Lezione 01/11/2022

Variabili aleatorie discrete

Definizione. *Si definisce* funzione di ripartizione *di una variabile aleatoria X una funzione $F : R \to [0, +\infty[$ tale che $F(k) = P(X \leq k)$ per ogni $k \in R$.*

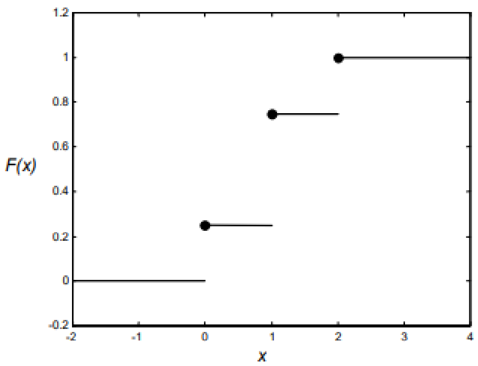
Esempio. *Considero il lancio di due monete e la v.a. X che conta il numero di teste. Calcolo i valori assunti dalla funzione di ripartizione:*

$$F(0) = P(X \leq 0) = P(X = 0) = \frac{1}{4} = 0.25;$$

$$F(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{4} + \frac{1}{2} = 0.74;$$

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1.$$

La funzione di ripartizione la rappresento con un grafico come quello seguente:



In generale se $X(\Omega) = \{x_1, \ldots, x_n\}$, allora

$$P(X \leq k) = \sum_{x_i \leq k} P(X = x_i).$$

Osservazione: Per la proprietà della probabilità dell'evento contrario:

$$P(X > a) = 1 - P(X \leq a);$$

$$P(X < a) = 1 - P(X \geq a);$$

$$P(X \geq a) = 1 - P(X < a);$$

$P(X \leq a) = P(X > a)$.

Le proprietà precedenti si usano negli esercizi quando vogliamo semplificare i calcoli. Il seguente è un esempio.

Esempio. *Sia X la v.a. che rappresenta la somma dei risultati di due lanci, allora $X(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ e $P(X \geq 3) = P(X = 3) + \ldots + P(X = 12)$, ma possiamo calcolare questa probabilità più facilmante perchè $P(X \geq 3) = 1 - P(X < 3) = 1 - P(X = 2)$.*

Variabile di Bernoulli

Definizione. *Si dice esperimento (o prova) di Bernoulli un esperimento aleatorio che può avere solo due possibili esiti. Conveniamo di chiamare* successo *l'esito che interessa e* insuccesso *l'altro esito possibile. La probabilità p di successo in un esperimento di Bernoulli si dice parametro dell'esperimento.*

Esempio.

- *Il lancio di una moneta (esiti: testa e croce, $p = \frac{1}{2}$ se la moneta non è truccata e successo=testa);*
- *Il lacio di un dado (esiti: numeri pari e numeri dispari, se successo=numero pari) allora $p = \dfrac{1}{2}$;*
- *Il lacio di un dado (esiti: numeri ≤ 2 e numeri > 2, se successo=numero ≤ 2 allora $p = \dfrac{1}{3}$);*
- *Estrazione di una pallina da un'urna con una pallina gialla e 3 palline verdi (esiti: pallina gialla o pallina non gialla, se successo=pallina gialla allora $p = \dfrac{1}{4}$);*
- *È noto che mediamente il 5% dei pezzi prodotti in una giornata da un'aziendahanno dei difetti. La scelta a caso di un pezzo tra quelli prodotti e la verifica se sia difettoso si possono assimilare a una prova di Bernoulli, dove si considera come* successo *l'evento che consiste nell'aver trovato un pezzo difettoso. Il parametro di questa prova di Bernoulli è $p = \frac{5}{100} = \frac{1}{20}$.*

Definizione. *Una* variabile di Bernoulli *è una variabile che descrive un espermento di Bernoulli. Si rappresenta così:*

$$X = \begin{cases} 1 \ \textit{(successo)} & \text{con probabilità } p \\ 0 \ \textit{(insuccesso)} & \text{con probabilità } 1 - p \end{cases}$$

Scriviamo $X \sim Ber(p)$ per indicare che X è una variabile di Bernoulli di parametro p.

Variabile Binomiale

Definizione (Processo di Bernoulli)**.** *Si chiama processo di Bernoulli l'esperimento aleatorio consistente nella ripetizione di n prove di Bernoulli identiche e indipendenti. Per esempio, sono processi di Bernoulli il lancio ripetuto per n volte di una moneta, oppure l'estrazione con reinserimento, per n volte successive, di una pallina da un'urna che contiene palline di due colori.*

Esempio. *Considerando i punti 1-4 dell'esempio precedente, i seguenti sono esempi di processi di Bernoulli:*

1. *Il lancio di una moneta per 5 volte (esiti di un singolo lancio: testa e croce, $p = \frac{1}{2}$ se la moneta non è truccata);*
2. *Il lacio di un dado per 4 volte (esiti di un singono lancio: numeri pari e numeri dispari, se successo=numeri pari allora $p = \dfrac{1}{2}$);*
3. *Il lacio di un dado per 20 volte (esiti di un singolo lancio: numeri ≤ 2 e numeri > 2, se successo=numeri ≤ 2 allora $p = \dfrac{1}{3}$);*
4. *100 estrazioni con rimpiazzo di una pallina da un'urna con una pallina gialla e 3 palline verdi (esiti di un'estrazione: pallina gialla o pallina non gialla, se successo=pallina gialla, allora $p = \dfrac{1}{4}$).*

Definizione (Variabile binomiale)**.** *Consideriamo un processo di Bernoulli costituito da n prove di parametro p. La variabile aleatoria X che conta il numero complessivo di successi ottenuti nelle n prove si dice binomiale di parametri n e p.*

Scriviamo $X \sim Bin(n, p)$ per indicare che X è una variabile aleatoria BINOMILALE di parametri n e p.

Esempio. *Consideriamo i processi di Bernoulli del precedente esercizio, allora i seguenti sono esempi di variabili binomiali:*

1. *$X =$numero di teste. $X \sim Bin(5, \dfrac{1}{2})$;*
2. *$X =$numero di risultati pari. $X \sim Bin(4, \dfrac{1}{2})$;*

3. X *=numero di risultati* ≤ 2. $X \sim Bin(20, \frac{1}{3})$*;*

4. X *= numero di palline gialle.* $X \sim Bin(100, \frac{1}{4})$.

Il parametro n è un numero intero positivo, mentre il parametro p è un numero reale con $0 \leq p \leq 1$. Qual è la distribuzione di probabilità di una variabile aleatoria binomiale di parametri n e p?

Teorema. *Sia X una variabile aleatoria binomiale di parametri n e p. La distribuzione di probabilità di X è data dalla formula:*

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \ldots, n$$

Esercizio 1. *Lancio 6 volte una moneta tale che la probabilità che esce testa è $\frac{1}{3}$. Voglio calcolare la probabilità che nei 6 lanci escano esattamente due teste.*

Svolgimento. *Considero la variabile aleatoria X = numero di teste. Dato che il lancio della moneta per 6 volte è un processo di Bernoulli di parametri $n = 6$ e $p = \frac{1}{3}$, dove il successo è l'evento* esce testa*, dobbiamo calcolare $P(X = 2)$ tale che $X \sim Bin(6, \frac{1}{3})$.*

$P(X = 2) = P(\{110000\} \cup \{000011\} \cup \{100001\} \cup \ldots)$*(perchè $P(X = 2)$ è la probabilità di ottenere 2 successi e 4 insuccessi)=*$P(\{110000\}) + P(\{000011\}) + P(\{100001\}) + \ldots$ *(perchè gli eventi di cui calcolo la probabilità sono incompatibili tra loro, ad esempio* 110000 *e* 000011 *non posso verificarsi contemporaneamente) (***).*

Calcolo $P(110000)$*. Osservo che ogni esito della sequenza è un valore assunto dalla variabile di Bernoulli $X_i \sim Ber(\frac{1}{3})$:*

$$X_i = \begin{cases} 1 & \textit{se esce testa nel lancio } i \\ 0 & \textit{se esce croce nel lancio } i \end{cases}.$$

Quindi $P(110000)$ *è uguale a* $P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 0, X_5 = 0, X_6 = 0)$.

$P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 0, X_5 = 0, X_6 = 0) = P(X_1 = 1) \cdot P(X_2 = 1) \cdot P(X_3 = 0) \cdot P(X_4 = 0) \cdot P(X_5 = 0) \cdot P(X_6 = 0)$ *(dato che gli eventi sono indipendenti uso la proprietà $P(A \cap B) = P(A) \cdot P(B)$)=*
$=\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{1}{3}^2 \cdot \frac{2}{3}^4$.

Osservo che la probabilità che esca una qualsiasi sequenza con 2 uni e 4 zeri è sempre $P(110000)$, cioè $\frac{1}{3}^2 \cdot \frac{2}{3}^4$. In altre parole, $P(110000) = P(000011) = P(100001) = \ldots$ (Riprovando a calcolare la probabilità di qualsiasi sequenza di 2 uni e 4 zeri si ripete lo stesso ragionamento utilizzato per calcolare $P(110000)$, che è basato sul fatto che gli eventi sono indipendenti).

Di conseguenza, la somma $(* * *)$ si ottiene moltiplicando $\frac{1}{3}^2 \cdot \frac{2}{3}^4$ per il numero di sequenze di due uni e 4 zeri.

Quante sono le tutte le sequenze di due uni e 4 zeri?

Le sequenze di due uni e 4 zeri sono le combinazioni semplici di 6 oggetti e di classe 2, quindi il loro numero è $\binom{6}{2}$ (infatti scegliere la sequenza 110000 è equivalente a considerare il gruppo $\{X_1, X_2\}$ segliendo tra tutte le variabili (oggetti) $X_1, \ldots, X_6$; la sequenza 000011 è equivalente a considerare il gruppo $\{X_5, X_6\}$ segliendo tra tutte le variabili (oggetti) $X_1, \ldots, X_6$, ecc).

Infine,

$$P(X = 2) = \binom{6}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^4.$$