^{n} Var(X_k) \quad \textit{se le v.a. sono indipendenti.}$$

2.1 Tipi principali

- **Costante**

$$P(X = k) = 1,\ 0 \ \text{ altrimenti}$$

- **Bernoulli** $X \sim Ber(p)$

$$P(X = 1) = p,\ P(X = 0) = 1 - p$$

- **Binomiale** $X \sim B(n, p)$

  Con p prob. successo, k num. successi e n num. estrazioni.

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

- **Ipergeometrica** $X \sim H(n, m, r)$

  Con n num. elementi totali, m num. elementi di interesse, r num. estrazioni, k num. estrazioni elementi di interesse.

$$P(X = k) = \frac{\binom{m}{k}\binom{n-m}{r-k}}{\binom{n}{r}}$$

- **Geometrica** $X \sim G(p)$

  Con p prob. e $k - 1$ fallimenti prima del successo.

$$P_X(k) = (1 - p)^{k-1} * p$$

2.2 Spazi infiniti

Definizione Una v.a. è detta continua se:

$$\exists\, p : \mathbb{R} \to [0, +\infty) \mid P(X \in [a, b]) = \int_a^b p(t)\, dt \quad \text{con } p \text{ detta densità di probabilità}$$

Ovviamente il valore atteso diventa un integrale $\int_{-\infty}^{+\infty}$.

2.2.1 Tipi principali

- **Continua uniforme** $X \sim U(a, b)$

$$p(t) = \begin{cases} \frac{1}{b-a} & \text{se } t \in [a, b] \\ 0 & \text{altrimenti} \end{cases}$$

- **Poisson** $X \sim P(\lambda)$

  Con k num. di eventi per intervallo di tempo e λ num. medio di eventi per intervallo di tempo.

$$P_X(k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \forall k \in \mathbb{N}$$

  In alternativa si può ottenere come il limite della binomiale.

2.3 V.a. congiunte e condizionate

Definizione Date X, Y v.a. sullo stesso esperimento. La loro distribuzione congiunta è:

$$P_{XY}(x, y) = P(X = x, Y = y) = P(\{X = x\} \cap \{Y = y\})$$

Definizione Partendo da una distribuzione congiunta si ottengono le probabilità marginali di X, Y:

$$P_X(x) = P(X = x) = \sum_y P_{XY}(x, y)$$
$$P_Y(y) = P(Y = y) = \sum_x P_{XY}(x, y)$$

Partendo da questa tabella:

X	0	0	0	1	0	1	1	0
Y	1	0	0	2	3	1	1	0

Ricavo quella della distribuzione congiunta:

$Y \setminus X$	0	1
0	$\frac{3}{8}$	0
1	$\frac{1}{8}$	$\frac{1}{4}$
2	0	$\frac{1}{8}$
3	$\frac{1}{8}$	0

Risulta:

- $P_{XY}(1, 2) = \frac{1}{8}$
- $P_Y(1) = \frac{1}{4} + \frac{1}{8}$

Definizione 2 v.a. sono dette indipendenti se $P_{XY}(x, y) = P_X(x) * P_Y(y)$.

Definizione Data X v.a. e A evento. Il valore atteso di X condizionato da A è:

$$E[X \mid A] = \frac{E[X \cap A]}{P(A)}$$

3 Teoremi limite

Definizione Data X v.a. a valori solo positivi. La disuguaglianza di Markov è definita come:

$$\forall\ l > 0\ \ P(x \geq l) \leq \frac{E[X]}{l}$$

Definizione Data X v.a. a valori solo positivi. La disuguaglianza di Čebyšëv è definita come:

$$\forall\ l > 0\ \ P(|X - E[X]| \geq l) \leq \frac{Var[X]}{l^2}$$

Definizione Date $X_1, \ldots, X_n$ v.a. indipendenti tutte con stessa distribuzione e stesso valore atteso (μ). La legge debole dei grandi numeri è definita come:

$$\forall\ \epsilon > 0\ \lim_{n \to +\infty} P\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| \geq \epsilon\right) = 0$$

Quella forte invece:

$$P\left(\lim_{n \to +\infty} \frac{X_1 + \ldots + X_n}{n} = \mu\right) = 1$$

4 Catene di Markov (Introduzione)

Definizione Un processo stocastico è un insieme di v.a.

Definizione Gli stati sono il codominio delle v.a.

Definizione Una catena di Markov è un processo stocastico $\{X_t\}$ che rispetta le seguenti proprietà:

1. Per ogni istante di tempo t, coppia di stati i, j e sequenza di stati $k_0, \dots, k_{t-1}$ risulta:

$$P\{X_{t+1} = j \mid X_0 = k_0, \dots, X_{t-1} = k_{t-1}, X_t = i\} = P\{X_{t+1} = j \mid X_t = i\}$$

2. Per ogni coppia di stati i, j risulta:

$$\forall\ t \geq 0 \quad P\{X_{t+1} = j \mid X_t = i\} = P\{X_1 = j \mid X_0 = i\}$$

Estendendo il secondo punto si ottiene la probabilità di transizione in n passi:

$$p_{i,j}^{(n)} = P\{X_n = j \mid X_0 = i\} \quad \text{con } n \in \mathbb{N}$$

Definizione Una catena è detta omogenea se la transizione avvenuta in un certo tempo dipende soltanto dallo stato precedente.

Nel caso gli stati siano finiti si usa una matrice apposita per indicare le transizioni, la probabilità che ci si trovi in un certo stato dopo n passi è dato dalla moltiplicazione della matrice con se stessa per n volte.

Definizione Una catena ha distribuzione stazionaria se la sua distribuzione si mantiene costante nell'evolversi del tempo.

Definizione La matrice è detta ergodica se:

$$\exists n \in \mathbb{N} \mid \forall\ i, j \quad p_{i,j}^{(n)} > 0$$