VARIABILI ALEATORIE DISCRETE NOTEVOLI

La densità discreta contiene la descrizione probabilistica di una variabile aleatoria. Quindi due v. al. X e Y con la stessa densità sono probabilisticamente indistinguibili (o equidistribuite o identicamente distribuite), nel senso che $P(X \in B) = P(Y \in B)$ per ogni B sottoinsieme dell'alfabeto (Attenzione! Questo non significa $X = Y$!)
Assegnata una densità $p(\cdot)$ si può allora associare la famiglia delle v. al. X che hanno densità $p_X(\cdot) = p(\cdot)$.

CASO DI ALFABETO FINITO

Variabili aleatorie di Bernoulli

$X \sim Be(p)$ con $p \in [0, 1]$

se ha
- alfabeto $\mathcal{X} = \{0, 1\}$
- dens. discr. $p_X(1) = p$, $p_X(0) = 1 - p$

Inoltre, $E(X) = p$ e $Var(X) = p(1-p)$

Esempi

1. X v.al. che assume valore 1 se lanciando una moneta ottengo testa e assume valore 0 altrimenti. Si ha $X \sim Be(1/2)$.

2. Y v.al. che assume valore 1 se lanciando un dado ottengo un nr. pari e assume valore 0 altrimenti. Si ha $Y \sim Be(1/2)$.

X e Y sono probab. indistinguibili ma $X \neq Y$!

3. Sia $(\Omega, \mathcal{F}, P)$ uno sp. di prob. e $E \in \mathcal{F}$ un evento. Definiamo la v.al.

$$\mathbb{1}_E(\omega) = \begin{cases} 1 & \text{se } \omega \in E \\ 0 & \text{altrimenti} \end{cases}$$

Si ha $\mathbb{1}_E \sim Be(P(E))$.

Infatti otteniamo:

$p = P(\mathbb{1}_E = 1) = P(\{\omega \in \Omega : \omega \in E\}) = P(E)$.

SCHEMA DI BERNOULLI (O SCHEMA A PROVE INDIPENDENTI)

Contesto sperimentale:

- un certo numero $n \geq 1$ di prove identiche effettuate in sequenza
- ogni prova ha 2 esiti possibili codificati con 0 e 1
- il risultato di ciascuna prova non influenza il risultato delle altre.

Modello probabilistico:

- consideriamo $X_1, \dots, X_n$ v.al. identicamente distribuite
- $X_i \sim Be(p)$ (con $i = 1, \dots, n$) e p è la probabilità di ottenere 1
- gli eventi $\{X_1 = 1\}, \dots, \{X_n = 1\}$ sono tra loro indipendenti

Esempi

1. Una rete è composta da 150 terminali connessi ad un server. Controllo quali terminali sono pronti per trasmettere un lavoro. Per $i = 1, \dots, 150$ si ha

$$X_i = \begin{cases} 1 & \text{se l'i-esimo terminale è pronto per trasmettere un lavoro} \\ 0 & \text{altrimenti} \end{cases}$$

2. Controllo di qualità in una linea di produzione di chip. Ogni giorno ne vengono testati 1000.
Per $i = 1, \dots, 1000$ si ha

$$X_i = \begin{cases} 1 & \text{se l'i-esimo chip è difettoso} \\ 0 & \text{altrimenti} \end{cases}$$

3. Errori in una trasmissione digitale di 1'200'000 bit.
Per $i = 1, \dots, 1'200'000$ si ha

$$X_i = \begin{cases} 1 & \text{se l'i-esimo bit è stato trasmesso errato} \\ 0 & \text{altrimenti} \end{cases}$$

Variabili aleatorie binomiali

$X \sim \text{Bin}(n, p)$ con $n \in \mathbb{N} = \{1, 2, \dots\}$ e $p \in [0, 1]$

se ha
- alfabeto $\mathcal{X} = \{0, 1, \dots, n\}$
- dens. discr. $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \; k \in \mathcal{X}$

Inoltre, $E(X) = np$ e $Var(X) = np(1-p)$

La v.al. binomiale conta il nr. di successi (cioè di numeri '1') in uno schema di Bernoulli con n prove (indip.), dove la prob. di successo è p.

<u>Contesto</u>: le v. al. $X_i \sim Be(p)$ $(i=1,\dots,n)$ sono gli esiti delle n prove e la loro somma

$$X = \sum_{i=1}^{n} X_i \sim Bin(n,p)$$

è il nr. di successi ottenuti nelle n prove.
Cruciale: indip. degli eventi $\{X_1=1\},\dots,\{X_n=1\}$.

<u>Interpretazione densità</u>:

$$p_X(k) = P(\text{ottenere } k \text{ successi in } n \text{ prove})$$

$$= P\left(\begin{array}{l}\text{ottenere una stringa binaria di } n \text{ cifre} \\ \text{con } k \text{ cifre } 1\end{array}\right)$$

$$= \underbrace{\binom{n}{k}}_{\substack{\text{nr. stringhe di} \\ \text{lunghezza } n \\ \text{con } k \text{ cifre } 1}} \cdot \underbrace{\underbrace{p^k}_{\substack{\text{prob. di avere} \\ k \text{ cifre } 1}} \cdot \underbrace{(1-p)^{n-k}}_{\substack{\text{prob. di avere} \\ n-k \text{ cifre } 0}}}_{\text{equidistribuzione + indipendenza}}$$

Per esempio, supponiamo di fissare $n=5$, $k=1$ e di voler calcolare la probabilità della stringa '00010'. Se $X_i \sim Be(p)$, si ottiene

$$
\begin{aligned}
P(00010) &= P(X_1 = X_2 = X_3 = 0,\ X_4 = 1,\ X_5 = 0) \\
&\xrightarrow{\text{indip.}} = P(X_1=0)P(X_2=0)P(X_3=0)P(X_4=1)P(X_5=0) \\
&\xrightarrow{\text{equidistr.}} = (1-p)^4\, p
\end{aligned}
$$

Analogamente, possiamo calcolare

$$
\begin{aligned}
P(01000) &= P(X_1=0)P(X_2=1)P(X_3=0)P(X_4=0)P(X_5=0) \\
&= (1-p)^4\, p
\end{aligned}
$$

Per indipendenza ed equidistribuzione, tutte le stringhe di lunghezza 5 con una sola cifra 1 sono equiprobabili.

$\boxed{\text{ESERCIZIO}}$

È più facile ottenere almeno un 6 lanciando 4 volte un dado oppure ottenere almeno un doppio 6 lanciando 24 volte una coppia di dadi?

<u>Soluzione</u>. <u>Lancio singolo</u>. Costruiamo uno schema di Bernoulli. Per ogni $i = 1, \dots, 4$ definiamo le v.al.

$$X_i = \begin{cases} 1 & \text{se all'i-esimo lancio ottengo 6} \\ 0 & \text{altrimenti} \end{cases}$$

Si ha $X_i \sim Be(1/6)$. Inoltre gli eventi $\{X_1 = 1\}, \dots, \{X_4 = 1\}$ sono indipendenti. Allora il nr. di punteggi 6 ottenuti nei 4 lanci è $X = \sum_{i=1}^{4} X_i \sim Bin(4, 1/6)$.

Calcoliamo

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (5/6)^4 \approx 0.518.$$

<u>Lancio doppio</u>. Costruiamo uno schema di Bernoulli. Per $i = 1, \dots, 24$ definiamo le v.al.

$$Y_i = \begin{cases} 1 & \text{se all'i-esimo lancio ottengo una coppia di 6} \\ 0 & \text{altrimenti} \end{cases}$$

Si ha $Y_i \sim Be(1/36)$. Inoltre gli eventi $\{Y_1 = 1\}, \dots, \{Y_{24} = 1\}$ sono indipendenti. Allora il nr. di volte che ottengo un doppio 6 nei 24 lanci è $Y = \sum_{i=1}^{24} Y_i \sim Bin(24, 1/36)$.

Calcoliamo

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \left(35/36\right)^{24} \approx 0.49.$$

Quindi è più facile ottenere almeno un 6 lanciando un dado.