Dato uno spazio campionario S, una *variabile aleatoria* (o casuale) è una funzione X definita su S che associa ad ogni evento elementare un numero reale.

In alcuni casi gli eventi elementari sono già numeri reali, ad esempio i numeri da 1 a 6 nel lancio di un dado, e allora sono essi stessi valori di una variabile aleatoria. In altri casi è necessaria un'opportuna codifica. Ad esempio, nel lancio della moneta dove lo spazio campionario è $S = \{T, C\}$, possiamo definire la variabile aleatoria X tale che $X(T) = 0$ e $X(C) = 1$. Oppure nel lancio di due monete possiamo definire una variabile aleatoria che conta le croci, percui $X(TT) = 0,\ X(TC) = 1,\ X(CC) = 1,\ X(CT) = 1$.

Una variabile aleatoria si dice *variabile aleatoria finita*, se essa assume soltanto un numero finito di valori distinti.

Se X è una variabile aleatoria finita, useremo le seguenti notazioni:

$$\begin{array}{ll} X = a & \text{indica l'evento "X assume il valore a"} \\ X \leq a & \text{indica l'evento "X assume valori } \leq \text{a"} \\ X \geq a & \text{indica l'evento "X assume valori } \geq \text{a"} \\ a < X < b & \text{indica l'evento "X assume valori compresi in (a, b)"} \end{array}$$

Ad esempio, se l'esperimento consiste nel lanciare due monete eque, e se X è la variabile definita come numero di croci, allora $P(X \leq 1) = 3/4$.

Se X è una variabile aleatoria finita che assume soltanto i valori distinti $x_1 < x_2, < \cdots < x_n$, la funzione

$$f(x_i) = P(X = x_i), \quad i = 1, n$$

si chiama *distribuzione di probabilità* della variabile aleatoria X. Si chiama *funzione di ripartizione* di X la funzione

$$F(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

Una distribuzione di probabilità si rappresenta con un diagramma a barre (a bastoncini).

Una funzione di ripartizione è definita su tutto $\mathbb{R}$, è monotona crescente in senso debole, ha valori compresi fra 0 e 1 ed il suo grafico è una funzione a gradino.

Esempio

L'esperimento consista nel lancio di due monete distinguibili non truccate e sia X la variabile aleatoria definita come il numero di volte che esce C (croce) in ciascun lancio. Sappiamo che $P(TT) = P(TC) = P(CT) = P(CC) = 1/4$. Quindi $P(X = 0) = 1/4,\ P(X = 1) = 1/2,\ P(X = 2) = 1/4$. Possiamo rappresentare la distribuzione di probabilità con la Tabella 1 o con il diagramma a barre Figura 1.

Tabella 1: Lancio di due monete

x_i	0	1	2
$f(x_i)$	1/4	1/2	1/4

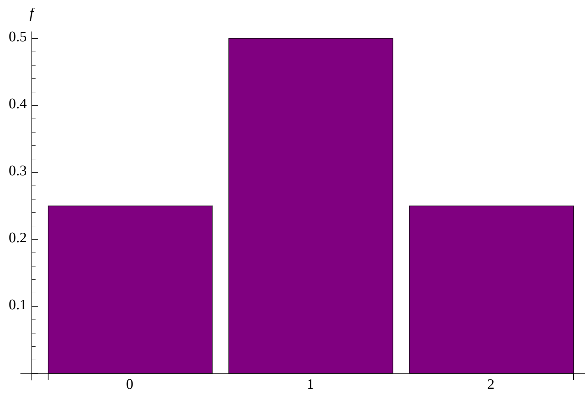

Figura 1: Lancio di due monete

Esempio

L'esperimento sia il lancio di un dado non truccato. Sia X il risultato sul dado. Allora X assume tutti i valori interi compresi fra 1 e 6 e la distribuzione di probabiltà è $P(X = i) = 1/6$ per ogni $i = 1, 6$. La funzione di ripartizione è F tale

$$F(x) = \begin{cases} 0 & se\ x < 1 \\ 1/6, & se\ 1 \leq x < 2 \\ 2/6, & se\ 2 \leq x < 3 \\ 3/6, & se\ 3 \leq x < 4 \\ 4/6, & se\ 4 \leq x < 5 \\ 5/6, & se\ 5 \leq x < 6 \\ 6/6, & se\ 6 \leq x \end{cases}$$

0.1 Valore atteso e varianza di una variabile aleatoria

Definizione 0.1.1 *Sia X una variabile aleatoria finita che può assumere solo i valori $x_1, \ldots, x_N$. Il* valore atteso *di X è il numero*

$$E[X] = \sum_{i=1}^{N} x_i P(X = x_i).$$

Esempi

1. Se la variabile aleatoria X è il numero di teste nel lancio di una moneta non truccata, allora il valore atteso di X è

$$E[X] = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}.$$

2. Se la variabile aleatoria X può assumere solo i valori 1 e 0 e $P(X = 1) = \frac{1}{3}$, allora il valore atteso di X è

$$E[X] = 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}.$$

3. Sia X il punteggio che si ottiene lanciando un dado non truccato, allora il valore atteso di X è

$$E[X] = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = 3.5.$$

   Si osservi che $E[X]$ non è necessariamente un valore di X.

Sia X una variabile aleatoria discreta che può assumere solo i valori $x_1, \ldots, x_N$. La *varianza* di X è il numero

$$Var[X] = \sum_{i=1}^{N} (x_i - E[X])^2 P(X = x_i),$$

e lo *scarto quadratico medio* è

$$\sigma = \sqrt{Var[X]}.$$

0.2 Variabili aleatorie bernoulliane e binomiali

Consideriamo un esperimento E i cui esiti sono due eventi incompatibili A e B. Si pensi al lancio di una moneta che ha solo i due esiti possibili Testa e Croce, oppure all'estrazione di un individuo da una popolazione i cui elementi sono raggruppati in due sole categorie, ad esempio Maschio e Femmina, oppure Infetto e Non infetto. Associamo convenzionalmente il valore 1 ad A, detto anche *successo* e il valore 0 a B detto anche *insuccesso*. Definiamo quindi la variabile aleatoria X che vale 1 in caso di successo e 0 in caso di insuccesso. Supponiamo inoltre che

$$P(A) = p \text{ e } P(B) = 1 - p.$$

Una variabile di questo tipo si chiama *variabile di Bernoulli o bernoulliana*.

Supponiamo ora di realizzare n ripetizioni indipendenti (dette anche prove bernoulliane) dello stesso esperimento E e indichiamo ora con X la variabile aleatoria che conta il numero totale di successi nelle n prove. Tale variabile aleatoria si chiama *variabile binomiale di parametri* (n, p). La distribuzione di probabilità associata ad una variabile binomiale di parametri (n, p) è data da

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i} = \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}, \ i = 0, n.$$

Infatti, fissata una precisa sequenza di i successi e $n-i$ insuccessi, la probabilità che si verifichi esattamente questa sequenza è $p^i(1-p)^{n-i}$ per l'indipendenza delle ripetizioni. D'altra parte si possono verificare i successi e $n-i$ insuccessi in $\binom{n}{i}$ modi distinti, che corrispondono ai modi distinti di scegliere i delle n ripetizioni cui attribuire il successo.

Risulta per la Formula del binomio di Newton

$$\sum_{i=0}^{n} P(X = i) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = (p + 1 - p)^n = 1.$$

Nelle Figure 2 e 3 potete osservare i grafici delle distribuzioni di probabilità di due variabili binomiali, la prima di parametri (40, 0.2) e l'altra di parametri (40, 0.5). Notate come la seconda sia simmetrica rispetto a $i = 20$ mentre la prima no.

Esempio

1. Una ditta produce lampadine che sono difettose con probabilità 0.01, indipendentemente l'una dall'altra. Queste lampadine vengono vendute

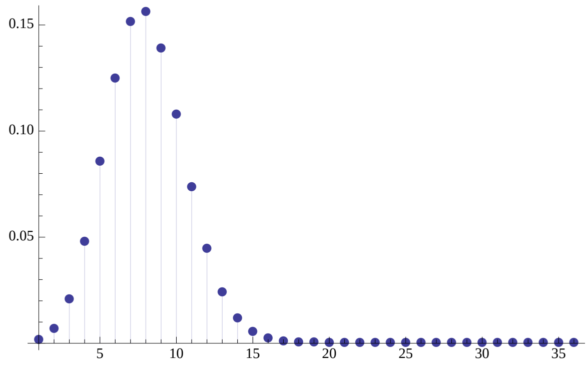

Figura 2: La distribuzione di una variabile binomiale di parametri (40, 0.2)

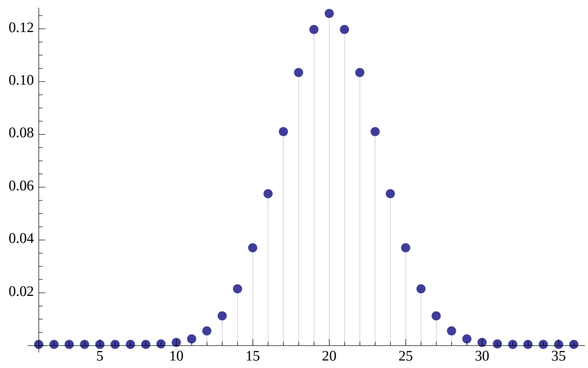

Figura 3: La distribuzione di una variabile binomiale di parametri (40, 0.5)

in scatole di 10 pezzi, con la garanzia di rimborso se nella scatola c'è più di una lampadina difettosa. Qual è la percentuale di confezioni che vengono rimborsate, assumendo che tutti i clenti usufruiscano della garanzia?

Svolgimento Sia X il numero di pezzi difettosi in una scatola di 10 pezzi. X è una variabile aleatoria binomiale di parametri (10, 0.01). Notate che in questo caso il successo corrisponde a trovare una lampadina difettosa nella confezione. La probabilità che una scatola contenga una lampadina difettosa e quindi venga restituita è pari a

$$\begin{aligned} P(X > 1) &= 1 - P(X = 0) - P(X = 1) = \\ &= 1 - \binom{10}{0} \cdot 0.01^0 \cdot 0.99^{10} - \binom{10}{1} \cdot 0.01^1 \cdot 0.99^9 \approx 0,0043. \end{aligned}$$

2. Calcolare la probabilità di ottenere 3 volte testa lanciando 7 volte una moneta non truccata.

   Svolgimento Il numero delle prove è 7 e il numero di successi è 3. La probabilità di successo è

   $$P(X = 3) = \binom{7}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{7-3} = 0.273437.$$

3. Calcolare la probabilità che, effettuando quattro estrazioni con reimbussolamento da un'urna contenente 20 palline bianche e 30 nere, venga estratta per esattamente tre volte una pallina bianca.

   Svolgimento La probabilità di successo cioè di estrarre una pallina bianca è 2/5. Il numero delle prove è 4. Il numero dei successi è 3.

   $$P(X = 3) = \binom{4}{3} \left(\frac{2}{5}\right)^3 \left(\frac{3}{5}\right)^{4-3} = 0.1536.$$

4. Calcolare la probabilità che, effettuando 6 estrazioni con reimbussolamento da un'urna contenente 20 palline bianche e 30 nere, venga estratta per almeno tre volte una pallina bianca.

   Svolgimento La probabilità di successo cioè di estrarre una pallina bianca è 2/5. Il numero delle prove è 6. L probabilità richiesta è uguale alla somma $P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) = 1 - P(X = 2) - P(X = 1) - P(X = 0) = 0.45568$.

0.2.1 Valore atteso e varianza di una variabile binomiale

Il valore atteso di una variabile binomiale $X = B(n, p)$ di parametri (n, p) è

$$E[X] = np$$

e la varianza è

$$Var[X] = np(1 - p).$$

Verifichiamo le formule precedenti nel caso $n = 1$:

$$E[B(1,p)] = 1 \cdot p + 0 \cdot (1-p) = p$$

$$Var[B(1,p)] = (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p) = (1-p)(p - p^2 + p^2) = p(1-p).$$