Calcolo delle Probabilità e Statistica Matematica – A.A. 2022/23

CAPITOLO 1 – Analisi combinatoria

1.1 Introduzione

1.2 Il principio fondamentale del calcolo combinatorio

1.3 Permutazioni

1.4 Disposizioni e combinazioni

N.B. La presente dispensa non sostituisce i libri di testo e non va resa accessibile sul web.

Gli argomenti indicati con ($\star\star$) vanno intesi come facoltativi.

1.1 Introduzione

Problema. Un sistema di comunicazione consiste di n antenne allineate. Ogni antenna può essere funzionante oppure difettosa. Quante sono le possibili configurazioni?

Avendo 2 casi per ogni antenna, il numero di possibili configurazioni è

$$2 \times 2 \times \cdots \times 2 = 2^n$$

Ad esempio, se $n = 4$ vi sono $2^4 = 16$ possibili configurazioni:

$$\begin{array}{cccc}
0\,0\,0\,0 & 0\,1\,0\,0 & 1\,0\,0\,0 & 1\,1\,0\,0 \\
0\,0\,0\,1 & 0\,1\,0\,1 & 1\,0\,0\,1 & 1\,1\,0\,1 \\
0\,0\,1\,0 & 0\,1\,1\,0 & 1\,0\,1\,0 & 1\,1\,1\,0 \\
0\,0\,1\,1 & 0\,1\,1\,1 & 1\,0\,1\,1 & 1\,1\,1\,1
\end{array}$$

(1 = antenna funzionante)

(0 = antenna **non** funzionante)

Supponiamo che il sistema sia funzionante se non vi sono 2 antenne difettose consecutive. Sapendo che esattamente m delle n antenne sono difettose, qual è la probabilità che il sistema sia funzionante?

Ad esempio, se $n = 4$ e $m = 2$ le possibili configurazioni sono 6:

$$
\begin{array}{lccccl}
 & \mathbf{0} & \mathbf{0} & 1 & 1: & \text{sistema } \textbf{non} \text{ funzionante} \\
 & 0 & 1 & 0 & 1: & \text{sistema funzionante} \\
(1 = \text{antenna funzionante}) & 0 & 1 & 1 & 0: & \text{sistema funzionante} \\
(0 = \text{antenna difettosa}) & 1 & \mathbf{0} & \mathbf{0} & 1: & \text{sistema } \textbf{non} \text{ funzionante} \\
 & 1 & 0 & 1 & 0: & \text{sistema funzionante} \\
 & 1 & 1 & \mathbf{0} & \mathbf{0}: & \text{sistema } \textbf{non} \text{ funzionante}
\end{array}
$$

Possiamo affermare che la probabilità che il sistema sia funzionante è la seguente?

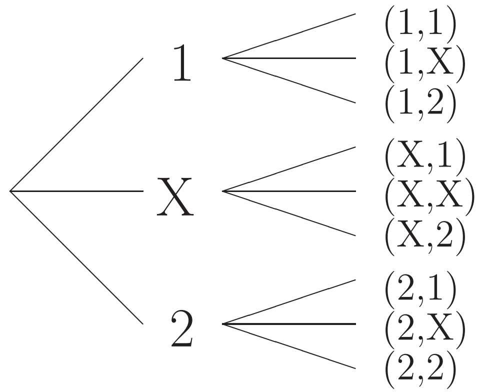
$$\text{prob.} = \frac{3}{6} = \frac{1}{2} = 0{,}5 \qquad (3 \text{ casi su } 6)$$

In molti problemi il calcolo delle probabilità si effettua semplicemente calcolando il numero di modi in cui avviene un dato evento; sarà questo l'argomento dell'*analisi combinatoria*.

1.2 Il principio fondamentale del calcolo combinatorio

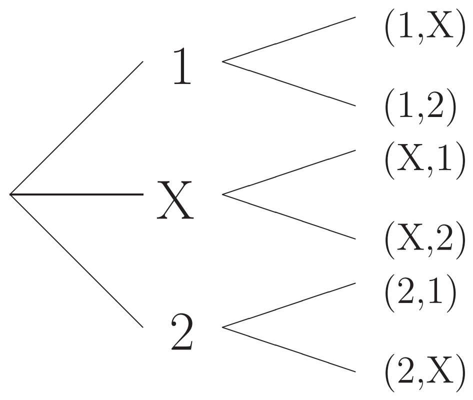
Esempio. Un giocatore scommette su 2 partite di calcio, con esiti 1, X, 2. In quanti modi può scegliere come scommettere?

Soluzione. $3 \times 3 = 9$



(i, j) : i è la scommessa sulla prima partita, j è la scommessa sulla seconda partita; è possibile effettuare qualsiasi scelta per entrambe le partite.

Esempio. Due giocatori (avversari) scommettono su una partita di calcio, con esiti 1, X, 2. Se ognuno può effettuare una scelta, in quanti modi si realizzano le scelte?
Soluzione. $3 \times 2 = 6$



(i, j) : i è la scommessa del primo giocatore, j è la scommessa del secondo giocatore; non è possibile effettuare la stessa scelta dai due giocatori.

Principio fondamentale del calcolo combinatorio. Si realizzino 2 esperimenti. Si supponga che il primo esperimento abbia m esiti possibili, e che per ognuno di questi il secondo esperimento abbia n esiti possibili. Se sequenze distinte di esiti dei due esperimenti producono esiti finali distinti, allora vi sono in tutto mn esiti possibili.

Dimostrazione. Elenchiamo tutti gli esiti dei due esperimenti:

$$m \text{ righe:} \quad \left\{ \begin{array}{cccccl} (1,1) & (1,2) & \cdots & (1,n) & \leftarrow n \text{ elementi} \\ (2,1) & (2,2) & \cdots & (2,n) & \leftarrow n \text{ elementi} \\ \vdots & \vdots & & \vdots & \\ (m,1) & (m,2) & \cdots & (m,n) & \leftarrow n \text{ elementi} \end{array} \right.$$

dove si intende che l'esito finale è la coppia ordinata (i, j) se il primo esperimento ha prodotto esito i e il secondo ha prodotto esito j. L'insieme dei possibili esiti consiste di m righe, ognuna contenente n elementi. Quindi vi sono in tutto mn esiti possibili.

Notiamo che sequenze distinte di esiti dei due esperimenti producono esiti finali distinti; in altri termini (i, j) è un risultato distinto da (j, i).

Esempio. Si lanciano a caso due dadi da gioco. Quanti sono i risultati possibili?
Soluzione. Il lancio del primo dado dà l'esito del primo esperimento, e il lancio del secondo dado dà l'esito del secondo esperimento; per il principio fondamentale del calcolo combinatorio vi sono in tutto $6 \times 6 = 36$ risultati possibili.

Esempio. Un campionato di calcio prevede 10 partite in una giornata, ed ognuna può dar luogo a 3 tipi di risultati (1, X, 2). Se si deve scegliere una partita e un risultato, quante sono le scelte possibili?
Soluzione. Si può vedere la scelta della partita come l'esito del primo esperimento e la scelta del risultato come l'esito del secondo esperimento; per il principio fondamentale del calcolo combinatorio vi sono in tutto $10 \times 3 = 30$ scelte possibili.

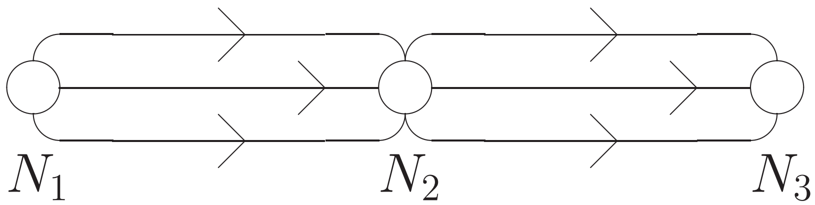
Esempio. Uno studente può inserire nel piano di studi 2 insegnamenti a scelta, di cui uno deve essere selezionato da un elenco di 9 insegnamenti e l'altro da un elenco distinto di 10 insegnamenti. In quanti modi diversi può completare il piano di studi?
Soluzione. Si hanno in totale $9 \times 10 = 90$ modi diversi.

Esempio. Quante sono le funzioni booleane definite su $\{0,1\}$?

Soluzione. Sia $f : \{0,1\} \to \{0,1\}$. Per ogni $x \in \{0,1\}$ risulta $f(x)$ uguale a 0 o 1, con $f(0)$ corrispondente all'esito del primo esperimento e $f(1)$ all'esito del secondo esperimento. Vi sono pertanto $2 \times 2 = 4$ funzioni siffatte.

x	$f(x) = 0$	$f(x) = x$	$f(x) = \overline{x} = 1 - x$	$f(x) = 1$
0	0	0	1	1
1	0	1	0	1

Esempio. Si cosideri il seguente grafo orientato. Quanti sono i percorsi distinti possibili che portano dal nodo N_1 al nodo N_3?



Soluzione. Vi sono $3 \times 3 = 9$ percorsi distinti.

Esempio. Si lancia a caso un dado da gioco. Se esce un numero pari si lancia nuovamente il dado. Se esce un numero dispari si lancia un dado truccato, in cui il 6 è stato modificato in 5. Quanti sono i risultati possibili?

Soluzione. Tenendo conto delle alternative relative al primo lancio, per il principio fondamentale del calcolo combinatorio vi sono in tutto

$$3 \times 6 + 3 \times 5 = 33 \quad \text{esiti possibili}$$

3×6 sono gli esiti in cui al primo lancio esce un numero pari,
3×5 sono gli esiti in cui al primo lancio esce un numero dispari.

Esempio. Un'urna contiene un dado regolare ed un dado con numeri 7, 8, 9, 10, 11, 11. Si lancia un dado estratto a caso dall'urna. Quanti sono i risultati possibili?

Soluzione. Tenendo conto delle 2 alternative relative all'estrazione, vi sono in tutto $6 + 5 = 11$ esiti possibili.

Esempio. Si effettuano due esperimenti. Il primo ha m possibili esiti. Se il primo esperimento produce l'esito $i \in \{1, 2, \ldots, m\}$, il secondo può avere k_i possibili esiti, con $i = 1, 2, \ldots, m$. Supponendo che sequenze distinte di esiti dei due esperimenti producano esiti finali distinti, quanti sono gli esiti finali possibili?

Soluzione. Gli esiti possibili sono

$$\sum_{i=1}^{m} k_i = k_1 + k_2 + \ldots + k_m.$$

Esempio. Si lanciano due dadi. Se il prodotto dei due numeri usciti è dispari allora vince Dario, se è pari allora vince Piero. Chi ha più possibilità di vincere?

Soluzione. Dario vince se il prodotto è dispari, ossia se entrambi i numeri usciti sono dispari. Ciò si realizza in $3 \times 3 = 9$ modi distinti.

Piero vince se il prodotto è pari, ossia se almeno uno dei numeri usciti è pari. Ciò si realizza in $3 \times 6 + 3 \times 3 = 18 + 9 = 27$ modi distinti (pari e qualsiasi oppure dispari e pari). Quindi Piero è avvantaggiato.

Principio fondamentale (generalizzato) del calcolo combinatorio. Si realizzino r esperimenti. Si supponga che il primo esperimento abbia n_1 esiti possibili, e che per ognuno di questi il secondo esperimento abbia n_2 esiti possibili, e ancora che per ognuno degli esiti dei primi 2 esperimenti il terzo esperimento abbia n_3 esiti possibili, ecc. Allora, se sequenze distinte di esiti degli r esperimenti producono esiti finali distinti, allora gli r esperimenti producono in tutto $n_1 \cdot n_2 \cdots n_r$ esiti possibili.

Esempio. (i) Quante sono le targhe automobilistiche formate da 7 caratteri, di cui 3 sono numeri e 4 sono lettere (scelte tra le 26 lettere dell'alfabeto anglosassone)?
(ii) Quante targhe vi sono escludendo le ripetizioni tra numeri e lettere?
Soluzione. Per il principio fondamentale (generalizzato) del calcolo combinatorio,
(i) vi sono $10 \times 10 \times 10 \times 26 \times 26 \times 26 \times 26 = 456\,976\,000$ targhe possibili;
(ii) vi sono $10 \times 9 \times 8 \times 26 \times 25 \times 24 \times 23 = 258\,336\,000$ targhe senza ripetizioni.

Esempio. Quanti sono i risultati possibili se si lancia a caso una moneta per n volte, se l'ordine è rilevante?

Soluzione. Ognuno degli n esperimenti consistenti nel lancio della moneta ha 2 possibili esiti, e quindi i risultati possibili sono $2 \times 2 \times \cdots \times 2 = 2^n$.

Ad esempio, per $n = 3$ si hanno 8 risultati: $ccc, cct, ctc, ctt, tcc, tct, ttc, ttt$.

Esempio. Si lancia un dado e poi si lanciano tante monete pari al risultato del lancio del dado. Quanti sono gli esiti finali possibili?

Soluzione. Gli esiti possibili sono

$$2 + 2^2 + 2^3 + 2^4 + 2^5 + 2^6 = \sum_{i=1}^{6} 2^i = 2^7 - 2 = 126,$$

avendo usato la seguente formula per $n = 6$ e $x = 2$:

$$\sum_{i=1}^{n} x^i = \frac{x^{n+1} - 1}{x - 1} - 1, \quad \text{valida per } x \neq 1.$$

Esempio. In quanti modi si possono scegliere r oggetti in un insieme di n oggetti distinti, se l'ordine delle scelte è rilevante? (Corrisponde a estrazioni senza reinserimento)

Soluzione. Ci sono n modi per scegliere il primo oggetto, per ognuno di questi vi sono $n-1$ modi per scegliere il secondo oggetto, e così via, vi sono $n-r+1$ modi per scegliere l'r-esimo oggetto, pertanto per il principio fondamentale (generalizzato) del calcolo combinatorio vi sono in totale

$$(n)_r = \underbrace{n(n-1)(n-2)\cdots(n-r+1)}_{r \text{ fattori}} = \frac{n!}{(n-r)!}$$

modi, dove il simbolo $(n)_r$ denota il fattoriale discendente. Notiamo che

$$n! = n(n-1)(n-2)\cdots(n-r+1)\cdot(n-r)(n-r-1)\cdots 1 = (n)_r \cdot (n-r)!$$

Esempio. Quanti sono i risultati possibili se si estraggono in sequenza 6 biglie da un'urna contente 90 biglie distinte (tenendo conto dell'ordine delle estrazioni)?

Soluzione. Nella prima estrazione vi sono 90 possibili risultati, nella seconda ve ne sono 89, e così via nella sesta ve ne sono 85, quindi i risultati possibili sono in tutto $(90)_6 = 90 \times 89 \times 88 \times 87 \times 86 \times 85 = 448\,282\,533\,600$.

Esempio. Si cosideri il seguente grafo. Avendo a disposizione 3 colori distinti, in quanti modi è possibile colorare i nodi del grafo dando colori diversi a nodi adiacenti? Cosa cambia se si aggiunge un arco che congiunge N_1 a N_4?

Soluzione. Vi sono $3 \times 2 \times 2 \times 2 = 24$ colorazioni possibili. Nel secondo caso 18.

Esempio. Risolvere il problema dell'esempio precedente quando i colori disponibili sono k.

Soluzione. Analogamente, nel primo caso vi sono $k\,(k-1)^3$ colorazioni possibili, mentre nel secondo ve ne sono $k\,(k-1)\,[k-1+(k-2)^2]$.

1.3 Permutazioni

In quanti modi si possono ordinare le lettere a, b, c? Per il principio fondamentale del calcolo combinatorio i casi possibili sono $3 \times 2 \times 1 = 6$: abc, acb, bac, bca, cab, cba. Le *permutazioni* di n oggetti distinti sono tutte le sequenze ordinate ottenibili scambiando gli oggetti in tutti i modi possibili.

Le permutazioni distinte di n oggetti sono $n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1 = n!$

$$
\begin{array}{ll}
0! = 1 & 10! = 3\,628\,800 \\
1! = 1 & 11! = 39\,916\,800 \\
2! = 2 & 12! = 479\,001\,600 \\
3! = 6 & 13! = 6\,227\,020\,800 \\
4! = 24 & 14! = 87\,178\,291\,200 \\
5! = 120 & 15! = 1\,307\,674\,368\,000 \\
6! = 720 & 16! = 20\,922\,789\,888\,000 \\
7! = 5\,040 & 17! = 355\,687\,428\,096\,000 \\
8! = 40\,320 & 18! = 6\,402\,373\,705\,728\,000 \\
9! = 362\,880 & 19! = 121\,645\,100\,408\,832\,000
\end{array}
$$

Esempio. Si eseguono simultaneamente 10 programmi, di cui 4 sono scritti in Linguaggio C e 6 in Java. Viene poi stilato l'elenco dei programmi nell'ordine di terminazione delle esecuzioni. Sapendo che le 10 esecuzioni hanno avuto termine in istanti diversi,

(a) quanti sono i possibili elenchi dei programmi?

(b) quanti sono i possibili elenchi, se i programmi in C e Java compaiono in 2 liste, con precedenza alla lista dei programmi in C?

Soluzione. (a) Ad ogni elenco corrisponde una possibile permutazione di 10 oggetti, quindi la risposta è $10! = 3\,628\,800$.

(b) Vi sono 4! elenchi dei programmi in C e 6! elenchi dei programmi in Java; per il principio fondamentale del calcolo combinatorio vi sono quindi $4! \cdot 6! = 24 \cdot 720 = 17\,280$ diversi elenchi se i programmi in C e Java compaiono in 2 liste, con precedenza alla lista dei programmi in C, ad esempio $C_4\ C_3\ C_2\ C_1\ J_6\ J_5\ J_4\ J_3\ J_2\ J_1$

Esempio. Quanti sono gli anagrammi di S T A T I S T I C A ?

Soluzione. Se le 10 lettere da permutare fossero distinte vi sarebbero $10! = 3\,628\,800$ permutazioni possibili. Tuttavia le lettere non sono distinte: se permutiamo le lettere S tra di loro, le lettere T tra di loro, le lettere A tra di loro, e le lettere I tra di loro, si ottiene comunque la stessa parola. Il numero di anagrammi distinti è quindi

$$\frac{10!}{2! \cdot 3! \cdot 2! \cdot 2!} = \frac{3\,628\,800}{2 \cdot 6 \cdot 2 \cdot 2} = 75\,600.$$

Permutazioni di oggetti non tutti distinti

Un ragionamento analogo a quello svolto nell'esempio precedente mostra che vi sono

$$\frac{n!}{n_1!\, n_2! \cdots n_r!}$$

permutazioni distinte di n oggetti presi da r categorie, dei quali n_1 sono identici fra loro, n_2 sono identici fra loro e distinti dai precedenti, ..., n_r sono identici fra loro e distinti dai precedenti, con

$$n = n_1 + n_2 + \cdots + n_r.$$

Esempio. Si eseguono simultaneamente 10 programmi, di cui 4 sono scritti in Linguaggio C e 6 in Java. Viene poi stilato l'elenco dei programmi nell'ordine di terminazione delle esecuzioni. Quanti sono i possibili elenchi dei programmi se è indicato solo il loro linguaggio?

Soluzione. Trattandosi di permutazioni di 10 oggetti non tutti distinti, appartenenti a 2 categorie (con 4 casi nella prima categoria e 6 casi nella seconda categoria), gli elenchi possibili sono

$$\frac{10!}{4!\,6!} = \frac{3\,628\,800}{24 \cdot 720} = 210.$$

Esempio. Quanti sono i vettori booleani di dimensione n costituiti da k bit pari a 1 e da $n-k$ bit pari a 0?

Soluzione. I possibili vettori siffatti sono $\dfrac{n!}{k!\,(n-k)!}$.

Per $n = 4$ e $k = 2$ i vettori sono $\dfrac{4!}{2!\,(4-2)!} = 6$: 0011, 0101, 0110, 1001, 1010, 1100.

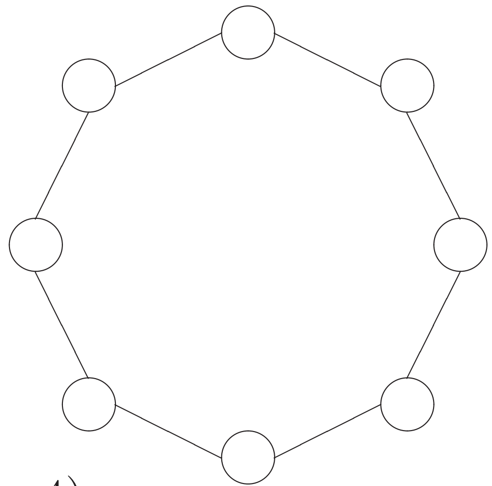
Esempio. In una giornata di un campionato di calcio vi sono 10 partite, ed ognuna può dar luogo a 3 tipi di risultati (1, X, 2). (i) Quanti sono i risultati possibili? (ii) Quante sono le sequenze di risultati possibili se vi sono 5 vincite in casa e 3 pareggi?

Soluzione. (i) I possibili risultati sono $3^{10} = 3 \times 3 \times \cdots \times 3 = 59\,049$.

(ii) Nel caso di 5 vincite in casa e 3 pareggi, allora le sequenze di risultati possibili sono

$$\frac{10!}{5!\,3!\,2!} = \frac{3\,628\,800}{120 \cdot 6 \cdot 2} = 2\,520.$$

Esercizio. Quante configurazioni si possono realizzare disponendo n oggetti in un allineamento circolare?



Soluzione. Il numero di *permutazioni circolari* è

$$\frac{n!}{n} = \frac{n(n-1)!}{n} = (n-1)!$$

poiché per ogni permutazione ve ne sono n equivalenti.

Ad esempio, per $n = 4$ si ha che (1, 4, 2, 3) equivale a (2, 3, 1, 4), quindi le $(n-1)! = 6$ permutazioni circolari sono

$(1,2,3,4) \quad (1,2,4,3) \quad (1,3,2,4) \quad (1,3,4,2) \quad (1,4,2,3) \quad (1,4,3,2)$

1.4 Disposizioni e combinazioni

Quanti insiemi di r oggetti si possono formare a partire da n oggetti distinti?

Per risolvere questo problema va specificato se gli insiemi da formare sono *ordinati* (in cui l'ordine è rilevante) o *non ordinati* (dove l'ordine non è rilevante).

Nel caso di insiemi *ordinati* le sequenze da formare si dicono **disposizioni semplici** se non sono ammesse ripetizioni, altrimenti si dicono **disposizioni con ripetizioni**.

Per il principio fondamentale del calcolo combinatorio

– il numero di disposizioni semplici di n oggetti raggruppati in r classi è

$$D_{n,r} = \underbrace{n(n-1)(n-2)\cdots(n-r+1)}_{r \text{ fattori}} = \frac{n!}{(n-r)!} = (n)_r$$

dove $(n)_r := n(n-1)(n-2)\cdots(n-r+1)$ è detto *fattoriale discendente*;

– il numero di disposizioni con ripetizioni di n oggetti raggruppati in r classi è

$$D'_{n,r} = \underbrace{n \cdot n \cdot n \cdots n}_{r \text{ fattori}} = n^r.$$

Esempio. Quante parole di lunghezza 2 si possono formare da un alfabeto di 4 lettere
(a) se le lettere non possono ripetersi?
(b) se le lettere possono ripetersi?

Soluzione. (a) Si tratta di disposizioni semplici di $n = 4$ oggetti raggruppati in $r = 2$ classi, quindi $D_{4,2} = (4)_2 = 4 \cdot 3 = 12$.
(ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc)
(b) Si tratta di disposizioni con ripetizioni di $n = 4$ oggetti raggruppati in $r = 2$ classi, quindi $D'_{4,2} = 4^2 = 16$.
(<u>aa</u>, ab, ac, ad, ba, <u>bb</u>, bc, bd, ca, cb, <u>cc</u>, cd, da, db, dc, <u>dd</u>)

Esempio. Da un'urna che contiene n biglie numerate da 1 a n si effettuano k estrazioni a caso. Quante sono le sequenze ordinate distinte ottenibili se le estrazioni
(a) si effettuano senza reinserimento?
(b) si effettuano con reinserimento?

Soluzione. (a) $(n)_k = n(n-1)(n-2)\cdots(n-k+1)$; (b) n^k.

($\star\star$) **Esempio.** Quante sono le operazioni distinte definite su 2 variabili booleane?

Soluzione. Essendo $x_1 \in \{0,1\}$ e $x_2 \in \{0,1\}$, vi sono 4 coppie (x_1, x_2) distinte. Ad ognuna di queste si può assegnare valore 0 o 1. Si tratta quindi di determinare il numero di disposizioni con ripetizioni di 2 oggetti (i valori 0 e 1) in quattro classi (le coppie (x_1, x_2)); vi sono pertanto $D'_{2,4} = 2^4 = 16$ operazioni distinte.

x_1	x_2	0	$x_1 \wedge x_2$	$x_1 \wedge \overline{x_2}$	x_1	$\overline{x_1} \wedge x_2$	x_2	$x_1 \oplus x_2$	$x_1 \vee x_2$
0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1

x_1	x_2	$\overline{x_1} \wedge \overline{x_2}$	$\overline{x_1 \oplus x_2}$	$\overline{x_2}$	$x_1 \vee \overline{x_2}$	$\overline{x_1}$	$\overline{x_1} \vee x_2$	$\overline{x_1} \vee \overline{x_2}$	1
0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1

($\wedge$ = AND, $\vee$ = OR, $\oplus$ = OR ESCLUSIVO)

Consideriamo il problema di determinare quanti insiemi non ordinati di r oggetti si possono formare a partire da n oggetti distinti.

Nel caso di insiemi *non ordinati* le sequenze si dicono **combinazioni semplici** se non sono ammesse ripetizioni, altrimenti si dicono **combinazioni con ripetizioni**.

In generale, dato che $(n)_r = n(n-1)\cdots(n-r+1)$ rappresenta il numero di scelte di r oggetti tra n, tenendo conto dell'ordine nel quale questi vengono selezionati, e dato che ogni insieme di r oggetti viene in tal modo contato $r!$ volte, si ha che il numero di sottoinsiemi di r oggetti che si possono formare da un insieme di n oggetti è

$$C_{n,r} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{(n-r)!\,r!} = \frac{(n)_r}{r!}.$$

Notiamo anche che $D_{n,r} = C_{n,r} \cdot r!$, in quanto il numero di sequenze ordinate è uguale al numero di sequenze non ordinate per il numero di permutazioni di r oggetti, e pertanto

$$C_{n,r} = \frac{D_{n,r}}{r!} = \frac{(n)_r}{r!}$$

Esempio. Al termine del campionato di calcio di serie A le prime 4 squadre ricevono dei premi in denaro di importi differenti. In quanti modi diversi si può costruire l'elenco delle squadre che riceveranno i quattro premi?

Soluzione. Si tratta di disposizioni semplici, poiché l'ordine è rilevante, pertanto la soluzione è

$$D_{20,4} = (20)_4 = 20 \cdot 19 \cdot 18 \cdot 17 = 116\,280$$

Esempio. Al termine del campionato di calcio di serie A le prime 4 squadre delle 20 partecipanti accedono alla UEFA Champions League. In quanti modi diversi si possono individuare le squadre che accederanno al torneo europeo?

Soluzione. Tra gli elenchi riferiti all'esempio precedente, ogni gruppo di 4 squadre compare in tutte le 4! permutazioni possibili. Poiché ora l'ordine non è rilevante, occorre dividere il numero di disposizioni semplici per quante volte una sequenza viene ripetuta, pertanto la soluzione è data dal numero di combinazioni semplici

$$C_{20,4} = \frac{D_{20,4}}{4!} = \frac{(20)_4}{4!} = \frac{116\,280}{24} = 4845$$

Per $r = 0, 1, \ldots, n$ definiamo il *coefficiente binomiale*

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!} = \frac{(n)_r}{r!} = \frac{n(n-1)\cdots(n-r+1)}{r!} > 0.$$

Se $r < 0$ o se $r > n$ si pone

$$\binom{n}{r} = 0.$$

Il numero $C_{n,r}$ di combinazioni semplici di n oggetti raggruppati in r classi è

$$C_{n,r} = \binom{n}{r} = \frac{n!}{(n-r)!\,r!} = \frac{(n)_r}{r!} = \frac{n(n-1)\cdots(n-r+1)}{r!}.$$

Questo denota il numero di sottoinsiemi di dimensione r che si possono formare con gli elementi di un insieme di dimensione n senza tener conto dell'ordine della selezione.

Il numero $C'_{n,r}$ di combinazioni con ripetizioni di n oggetti raggruppati in r classi è

$$C'_{n,r} = \binom{n+r-1}{r} = \frac{(n+r-1)!}{(n-1)!\,r!} = \frac{(n+r-1)_r}{r!} = \frac{(n+r-1)(n+r-2)\cdots n}{r!}.$$

Tabella riepilogativa

Quanti insiemi di k oggetti si possono formare a partire da n oggetti distinti?

	Disposizioni (l'ordine è rilevante)	**Combinazioni** (l'ordine non è rilevante)
semplici (senza ripetizioni)	$D_{n,k} = (n)_k = \dfrac{n!}{(n-k)!}$	$C_{n,k} = \binom{n}{k} = \dfrac{(n)_k}{k!}$
composte (con ripetizioni)	$D'_{n,k} = n^k$	$C'_{n,k} = \binom{n+k-1}{k}$

Esempio. Quante combinazioni di 4 oggetti in gruppi di 2 si possono formare
(a) nel caso di combinazioni semplici? (gli oggetti non possono ripetersi)
(b) nel caso di combinazioni con ripetizioni? (gli oggetti possono ripetersi)

Soluzione. (a) $C_{4,2} = (4)_2/2! = (4 \cdot 3)/2 = 6 \;\; (ab, ac, ad, bc, bd, cd)$.
(b) $C'_{4,2} = (5)_2/2! = (5 \cdot 4)/2 = 10 \;\; (\underline{aa}, ab, ac, ad, \underline{bb}, bc, bd, \underline{cc}, cd, \underline{dd})$.

Esempio. In quanti modi si possono scegliere 3 oggetti da un insieme di 20?

Soluzione. $C_{20,3} = \binom{20}{3} = \dfrac{(20)_3}{3!} = \dfrac{20 \cdot 19 \cdot 18}{3 \cdot 2 \cdot 1} = \dfrac{6840}{6} = 1140.$

Esempio. Una classe di tango argentino ha 22 studenti, 10 donne e 12 uomini. In quanti modi si possono formare 5 coppie?

Soluzione. Vi sono $\binom{10}{5}$ modi di selezionare 5 persone da 10 donne, $\binom{12}{5}$ modi di selezionare 5 persone da 12 uomini, e 5! modi di accoppiare 5 donne con 5 uomini, quindi la soluzione è $\binom{10}{5}\binom{12}{5} 5! = 252 \cdot 792 \cdot 120 = 23\,950\,080.$

Esempio. (a) Quanti comitati composti da 2 donne e 3 uomini si possono formare da un gruppo di 5 donne e 7 uomini? (b) Quanti sono i comitati se 2 uomini che hanno litigato rifiutano di sedere insieme nel comitato?

Soluzione. (a) Ci sono $\binom{5}{2}$ possibili insiemi con 2 donne, e $\binom{7}{3}$ possibili insiemi di 3 uomini; segue allora dal principio fondamentale del calcolo combinatorio che vi sono

$$\binom{5}{2}\binom{7}{3} = \frac{5 \cdot 4}{2 \cdot 1} \cdot \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 350 \quad \text{comitati possibili formati da 2 donne e 3 uomini.}$$

(b) Se due uomini rifiutano di far parte insieme nel comitato, vi sono $\binom{2}{0}\binom{5}{3}$ insiemi di 3 uomini che non contengono nessuno dei 2 litiganti, e $\binom{2}{1}\binom{5}{2}$ insiemi di 3 uomini che contengono esattamente 1 dei 2 litiganti, e quindi vi sono $\binom{2}{0}\binom{5}{3} + \binom{2}{1}\binom{5}{2} = 30$ gruppi di 3 uomini senza i 2 litiganti. In totale vi sono $30 \cdot \binom{5}{2} = 300$ comitati.

Notiamo che il numero di comitati che contengono i 2 uomini che hanno litigato è: $\binom{5}{2}\binom{2}{2}\binom{5}{1} = 50$. Inoltre, risulta $\binom{7}{3} = \binom{2}{0}\binom{5}{3} + \binom{2}{1}\binom{5}{2} + \binom{2}{2}\binom{5}{1}$

Esempio. Quanti sono i risultati possibili nel gioco del Superenalotto?

Soluzione. Se si tiene conto dell'ordine delle estrazioni, i possibili risultati sono

$$D_{90,6} = (90)_6 = 90 \times 89 \times 88 \times 87 \times 86 \times 85 = 448\,282\,533\,600.$$

Tra tali risultati ve ne sono alcuni in cui si presentano gli stessi numeri permutati. Quindi, dividendo per il numero di possibili permutazioni dei numeri estratti, pari a

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720,$$

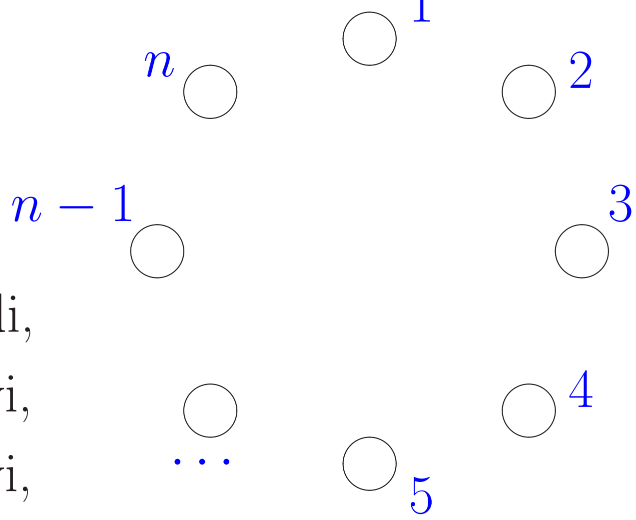
si ottiene il numero dei risultati possibili nel gioco del Superenalotto (ossia senza tener conto dell'ordine delle estrazioni):

$$\frac{90 \times 89 \times 88 \times 87 \times 86 \times 85}{6!} = \frac{448\,282\,533\,600}{720} = 622\,614\,630.$$

Notiamo che tale risultato coincide con le combinazioni semplici di 90 oggetti distinti in 6 classi:

$$C_{90,6} = \binom{90}{6} = \frac{D_{90,6}}{6!} = \frac{(90)_6}{6!} = \frac{90 \times 89 \times 88 \times 87 \times 86 \times 85}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = 622\,614\,630.$$

Esempio. Un network è costituito da n nodi. Quanti sono i collegamenti diretti che si possono attivare tra ciascun nodo e tutti i nodi rimanenti?



Soluzione. Il numero di collegamenti diretti è
$(n-1)+(n-2)+(n-3)+\ldots+2+1=\sum_{k=1}^{n-1} k$.
Infatti, ci sono

- $n-1$ collegamenti dal nodo 1 a tutti i rimanenti nodi,
- $n-2$ collegamenti dal nodo 2 a tutti i nodi successivi,
- $n-3$ collegamenti dal nodo 3 a tutti i nodi successivi,

  $\ldots$
- 2 collegamenti dal nodo $n-2$ ai nodi successivi,
- 1 collegamento dal nodo $n-1$ al nodo n.

Inoltre la soluzione coincide col numero di combinazioni semplici di n oggetti in 2 classi (ossia il numero di coppie non ordinate che si possono formare da n oggetti):

$$\mathcal{C}_{n,2}=\binom{n}{2}=\frac{n(n-1)}{2}, \qquad \text{pertanto} \quad \sum_{k=1}^{n-1} k=\binom{n}{2}=\frac{n(n-1)}{2}.$$

Esempio. Se 4 computer distinti devono smaltire 3 carichi di lavoro identici,
(a) quante sono le possibili distribuzioni dei 3 carichi?
(b) quante sono le possibili distribuzioni dei 3 carichi se ciascun computer può smaltire al più un solo carico?

Soluzione. (a) Si tratta di combinazioni (l'ordine dei carichi non è rilevante) con ripetizioni (un computer può ricevere più carichi). Quindi le possibili distribuzioni sono

$$\mathcal{C}'_{4,3} = \binom{4+3-1}{3} = \binom{6}{3} = \frac{(6)_3}{3!} = \frac{6 \times 5 \times 4}{3 \times 2 \times 1} = 20,$$

e sono così rappresentabili (**xyuv** significa che il computer 1 riceve **x** carichi, il computer 2 riceve **y** carichi, il computer 3 riceve **u** carichi, il computer 4 riceve **v** carichi):

$$\begin{array}{c} \mathbf{0111},\ \mathbf{1011},\ \mathbf{1101},\ \mathbf{1110}, \\ \mathbf{0012},\ \mathbf{0102},\ \mathbf{1002},\ \mathbf{0120},\ \mathbf{1020},\ \mathbf{1200}, \\ \mathbf{0021},\ \mathbf{0201},\ \mathbf{2001},\ \mathbf{0210},\ \mathbf{2010},\ \mathbf{2100}, \\ \mathbf{0003},\ \mathbf{0030},\ \mathbf{0300},\ \mathbf{3000}. \end{array}$$

Notiamo che $\binom{6}{3} = \binom{4}{3} + \binom{4}{2} \times 2 + \binom{4}{1} = 4 + 6 \times 2 + 4 = 20$.

(b) Poiché non sono ammesse ripetizioni la soluzione è $\mathcal{C}_{4,3} = \binom{4}{3} = 4$.

Tavola (di Tartaglia-Newton) dei coefficienti binomiali

$n\backslash k$	0	1	2	3	4	5	6	7	somma
0	1	0	0	0	0	0	0	0	$1 = 2^0$
1	1	1	0	0	0	0	0	0	$2 = 2^1$
2	1	2	1	0	0	0	0	0	$4 = 2^2$
3	1	3	3	1	0	0	0	0	$8 = 2^3$
4	1	4	6	4	1	0	0	0	$16 = 2^4$
5	1	5	10	10	5	1	0	0	$32 = 2^5$
6	1	6	15	20	15	6	1	0	$64 = 2^6$
7	1	7	21	35	35	21	7	1	$128 = 2^7$

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{(n)_k}{k!}$$

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n;$$

$$\binom{n}{0} = \frac{n!}{0!\,n!} = \frac{n!}{n!} = 1; \qquad \binom{n}{n} = \frac{n!}{n!\,0!} = \frac{n!}{n!} = 1;$$

$$\binom{n}{1} = \frac{n!}{1!\,(n-1)!} = \frac{n\,(n-1)!}{(n-1)!} = n; \qquad \binom{n}{n-1} = \frac{n!}{(n-1)!\,1!} = \frac{n\,(n-1)!}{(n-1)!} = n;$$

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n!}{(n-k)!\,k!} = \binom{n}{n-k}.$$

Formula di ricorrenza dei coefficienti binomiali

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \qquad 1 \le r \le n.$$

(⋆⋆) Dimostrazione analitica

$$\binom{n-1}{r-1} + \binom{n-1}{r} = \frac{(n-1)!}{(r-1)!\,(n-r)!} + \frac{(n-1)!}{r!\,(n-r-1)!}$$

$$= \frac{(n-1)!}{(r-1)!\,(n-r)\,(n-r-1)!} + \frac{(n-1)!}{r\,(r-1)!\,(n-r-1)!}$$

$$= \left[\frac{1}{n-r} + \frac{1}{r}\right] \frac{(n-1)!}{(r-1)!\,(n-r-1)!} = \frac{n}{r\,(n-r)}\,\frac{(n-1)!}{(r-1)!\,(n-r-1)!} = \binom{n}{r}.$$

(⋆⋆) Esercizio. Realizzare un programma che usi la formula di ricorrenza dei coefficienti binomiali per ottenere la tavola di Tartaglia-Newton per $\binom{n}{r}$, $0 \le r \le n \le N$, con N assegnato. Quali sono le condizioni iniziali da usare?

Formula di ricorrenza dei coefficienti binomiali

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \qquad 1 \leq r \leq n.$$

(⋆⋆) Dimostrazione combinatoria

Si basa sul fatto che i sottoinsiemi di r elementi di un insieme di n oggetti sono $\binom{n}{r}$. Consideriamo un insieme di n oggetti e fissiamo l'attenzione su uno di essi, che chiamiamo *oggetto 1*. Vi sono $\binom{n-1}{r-1}\binom{1}{1}$ sottoinsiemi di r elementi che contengono l'oggetto 1. Inoltre vi sono $\binom{n-1}{r}\binom{1}{0}$ sottoinsiemi di r elementi che non contengono l'oggetto 1. La somma dei termini $\binom{n-1}{r-1} + \binom{n-1}{r}$ fornisce il numero $\binom{n}{r}$ di sottoinsiemi di r elementi.

Ogni sottoinsieme A di un insieme $\{a_1, a_2, \ldots, a_n\}$ si può rappresentare con un vettore booleano $(x_1, x_2, \ldots, x_n)$, dove $x_i = \mathbf{1}$ se $a_i \in A$ e $x_i = \mathbf{0}$ altrimenti, per $i = 1, 2, \ldots, n$. Quindi ad ogni sottoinsieme corrisponde un solo vettore e viceversa. Poiché in totale vi sono 2^n vettori, tanti sono i possibili sottoinsiemi. Poiché vi sono $\binom{n}{k}$ vettori aventi k bit pari a $\mathbf{1}$, tanti sono i sottoinsiemi di cardinalità k, per $k = 0, 1, \ldots, n$.

Esempio. Per $n = 4$ vi sono $2^4 = 16$ sottoinsiemi di $\{a, b, c, d\}$.

$$\binom{4}{0} = 1, \qquad \binom{4}{1} = 4, \qquad \binom{4}{2} = 6, \qquad \binom{4}{3} = 4, \qquad \binom{4}{4} = 1, \qquad \sum_{k=0}^{4} \binom{4}{k} = 2^4 = 16$$

	k		k
$(\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}) : \{\}$	0	$(\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0}) : \{a\}$	1
$(\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{1}) : \{d\}$	1	$(\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{1}) : \{a, d\}$	2
$(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0}) : \{c\}$	1	$(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{0}) : \{a, c\}$	2
$(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{1}) : \{c, d\}$	2	$(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{1}) : \{a, c, d\}$	3
$(\mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{0}) : \{b\}$	1	$(\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{0}) : \{a, b\}$	2
$(\mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{1}) : \{b, d\}$	2	$(\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{1}) : \{a, b, d\}$	3
$(\mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{0}) : \{b, c\}$	2	$(\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{0}) : \{a, b, c\}$	3
$(\mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{1}) : \{b, c, d\}$	3	$(\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}) : \{a, b, c, d\}$	4

Teorema del binomio

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k\, y^{n-k}, \qquad n \geq 1.$$

$\binom{n}{k}$ è detto *coefficiente binomiale* perché interviene nello sviluppo del binomio

$$(x+y)^2 = \sum_{k=0}^{2} \binom{2}{k} x^k\, y^{2-k} = y^2 + 2yx + x^2,$$

$$(x+y)^3 = \sum_{k=0}^{3} \binom{3}{k} x^k\, y^{3-k} = y^3 + 3y^2x + 3yx^2 + x^3.$$

Esempio. Quanti sono i sottoinsiemi di un insieme di n elementi?

Soluzione. Poiché vi sono $\binom{n}{k}$ sottoinsiemi di dimensione k, dal teorema del binomio per $x = y = 1$ si ha (cf. anche l'ultima colonna della tavola dei coefficienti binomiali)

$$\sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n.$$

Nella somma $\displaystyle\sum_{k=0}^{n}\binom{n}{k}=2^n$ è incluso il caso $k=0$ che corrisponde all'insieme vuoto.
Quindi il numero di sottoinsiemi non vuoti di un insieme di n elementi è

$$\sum_{k=1}^{n}\binom{n}{k}=\sum_{k=0}^{n}\binom{n}{k}-\binom{n}{0}=2^n-1.$$

Analogamente, il numero di sottoinsiemi costituiti da almeno 2 elementi è

$$\sum_{k=2}^{n}\binom{n}{k}=\sum_{k=0}^{n}\binom{n}{k}-\binom{n}{1}-\binom{n}{0}=2^n-n-1.$$

Esercizio. Utilizzare il teorema del binomio per dimostrare che

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}=0, \qquad n>0.$$

($\star\star$) **Esercizio.** Dimostrare il teorema del binomio per induzione su n. In modo analogo, mostrare che risulta

$$(x+y+z)^n=\sum_{k=0}^{n}\sum_{r=0}^{n-k}\frac{n!}{k!\,r!\,(n-k-r)!}\,x^k\,y^r\,z^{n-k-r}, \qquad n\geq 1.$$

(⋆⋆) Proposizione. Per ogni $n \geq k \geq 0$ sussiste la seguente identità:

$$\sum_{j=k}^{n} \binom{j}{k} = \binom{n+1}{k+1}.$$

Dimostrazione. Procedendo per induzione su n, notiamo che per $n = k$ l'identità è valida, essendo

$$\binom{k}{k} = \binom{k+1}{k+1} = 1.$$

Supponendo valida l'identità per n, vediamo che essa sussiste anche per $n+1$. Si ha

$$\sum_{j=k}^{n+1} \binom{j}{k} = \sum_{j=k}^{n} \binom{j}{k} + \binom{n+1}{k} = \binom{n+1}{k+1} + \binom{n+1}{k}.$$

Ricordando la formula di ricorrenza $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$ si ottiene $\sum_{j=k}^{n+1} \binom{j}{k} = \binom{n+2}{k+1}$, da cui segue immediatamente la tesi.

Esempio. Calcolare quanti sono i vettori $(x_1, \ldots, x_k)$ nei quali
(a) ogni x_i è un intero positivo tale che $1 \le x_i \le n$;
(b) ogni x_i è un intero positivo tale che $1 \le x_i \le n$, ed inoltre ogni x_i è diverso da ciascun intero $x_1, x_2, \ldots, x_{i-1}$;
(c) ogni x_i è un intero positivo tale che $1 \le x_i \le n$, ed inoltre $x_1 < x_2 < \cdots < x_k$;
(d) ogni x_i è un intero positivo tale che $1 \le x_i \le n$, ed inoltre $x_1 \le x_2 \le \cdots \le x_k$.

Soluzione. (a) $D'_{n,k} = n^k$; (b) $D_{n,k} = (n)_k$; (c) $C_{n,k} = \binom{n}{k}$; (d) $C'_{n,k} = \binom{n+k-1}{k}$.

(⋆⋆) Esercizio. Dimostrare le seguenti uguaglianze:

$$\sum_{i_1=1}^{n} \sum_{\substack{i_2=1 \\ i_2 \neq i_1}}^{n} \cdots \sum_{\substack{i_k=1 \\ i_k \neq i_r \ \forall r<k}}^{n} 1 = (n)_k, \qquad \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} 1 = n^k,$$

$$\sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1-1} \cdots \sum_{i_k=1}^{i_{k-1}-1} 1 = \binom{n}{k}, \qquad \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1} \cdots \sum_{i_k=1}^{i_{k-1}} 1 = \binom{n+k-1}{k}.$$

Proposizione. Sussiste la seguente uguaglianza (detta di Vandermonde):

$$\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} = \binom{n}{0}\binom{m}{k} + \binom{n}{1}\binom{m}{k-1} + \cdots + \binom{n}{k}\binom{m}{0}.$$

Per la dimostrazione, si può procedere per via combinatoria. Ad esempio, considerando n biglie nere e m biglie bianche. In quanti modi si può formare un sacchetto di k biglie? Possiamo prendere 0 biglie nere e k biglie bianche, oppure 1 biglia nera e $k-1$ biglie bianche, e così via, fino a k biglie nere e 0 biglie bianche.

Ad esempio, supponiamo che per raggiungere una destinazione vi sono 7 percorsi, di cui 4 con pedaggio e 3 senza pedaggio. Se occorre scegliere 3 percorsi, tale operazione può compiersi in un numero di modi pari a

$$\binom{7}{3} = \frac{(7)_3}{3!} = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 35.$$

La formula di Vandemonde dà il risultato in base ai numeri di percorsi di un dato tipo:

$$\binom{7}{3} = \binom{4}{0}\binom{3}{3} + \binom{4}{1}\binom{3}{2} + \binom{4}{2}\binom{3}{1} + \binom{4}{3}\binom{3}{0} = 1 + 12 + 18 + 4 = 35.$$

($\star\star$) Esercizio. Dimostrare l'uguaglianza [che è analoga a $(a+b)^k = \sum_{i=0}^k \binom{k}{i} a^i b^{k-i}$]

$$(n+m)_k = \sum_{i=0}^{k} \binom{k}{i} (n)_i (m)_{k-i}.$$

Soluzione. Dall'uguaglianza di Vandermonde

$$\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i},$$

svolgendo i coefficienti binomiali si ha:

$$\frac{(n+m)_k}{k!} = \sum_{i=0}^{k} \frac{(n)_i}{i!} \frac{(m)_{k-i}}{(k-i)!}.$$

Quindi si ricava

$$(n+m)_k = \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} (n)_i (m)_{k-i} = \sum_{i=0}^{k} \binom{k}{i} (n)_i (m)_{k-i}$$

da cui segue la tesi.

Esempio. Una squadra di calcetto è composta da 6 dilettanti e da 2 professionisti.

(a) In quanti modi si può scegliere una terna di giocatori per uno stage nazionale?

(b) Stabilire in quanti modi si può scegliere una terna di giocatori composta da

- 1 dilettante e 2 professionisti,
- 2 dilettanti e 1 professionista,
- 3 dilettanti.

(c) Fare uso dell'uguaglianza di Vandermonde per verificare la somma dei valori.

Soluzione. (a) Si tratta di combinazioni (ordine non rilevante) senza ripetizioni:

$$C_{8,3} = \binom{8}{3} = \frac{(8)_3}{3!} = \frac{8 \times 7 \times 6}{3 \times 2 \times 1} = 56$$

(b)

$$\binom{6}{1}\binom{2}{2} = 6 \cdot 1 = 6, \qquad \binom{6}{2}\binom{2}{1} = \frac{6 \cdot 5}{2} \cdot 2 = 30, \qquad \binom{6}{3}\binom{2}{0} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2} \cdot 1 = 20$$

(c)

$$56 = \binom{8}{3} = \sum_{i=0}^{2} \binom{6}{i}\binom{2}{3-i} = 0 + \binom{6}{1}\binom{2}{2} + \binom{6}{2}\binom{2}{1} + \binom{6}{3}\binom{2}{0} = 6 + 30 + 20.$$

Esempio. Ad una corsa campestre partecipano 6 dilettanti e 2 professionisti.

(a) In quanti modi si può formare il podio dei 3 vincitori?

(b) Stabilire in quanti modi si può formare il podio dei 3 vincitori composto da

- 1 dilettante e 2 professionisti,
- 2 dilettanti e 1 professionista,
- 3 dilettanti.

(c) Verificare la somma dei valori ottenuti.

Soluzione. (a) Si tratta di disposizioni (ordine rilevante) senza ripetizioni:

$$D_{8,3} = (8)_3 = 8 \times 7 \times 6 = 336$$

(b)

$$(6)(2 \cdot 1) \times 3 = 36, \qquad (6 \cdot 5)(2) \times 3 = 180, \qquad (6 \cdot 5 \cdot 4) = 120$$

(c) Si ha

$$336 = (8)_3 = 0 + (6)(2 \cdot 1) \times 3 + (6 \cdot 5)(2) \times 3 + (6 \cdot 5 \cdot 4) = \sum_{i=0}^{3} \binom{3}{i} (6)_i (2)_{3-i},$$

che corrisponde all'uguaglianza $(n+m)_k = \sum_{i=0}^{k} \binom{k}{i} (n)_i (m)_{k-i}$.

Esempio. Da un lotto di $N = 20$ pezzi costituito da 4 pezzi difettosi e 16 buoni si estraggono $n = 10$ pezzi a caso.

(a) Quanti sono i possibili campioni?

(b) Quanti sono i possibili campioni che non contengono pezzi difettosi?

(c) Quanti sono i possibili campioni che contengono 1 pezzo difettoso e 9 buoni?

(d) Quanti sono i possibili campioni che contengono almeno 2 pezzi difettosi?

Soluzione. (a) I possibili campioni sono: $\binom{20}{10} = \frac{(20)_{10}}{10!} = 184\,756.$

(b) I possibili campioni che non contengono pezzi difettosi sono:

$$\binom{4}{0}\binom{16}{10} = 1\binom{16}{6} = \frac{(16)_6}{6!} = 8\,008.$$

(c) I possibili campioni che contengono 1 pezzo difettoso e 9 buoni sono:

$$\binom{4}{1}\binom{16}{9} = 4\binom{16}{7} = 4\,\frac{(16)_7}{7!} = 45\,760.$$

(d) I possibili campioni che contengono almeno 2 pezzi difettosi sono:

$$\binom{20}{10} - \binom{4}{0}\binom{16}{10} - \binom{4}{1}\binom{16}{9} = 184\,756 - 8\,008 - 45\,760 = 130\,988.$$

Equivalentemente, i possibili campioni che contengono almeno 2 pezzi difettosi sono:

$$\binom{4}{2}\binom{16}{8}+\binom{4}{3}\binom{16}{7}+\binom{4}{4}\binom{16}{6}=130\,988.$$

Esempio. Un lotto di 4 computer distinti deve essere assegnato a 2 laboratori.

(a) In quanti modi può essere fatto?

(b) E se ogni laboratorio deve ricevere almeno un computer?

Soluzione. (a) Assegnando k computer al primo laboratorio e $4-k$ al secondo laboratorio, l'assegnazione dei computer si può effettuare in $\binom{4}{k}$ modi; si ha quindi:

$$\sum_{k=0}^{4}\binom{4}{k}=2^4=16;$$

tenendo conto che ogni computer può essere assegnato a ciascuno dei 2 laboratori, la soluzione è anche esprimibile come $D'_{2,4}=2^4=16$.

(b) In questo caso si ha:

$$\sum_{k=1}^{3}\binom{4}{k}=2^4-\binom{4}{0}-\binom{4}{4}=14.$$

Esercizi per casa

1.1) Quanti sono i vettori booleani di lunghezza n
(i) che hanno almeno un elemento uguale a $\mathbf{1}$?
(ii) che hanno solo i primi due elementi uguali a $\mathbf{1}$?
(iii) che hanno tutti gli elementi uguali?

1.2) In una città ci sono 5 alberghi.
(i) In quanti modi 3 persone possono scegliere un albergo dove pernottare?
(ii) Cosa cambia se ogni persona deve scegliere un albergo diverso?

1.3) (i) Quanti sono i vettori booleani di lunghezza n?
(ii) Quanti sono i suddetti vettori con esattamente k elementi uguali a $\mathbf{0}$?
(iii) Quanti sono i suddetti vettori se i primi k elementi sono uguali?

1.4) Sia $S = \{1, 2, \ldots, 2n\}$; determinare
(i) il numero di sottoinsiemi di S aventi cardinalità k e che contengono il numero 1;
(ii) il numero dei sottoinsiemi di S aventi cardinalità k e che non contengono numeri pari.

1.5) Una squadra è formata da 11 titolari e 9 riserve. Selezionando 4 persone della squadra,
(i) quanti raggruppamenti distinti si possono formare?
(ii) quanti raggruppamenti distinti si possono formare, contenenti un solo titolare?
(iii) quanti raggruppamenti distinti si possono formare, contenenti almeno un titolare?

1.6) In quanti modi si possono disporre in fila 5 donne e 4 uomini in modo che 2 uomini non siano mai consecutivi?

1.7) Quanti sono i numeri di 6 cifre che contengono esattamente due volte la cifra 1, esattamente due volte la cifra 2 e non contengono lo 0?

1.8) Stabilire quante sono le sequenze del tipo abc tali che
(i) a è un intero tra 1 e 9, b e c sono interi tra 0 e 9, e risulta $a > b > c$;
(ii) a è un intero tra 1 e 9, b e c sono interi tra 0 e 9, e risulta $a \geq b \geq c$.

1.9) Calcolare le seguenti somme:

$$\text{(i)} \quad \binom{10}{2} + \binom{10}{3} \qquad\qquad \text{(ii)} \quad \sum_{k=1}^{7} \binom{8}{k} \qquad\qquad \text{(iii)} \quad \sum_{i=0}^{3} \binom{4}{i} \binom{4}{3-i}$$

1.10) Calcolare le seguenti espressioni:

$$\text{(i)} \quad \sum_{k=0}^{4} \binom{4}{k} 2^k \qquad\qquad \text{(ii)} \quad \sum_{k=0}^{3} \binom{3}{k} 2^n \qquad\qquad \text{(iii)} \quad \sum_{k=1}^{3} \binom{4}{k} 2^k 3^{n-k}$$

CAPITOLO 2 – Assiomi della probabilità

2.1 Introduzione

In questo capitolo introduciamo il concetto di

probabilità di un evento

e quindi mostriamo come le probabilità possano essere calcolate in certe situazioni.

Preliminarmente avremo però bisogno di definire i concetti di

spazio campionario

e di

evento di un esperimento.

2.2 Spazio campionario ed eventi

Chiameremo *esperimento* qualunque fenomeno il cui risultato non possa essere previsto con certezza. Sebbene l'esito dell'esperimento non sia noto a priori, supponiamo di conoscere l'insieme di tutti i possibili esiti, che chiameremo *spazio campionario* dell'esperimento e denoteremo con S; i suoi elementi sono detti *eventi elementari*.

Esempio. Nell'esperimento che consiste nel lancio di 2 dadi, lo spazio campionario consiste di $D'_{6,2} = 6^2 = 36$ elementi:

$$S = \{(i,j) : i,j = 1,2,3,4,5,6\},$$

dove (i,j) indica che il primo dado mostra il numero i e l'altro dado il numero j.

Esempio. Se l'esito dell'esperimento è l'ordine di arrivo di una competizione sportiva con n partecipanti, lo spazio campionario consiste di $n!$ elementi:

$$S = \{\text{tutte le } n! \text{ permutazioni di } (1,2,\ldots,n)\}$$

$$= \{(\omega_1,\omega_2,\ldots,\omega_n) : \omega_i \in \{1,2,\ldots,n\}\ \forall i,\ \omega_i \neq \omega_j\ \forall i \neq j\}$$

Esempio. Se l'esperimento consiste nel lanciare successivamente n monete, lo spazio campionario è costituito da $D'_{2,n} = 2^n$ elementi:

$$S = \{\omega_1\omega_2\cdots\omega_n : \forall i,\ \omega_i \in \{c,t\}\}; \qquad (S \text{ è finito})$$

$$\text{per } n=3: \qquad S = \{ccc, cct, ctc, ctt, tcc, tct, ttc, ttt\}$$

Esempio. Un esperimento consiste nel lanciare ripetutamente una moneta. Consideriamo come esito dell'esperimento il numero d'ordine del lancio in cui compare testa per la prima volta. Lo spazio campionario è l'insieme dei numeri naturali:

$$S = \{n : n = 1, 2, \ldots\} \qquad (S \text{ è infinito numerabile})$$

Esempio. Se l'esperimento consiste nel misurare il tempo di vita di un dispositivo elettronico, lo spazio campionario consiste nell'insieme dei numeri reali non negativi:

$$S = \{x \in \mathbb{R} : 0 \le x < \infty\} \qquad (S \text{ è infinito non numerabile})$$

Un sottoinsieme A dello spazio campionario sarà detto *evento*. Un evento è quindi un insieme di possibili esiti di un esperimento. Se l'esito di un esperimento è contenuto in A, diremo che l'evento A si verifica, oppure che A si realizza, oppure che A occorre.

Esempio. Nell'esperimento del lancio di 2 dadi, l'evento

$$A = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

si verifica quando la somma dei 2 dadi è 7.

Esempio. Nell'esperimento del lancio di 3 monete, l'evento

$$A = \{ccc, cct, ctc, ctt\}$$

si verifica quando al primo lancio esce croce.

Esempio. Nell'esperimento che consiste nel misurare il tempo di vita di un dispositivo elettronico, l'evento

$$A = \{x \in \mathbb{R} : 0 \leq x \leq 5\}$$

si verifica quando il dispositivo dura 5 ore o meno.

Operazioni tra eventi

• Dati due eventi A e B, definiamo il nuovo evento $A \cup B$, detto *unione* di A e B, formato da tutti gli esiti dell'esperimento che stanno in A o in B o in entrambi.

• Analogamente, dati due eventi A e B, definiamo il nuovo evento $A \cap B$, detto *intersezione* di A e B, formato da tutti gli esiti dell'esperimento che sono sia in A che in B. (Talora $A \cap B$ si indica con AB).

• Per ogni evento A definiamo il nuovo evento $\overline{A}$, detto *complementare* di A, formato da tutti gli esiti dell'esperimento che non sono in A. (Talvolta $\overline{A}$ si indica con A^c).

Esempio. Nell'esperimento del lancio di 2 monete, con $S = \{cc, ct, tc, tt\}$, se

$$A = \{cc, ct\} = \{\text{croce al primo lancio}\},$$

$$B = \{cc, tt\} = \{\text{nei due lanci si ha lo stesso risultato}\},$$

si ha

$$A \cup B = \{cc, ct, tt\}, \qquad A \cap B = \{cc\}, \qquad \overline{A} = \{tc, tt\}, \qquad \overline{B} = \{ct, tc\}.$$

• Il risultato di qualunque esperimento appartiene certamente allo spazio campionario; pertanto S viene detto *evento certo*.

• Un evento si dice *impossibile*, e si indica con $\emptyset$, se non contiene esiti dell'esperimento. ($\emptyset$ corrisponde all'insieme vuoto). Risulta ovviamente $\overline{S} = \emptyset$ e $\overline{\emptyset} = S$.

• L'evento complementare di $\overline{A}$ è l'evento A, ossia $\overline{\overline{A}} = A$ per ogni evento A.

• Due eventi A e B si dicono *incompatibili* se $A \cap B = \emptyset$ (la loro intersezione è vuota).

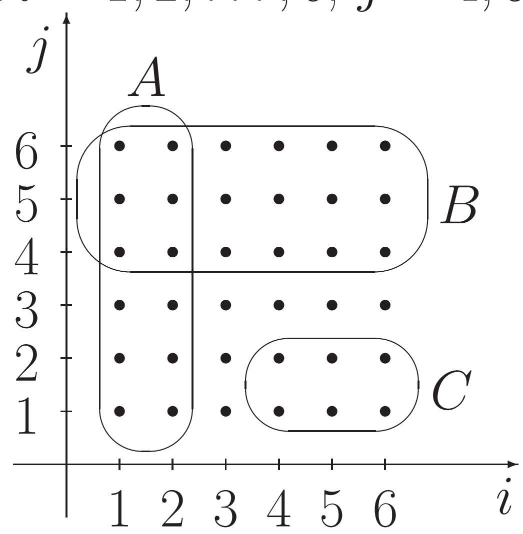
Esempio. Nell'esperimento del lancio di due dadi sia

$A = \{$al 1° lancio esce un numero minore di 3$\} = \{(i,j)\colon i = 1,2;\ j = 1,2,\ldots,6\}$

$B = \{$al 2° lancio esce un numero maggiore di 3$\} = \{(i,j)\colon i = 1,2,\ldots,6;\ j = 4,5,6\}$

$C = \{$al 1° lancio esce un numero maggiore di 3 e al 2° lancio esce un numero minore di 3$\}$

$= \{(i,j)\colon i = 4,5,6;\ j = 1,2\}$.



Risulta

$A \cap B = \{(i,j)\colon i = 1,2;\ j = 4,5,6\} \neq \emptyset$,

$A \cap C = \emptyset, \quad B \cap C = \emptyset$.

- $A_1, A_2, \ldots$ si dicono *a due a due incompatibili* se $A_i \cap A_j = \emptyset$ per ogni $i \neq j$.
- Più eventi (in numero finito o infinito) si dicono *necessari* se la loro unione è S.
- Gli eventi $A_1, A_2, \ldots$ costituiscono una *partizione* di S se sono non vuoti, necessari e a due a due incompatibili.

Notiamo che

– se più eventi sono a due a due incompatibili, allora se ne verifica al più uno di essi,

– se più eventi sono necessari, allora se ne verifica almeno uno di essi,

– se più eventi sono necessari e a due a due incompatibili, allora se ne verifica uno e uno solo di essi (ossia esattamente uno solo di essi).

Esempio. Sia $S = \{ccc, cct, ctc, ctt, tcc, tct, ttc, ttt\}$ lo spazio campionario nell'esperimento del lancio di 3 monete. Posto $A_k = \{\text{esce testa } k \text{ volte}\}$ per $k = 0, 1, 2, 3$, si ha che A_0, A_1, A_2, A_3 formano una partizione, essendo

$$A_0 = \{ccc\}, \quad A_1 = \{cct, ctc, tcc\}, \quad A_2 = \{ctt, tct, ttc\}, \quad A_3 = \{ttt\},$$

$$A_i \cap A_j = \emptyset \quad \text{per ogni } i \neq j, \qquad \bigcup_{k=0}^{3} A_k = S.$$

• Se $A_1, A_2, \ldots$ sono eventi, si definiscono l'unione e l'intersezione di questi come

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \cdots, \qquad \bigcap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \cdots;$$

$\bigcup_{i=1}^{\infty} A_i$ è l'evento formato dagli esiti compresi in almeno uno degli eventi $A_1, A_2, \ldots$;

$\bigcap_{i=1}^{\infty} A_i$ è l'evento formato dagli esiti che sono compresi in tutti gli eventi $A_1, A_2, \ldots$.

• Per ogni evento A risulta

$$A \cup \overline{A} = S, \quad A \cap \overline{A} = \emptyset, \quad A \cup S = S, \quad A \cap S = A, \quad A \cup \emptyset = A, \quad A \cap \emptyset = \emptyset.$$

• Dati due eventi A e B, se tutti gli esiti di A sono anche in B, allora diciamo che A è contenuto in B, oppure che A implica B, e scriviamo $A \subset B$.

• Se $A \subset B$ e $B \subset A$, diciamo che A e B coincidono, e scriviamo $A = B$.

• Per ogni coppia di eventi A e B si ha:

$$A \cap B \subset A \subset A \cup B, \qquad A \cap B \subset B \subset A \cup B.$$

• Risulta $A \subset B$ se e solo se $\overline{B} \subset \overline{A}$.

Esempio. Sia $S = \{ccc, cct, ctc, ctt, tcc, tct, ttc, ttt\}$ lo spazio campionario nell'esperimento del lancio di 3 monete. Posto

$$A = \{\text{esce sempre testa}\} = \{ttt\},$$

$$B = \{\text{esce almeno una volta testa}\} = \{cct, ctc, ctt, tcc, tct, ttc, ttt\},$$

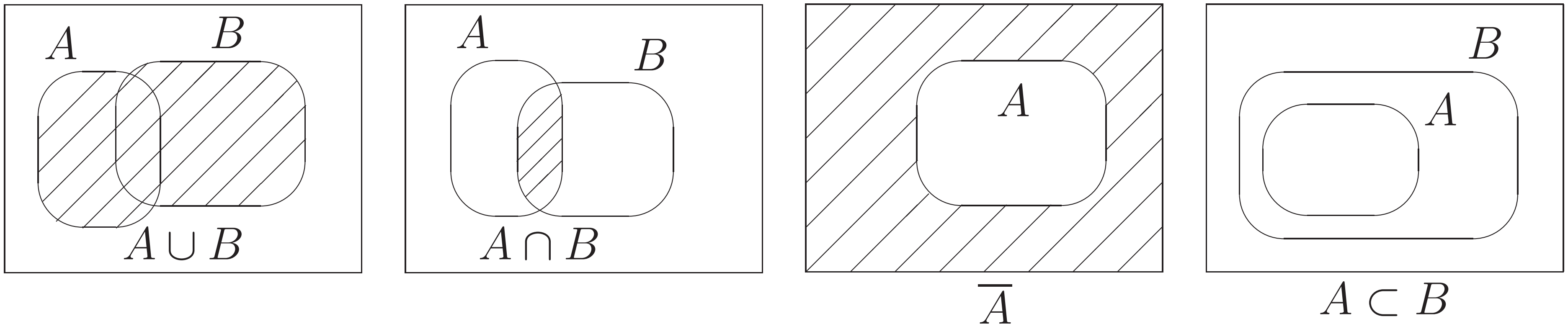
si ha $A \subset B$, e quindi

$$\overline{A} = \{\text{esce almeno una volta croce}\} = \{ccc, cct, ctc, ctt, tcc, tct, ttc\},$$

$$\overline{B} = \{\text{esce sempre croce}\} = \{ccc\},$$

da cui segue $\overline{B} \subset \overline{A}$.

Diagrammi di Venn



Esempio. Notiamo che se $A \subset B$, allora

(i) A e $\overline{B}$ sono incompatibili;

(ii) $A \cup (\overline{A} \cap B) = B$;

(iii) $A \cap (\overline{A} \cap B) = \emptyset$.

Tali relazioni sono facilmente verificabili facendo uso dei diagrammi di Venn.

Esempio. Nell'esperimento dell'estrazione di una biglia da un'urna contenente n biglie numerate da 1 a n, con $S = \{1, 2, \ldots, n\}$, consideriamo gli eventi che si realizzano quando viene estratta una biglia compresa tra 1 e k, per $k = 1, 2, \ldots, n$:

$$A_1 = \{1\}, \quad A_2 = \{1, 2\}, \quad \cdots \quad A_k = \{1, 2, \ldots, k\}, \quad \cdots \quad A_n = \{1, 2, \ldots, n\}.$$

Gli eventi $A_1, A_2, \ldots, A_n$ sono necessari, essendo

$$A_1 \cup A_2 \cup \cdots \cup A_n = \{1, 2, \ldots, n\} \equiv S.$$

Essi non sono a 2 a 2 incompatibili, essendo

$$A_i \cap A_j = A_{\min(i,j)} \neq \emptyset.$$

Ad esempio,

$$A_2 \cap A_4 = \{1, 2\} \cap \{1, 2, 3, 4\} = \{1, 2\} = A_2 \neq \emptyset.$$

Si vede facilmente che risulta

$$A_1 \subset A_2 \subset \cdots \subset A_k \subset \cdots \subset A_n.$$

Funzione indicatrice: $I[P] = 1$ se P è una proposizione vera, e $I[P] = 0$ altrimenti.

Esempio. Sia $S = \{(\omega_1, \omega_2, \ldots, \omega_n) : \omega_i \in \{\mathbf{0}, \mathbf{1}\} \ \forall i\}$ lo spazio campionario dell'esperimento che consiste nello scegliere a caso un vettore booleano di lunghezza n. Indichiamo con A_k l'evento costituito dai vettori aventi esattamente k bit pari a $\mathbf{1}$:

$$A_k = \{(\omega_1, \omega_2, \ldots, \omega_n) \in S : \sum_{i=1}^{n} I[\omega_i = \mathbf{1}] = k\}, \qquad k = 0, 1, \ldots, n$$

dove $I[\omega_i = \mathbf{1}] = 1$ se il bit ω_i è pari a $\mathbf{1}$, e $I[\omega_i = \mathbf{1}] = 0$ altrimenti, e pertanto $\sum_{i=1}^{n} I[\omega_i = \mathbf{1}]$ corrisponde al numero totale di bit pari a $\mathbf{1}$.

Gli eventi $A_0, A_1, \ldots, A_n$ sono necessari, ossia $A_0 \cup A_1 \cup \cdots \cup A_n = S$, in quanto tale unione fornisce l'insieme di vettori booleani per i quali il numero di bit pari a $\mathbf{1}$ è 0, o 1, o 2, …, o n. Gli eventi $A_0, A_1, \ldots, A_n$ sono non vuoti, sono a due a due incompatibili, essendo $A_i \cap A_j = \emptyset \ \ \forall i \neq j$, e quindi costituiscono una partizione di S.

Ad esempio, nel caso $n = 3$ si ha:

$$A_0 = \{\mathbf{000}\}, \quad A_1 = \{\mathbf{001}, \mathbf{010}, \mathbf{100}\}, \quad A_2 = \{\mathbf{011}, \mathbf{101}, \mathbf{110}\}, \quad A_3 = \{\mathbf{111}\}.$$

Proprietà

- commutative:

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A;$$

- associative:

$$(A \cup B) \cup C = A \cup (B \cup C), \qquad (A \cap B) \cap C = A \cap (B \cap C);$$

- distribuitive:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \qquad (A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

- formule di De Morgan:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \qquad \overline{A \cap B} = \overline{A} \cup \overline{B},$$

valide anche per un insieme finito di eventi $A_1, A_2, \ldots, A_n$:

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A}_i, \qquad \overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A}_i.$$

Esercizio. Siano A e B eventi distinti; stabilire se le seguenti identità sono vere:

(i) $A = (A \cap B) \cup (A \cap \overline{B})$;

(ii) $A \cap \overline{B} = \overline{B \cup \overline{A}}$;

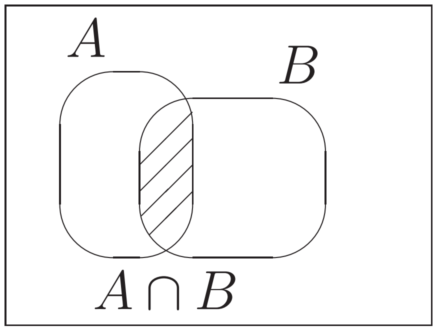
(iii) $\overline{A} \cup (A \cap B) = \overline{B \cap \overline{A}}$;

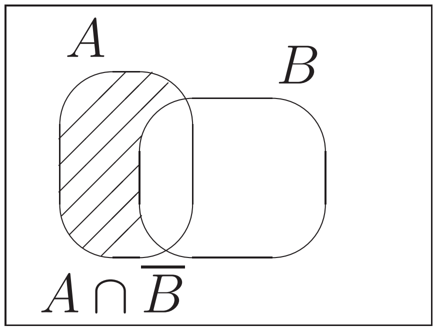
(iv) $A \cup \overline{(A \cup B)} = A \cap \overline{B}$;

(v) $(A \cap B) \cup \overline{(A \cup B)} = (A \cup \overline{B}) \cap (B \cup \overline{A})$.

Soluzione. (i) vera; (ii) vera; (iii) falsa; (iv) falsa; (v) vera.

Notiamo che gli eventi $A \cap B$ e $A \cap \overline{B}$ sono incompatibili.

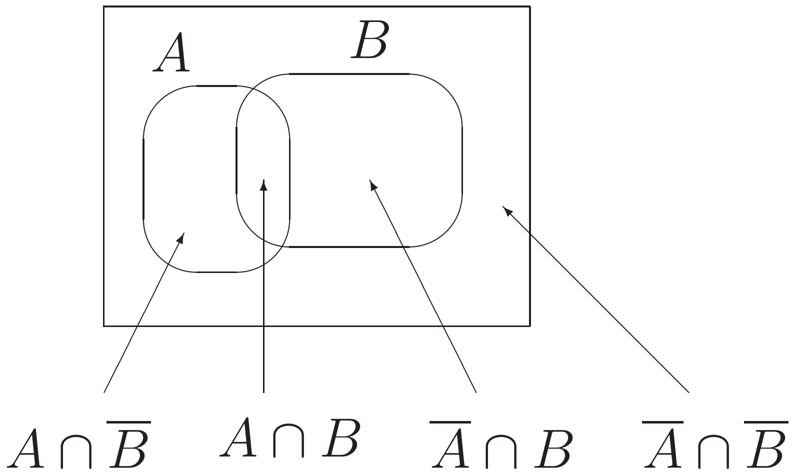




Osserviamo che, dati 2 eventi A e B qualsiasi, allora gli eventi

$$A \cap B, \qquad A \cap \overline{B}, \qquad \overline{A} \cap B, \qquad \overline{A} \cap \overline{B}$$

costituiscono una collezione di eventi necessari e a 2 a 2 incompatibili.



Dal grafico è facile verificare che

$$|A \cap B| + |A \cap \overline{B}| = |A|, \qquad |A \cup B| = |A| + |B| - |A \cap B|,$$

dove $|A|$ denota la cardinalità (ossia il numero di elementi) di A.

Esercizio. Un corso universitario è seguito da 100 studenti. Sappiamo che 80 studenti sono iscritti al 2º anno, e sappiamo che 30 studenti hanno svolto attività di tirocinio. Inoltre, sappiamo che 90 studenti sono del 2º anno o hanno svolto attività di tirocinio. Quanti studenti sono iscritti al 2º anno e non hanno svolto attività di tirocinio?

Soluzione. Consideriamo gli eventi $A = \{\text{studenti iscritti al 2º anno}\}$ e $B = \{\text{studenti che hanno svolto attività di tirocinio}\}$. Sappiamo che $|S| = 100$, $|A| = 80$, $|B| = 30$ e $|A \cup B| = 90$. Dovendo calcolare $|A \cap \overline{B}|$, notiamo che

$$|A \cup B| = |A| + |B| - |A \cap B|, \qquad |A| = |A \cap B| + |A \cap \overline{B}|,$$

e quindi $|A \cap B| = |A| + |B| - |A \cup B| = 80 + 30 - 90 = 20$, da cui risulta $|A \cap \overline{B}| = |A| - |A \cap B| = 80 - 20 = 60$. Riportiamo queste informazioni in tabella:

	B	$\overline{B}$	
A	$\lvert A \cap B\rvert$	$\lvert A \cap \overline{B}\rvert$	$\lvert A\rvert$
$\overline{A}$	$\lvert\overline{A} \cap B\rvert$	$\lvert\overline{A} \cap \overline{B}\rvert$	$\lvert\overline{A}\rvert$
	$\lvert B\rvert$	$\lvert\overline{B}\rvert$	$\lvert S\rvert$

	B	$\overline{B}$	
A	20	60	80
$\overline{A}$	10	10	20
	30	70	100

La classe degli eventi

Abbiamo già visto che un sottoinsieme A dello spazio campionario è detto evento. Più precisamente, la *classe degli eventi* $\mathcal{F}$ è una famiglia di sottoinsiemi di S tale che

(i) $S \in \mathcal{F}$;

(ii) se $A \in \mathcal{F}$ allora $\overline{A} \in \mathcal{F}$;

(iii) se $A_1, A_2, \ldots \in \mathcal{F}$ allora $A_1 \cup A_2 \cup \cdots \in \mathcal{F}$.

Da tali proprietà segue che $\mathcal{F}$ è una σ-algebra (sigma-algebra) di eventi, ed inoltre:

(iv) $\emptyset \in \mathcal{F}$;

(v) se $A_1, A_2, \ldots \in \mathcal{F}$ allora $A_1 \cap A_2 \cap \cdots \in \mathcal{F}$.

Quindi $\mathcal{F}$ è una famiglia di sottoinsiemi di S che contiene l'evento certo e l'evento impossibile, ed è chiusa rispetto a unione, intersezione, complementazione.

Esempio. Alcuni esempi di classi degli eventi: • $\mathcal{F} = \{\emptyset, S\}$;

• $\mathcal{F} = \{\emptyset, A, \overline{A}, S\}$;

• $\mathcal{F} = \{\emptyset, \{cc\}, \{ct\}, \{cc, ct\}, \{tc, tt\}, \{cc, tc, tt\}, \{ct, tc, tt\}, \{cc, ct, tc, tt\}\}$;

• $\mathcal{F} = \mathcal{P}(S)$ (insieme potenza, o insieme delle parti di S), con $|S| = n$ e $|\mathcal{P}(S)| = 2^n$.

2.3 Impostazioni frequentista e soggettiva della probabilità

La probabilità di un evento può essere definita in termini della frequenza relativa.

Supponiamo che un esperimento, il cui spazio campionario è S, venga ripetuto varie volte sotto le medesime condizioni. Per ogni evento E dello spazio campionario S, definiamo $n(E)$ come *frequenza assoluta*, ossia il numero di volte che si è verificato E nelle prime n ripetizioni dell'esperimento. Notiamo che risulta $0 \leq n(E) \leq n$.

Ad esempio, se in $n = 10$ lanci di una moneta esce 6 volte testa, allora $n(E) = 6$, con $E = \{t\}$ e $S = \{c, t\}$.

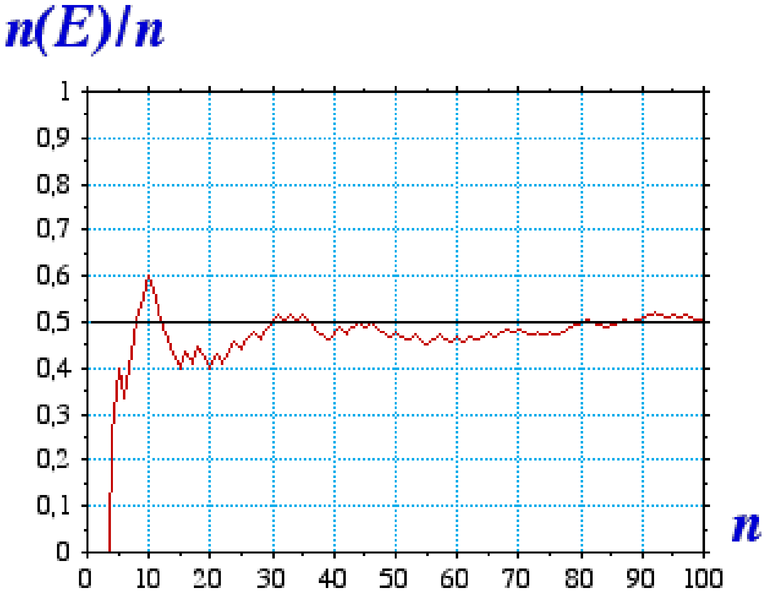
Allora, nell'**impostazione frequentista**, la *probabilità* dell'evento E è definita come

$$P(E) = \lim_{n\to\infty} \frac{n(E)}{n}$$

Cioè, $P(E)$ è definita come limite della *frequenza relativa* $n(E)/n$, ossia limite della proporzione del numero di volte che l'evento E si verifica.

Esempio. Lancio di una moneta ripetuto 100 volte; $S = \{c, t\}$; $E = \{t\}$

n	ω_n	$n(E)$	$n(E)/n$
1	c	0	0
2	c	0	0
3	c	0	0
4	t	1	0,25
5	t	2	0,4
6	c	2	0,33
7	t	3	0,43
8	t	4	0,5
9	t	5	0,56
10	t	6	0,6



Essendo $0 \le n(E) \le n$ e $n(S) = n$, notiamo che risulta

$$0 \le \frac{n(E)}{n} \le 1 \qquad \text{e} \qquad \frac{n(S)}{n} = 1.$$

Pertanto, dalla regola

$$P(E) = \lim_{n\to\infty} \frac{n(E)}{n}$$

ci si attende che debba risultare

$$0 \le P(E) \le 1 \qquad \text{e} \qquad P(S) = 1.$$

Inoltre, se A e B sono incompatibili, allora il numero di volte che si realizza un risultato presente in A oppure in B è pari al numero di volte che si realizza un risultato presente in A più numero di volte che si realizza un risultato presente in B, ossia risulta

$$n(A \cup B) = n(A) + n(B) \quad \text{se} \quad A \cap B = \emptyset,$$

e quindi ci si attende che sia

$$P(A \cup B) = P(A) + P(B) \quad \text{se} \quad A \cap B = \emptyset.$$

La definizione

$$P(E) = \lim_{n\to\infty} \frac{n(E)}{n}$$

presenta dei seri inconvenienti:

- come possiamo garantire che l'esperimento si possa ripetere più volte sotto le medesime condizioni?
- come possiamo sapere se $n(E)/n$ converge al limite ad una costante, che sia la stessa per ogni possibile successione di ripetizioni dell'esperimento?

Per esempio, supponiamo che l'esperimento consista nel lancio di una moneta.

- Come possiamo sapere che la proporzione di teste ottenute nei primi n lanci converga ad un valore quando n diventa grande?
- Inoltre, anche sapendo che questa converga, come possiamo sapere che ripetendo altre volte gli esperimenti, otterremo ancora lo stesso limite nella proporzione di teste?

È più ragionevole assumere un insieme di *assiomi* più semplici ed intuitivi per definire la probabilità, che includano le relazioni

$$0 \le P(E) \le 1, \quad P(S) = 1, \quad P(A \cup B) = P(A) + P(B) \ \text{ se } \ A \cap B = \emptyset.$$

Secondo l'**impostazione soggettiva**, alla quale ha dato impulso significativo il matematico Bruno De Finetti (1906-1985), la probabilità di un evento è il *grado di fiducia che un individuo ha nel verificarsi dell'evento*.

In questa affermazione la probabilità perde la caratteristica "assoluta" di numero legato intrinsecamente all'evento per dipendere dall'opinione soggettiva dell'osservatore.

Esprimendosi in termini di scommesse, nell'impostazione soggettiva, la probabilità di un evento è *il prezzo che un individuo ritiene equo pagare per ricevere* 1 *se l'evento si verifica e* 0 *se l'evento non si verifica.*

$$\omega = \text{risultato dell'esperimento} \qquad \begin{cases} \omega \in A & \Rightarrow \quad \text{riceviamo } 1 \\ \omega \in \overline{A} & \Rightarrow \quad \text{riceviamo } 0 \end{cases}$$

Va inoltre imposta la seguente condizione di coerenza:

Le probabilità degli eventi vanno attribuite in modo che non sia possibile ottenere con un insieme di scommesse una vincita certa o una perdita certa.

Chiariamo il ruolo della condizione di coerenza per la probabilità soggettiva.

Sia $P(A)$ la probabilità di un evento A secondo l'impostazione soggettiva. Nel pagare $P(A)$ e nel ricevere 1 oppure 0 si guadagna $1 - P(A)$ oppure $-P(A)$, quindi almeno $-P(A)$ e al massimo $1 - P(A)$. Se $P(A)$ fosse negativa si avrebbe certamente un guadagno positivo, mentre se $P(A)$ fosse maggiore di 1 si avrebbe certamente una perdita, e nei due casi la condizione di coerenza è violata. Si ha quindi $0 \leq P(A) \leq 1$.

Se consideriamo una scommessa sull'evento certo S, paghiamo $P(S)$ per ricevere certamente 1; si ha quindi certamente un guadagno pari a $1 - P(S)$. Se fosse $P(S) < 1$ si avrebbe una vincita certa mentre se fosse $P(S) > 1$ si avrebbe una perdita certa. Per la condizione di coerenza si ricava quindi $P(S) = 1$. Analogamente si può mostrare che nell'impostazione soggettiva si ha $P(A \cup B) = P(A) + P(B)$ se $A \cap B = \emptyset$.

L'impostazione soggettiva della probabilità è sottoposta a varie critiche. Lo schema di scommesse che la descrive è infatti contestabile perché la propensione o l'avversità al rischio cambia da persona a persona e per ogni individuo può variare con le circostanze connesse all'esperimento.

Le considerazioni inerenti l'impostazione frequentista e l'impostazione soggettiva suggeriscono di stabilire certe regole per definire la probabilità che includano le relazioni

$$0 \leq P(E) \leq 1, \quad P(S) = 1, \quad P(A \cup B) = P(A) + P(B) \text{ se } A \cap B = \emptyset,$$

anche estendendo l'ultima relazione al caso di successioni di eventi:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \text{ se } A_i \cap A_j = \emptyset \text{ per } i \neq j.$$

2.4 Assiomi della probabilità

Lo *spazio di probabilità* di un esperimento è $(S, \mathcal{F}, P)$, dove S è lo spazio campionario, $\mathcal{F}$ è la classe degli eventi, e $P : \mathcal{F} \to \mathbb{R}$ è una funzione tale che per ogni evento A esiste un reale $P(A)$, definito come *probabilità* di A, per cui valgono i seguenti 3 assiomi.

Assioma 1. Per ogni $A \in \mathcal{F}$ si ha

$$0 \leq P(A) \leq 1$$

Assioma 2.

$$P(S) = 1$$

Assioma 3. (Proprietà di additività numerabile) Per ogni successione di eventi $A_1, A_2, \ldots$ a due a due incompatibili (ossia tali che $A_i \cap A_j = \emptyset$ se $i \neq j$), si ha

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Proposizione.

$$P(\emptyset) = 0$$

($\star\star$) **Dimostrazione.** Consideriamo una successione di eventi $A_1, A_2, \ldots$, dove $A_1 = S$, $A_2 = A_3 = \ldots = \emptyset$; allora, essendo gli eventi $A_1, A_2, \ldots$ a due a due incompatibili ed essendo

$$S = \bigcup_{i=1}^{\infty} A_i$$

dall'Assioma 3 ricaviamo

$$P(S) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = P(S) + \sum_{i=2}^{\infty} P(\emptyset)$$

il che implica $\sum_{i=2}^{\infty} P(\emptyset) = 0$ e quindi

$$P(\emptyset) = 0$$

perciò l'evento impossibile ha probabilità 0 di verificarsi.

Proposizione. (Proprietà di additività finita) Per ogni collezione finita $A_1, A_2, \ldots, A_n$ di eventi a due a due incompatibili (ossia tali che $A_i \cap A_j = \emptyset$ se $i \neq j$),

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

($\star\star$) **Dimostrazione.** Consideriamo la successione $A_1, A_2, \ldots, A_n, \emptyset, \emptyset, \ldots$; essendo gli eventi di tale successione a due a due incompatibili, dall'Assioma 3 segue che

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\left(\bigcup_{i=1}^{n} A_i\right) \cup \emptyset \cup \emptyset \cup \ldots\right) = \sum_{i=1}^{n} P(A_i) + P(\emptyset) + P(\emptyset) + \ldots$$

Ricordando che $P(\emptyset) = 0$, si ottiene

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

da cui segue la tesi.

Esempio. Nell'esperimento del lancio di una moneta si ha $S = \{t, c\}$; occorre fissare due numeri reali, p_1 e p_2, in modo che

$$P(\{t\}) = p_1, \qquad P(\{c\}) = p_2.$$

Dall'Assioma 1 segue

$$0 \leq p_1 \leq 1, \qquad 0 \leq p_2 \leq 1.$$

Dall'Assioma 2 e dalla proprietà di additività finita si ha

$$1 = P(S) = P(\{t\} \cup \{c\}) = P(\{t\}) + P(\{c\}) = p_1 + p_2.$$

Ad esempio, se testa e croce si verificano con uguale probabilità avremo

$$P(\{t\}) = \frac{1}{2}, \qquad P(\{c\}) = \frac{1}{2};$$

se invece la moneta è truccata in modo da mostrare testa con probabilità doppia rispetto a quella di croce, avremo

$$P(\{t\}) = \frac{2}{3}, \qquad P(\{c\}) = \frac{1}{3}.$$

Esempio. Nel lancio di un dado, con $S = \{1, 2, 3, 4, 5, 6\}$, se supponiamo che i sei risultati siano equiprobabili, allora deve essere

$$P(\{i\}) = \frac{1}{6} \quad (i = 1, 2, 3, 4, 5, 6), \qquad \sum_{i=1}^{6} P(\{i\}) = 1.$$

Dalla proprietà di additività finita segue la probabilità di ottenere un numero pari:

$$P(\{2, 4, 6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}.$$

Più in generale, se il dado ha n facce, con $S = \{1, 2, \ldots, n\}$, e si suppone che le n facce siano equiprobabili, allora avremo

$$P(\{1\}) = P(\{2\}) = \ldots = P(\{n\}) = \frac{1}{n},$$

e la probabilità che in un lancio del dado si ottenga un numero compreso tra 1 e k è

$$P(\{1, 2, \ldots, k\}) = P(\{1\}) + P(\{2\}) + \ldots + P(\{k\}) = \sum_{i=1}^{k} \frac{1}{n} = \frac{k}{n}, \qquad (1 \leq k \leq n).$$

Ciò suggerisce che se i risultati sono equiprobabili, allora la probabilità di un evento è data dal rapporto della sua cardinalità diviso la cardinalità dello spazio campionario.

Esempio. Nel gioco della ruota della fortuna viene fatto ruotare a caso una ruota costituita da n settori circolari, con la regola che si vincono v_k euro se la ruota si ferma in corrispondenza del settore k-esimo, avente arco di lunghezza ℓ_k, per $k = 1, 2, \ldots, n$.

È ragionevole supporre che la probabilità che la ruota si fermi in corrispondenza del settore k-esimo sia proporzionale alla lunghezza ℓ_k dell'arco, per $k = 1, 2, \ldots, n$. Detto A_k tale evento, allora gli eventi $A_1, A_2, \ldots, A_n$ sono a 2 a 2 incompatibili e necessari, con

$$P(A_k) = c\,\ell_k, \qquad k = 1, 2, \ldots, n,$$

dove c è la costante di proporzionalità. Pertanto, usando la proprietà di additività finita si ha

$$1 = P(S) = P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k) = c \sum_{k=1}^{n} \ell_k \quad \Rightarrow \quad c = \frac{1}{\ell_1 + \ell_2 + \ldots + \ell_n}.$$

Quindi la probabilità di vincere v_k euro è

$$P(A_k) = \frac{\ell_k}{\ell_1 + \ell_2 + \ldots + \ell_n}, \qquad k = 1, 2, \ldots, n.$$

2.5 Alcune semplici proprietà

Proposizione. Per ogni evento A risulta

$$P(\overline{A}) = 1 - P(A)$$

Dimostrazione. Dall'Assioma 2 e dalla proprietà di additività finita, con A e $\overline{A}$ eventi necessari e incompatibili, segue

$$1 = P(S) = P(A \cup \overline{A}) = P(A) + P(\overline{A}),$$

da cui si giunge alla tesi.

Esempio. Nel lancio di 2 monete, con $S = \{cc, ct, tc, tt\}$, se la la probabilità di ottenere 2 volte testa è $P(A) = 1/4$, con $A = \{tt\}$, allora la probabilità di ottenere al più una volta testa deve essere $P(\overline{A}) = 1 - P(A) = 1 - 1/4 = 3/4$.

Esempio. Più in generale, se la probabilità di ottenere n volte croce lanciando n monete è $1/2^n$, allora la probabilità di ottenere almeno una volta testa deve essere $1 - 1/2^n$, essendo tali eventi uno il complementare dell'altro.

Proposizione. Se $A_1, A_2, \ldots, A_n$ sono eventi necessari e a 2 a 2 incompatibili, allora

$$\sum_{i=1}^{n} P(A_i) = 1.$$

Dimostrazione. Poiché $A_1, A_2, \ldots, A_n$ sono eventi a 2 a 2 incompatibili (ossia tali che $A_i \cap A_j = \emptyset$ per $i \neq j$), dalla proprietà di additività finita segue

$$\sum_{i=1}^{n} P(A_i) = P\left(\bigcup_{i=1}^{n} A_i\right).$$

Ricordando che gli eventi sono necessari (ossia tali che $\bigcup_{i=1}^{n} A_i = S$), si ha

$$\sum_{i=1}^{n} P(A_i) = P\left(\bigcup_{i=1}^{n} A_i\right) = P(S) = 1.$$

Notiamo che se gli eventi $A_1, A_2, \ldots, A_n$ sono necessari, ma non sono a 2 a 2 incompatibili, allora la somma delle loro probabilità può essere maggiore di 1.

Esercizio. Gli addetti di un *call center* vengono assegnati in tre sale secondo le seguente regole:
– nella prima sala viene assegnato un numero doppio di addetti rispetto alla seconda,
– nella seconda sala viene assegnato un numero doppio di addetti rispetto alla terza.
Se si sceglie un addetto a caso, qual è la probabilità che venga assegnato alla sala i-esima? ($i = 1, 2, 3$).

Soluzione. Poniamo $A_i = \{\text{un addetto scelto a caso viene assegnato alla sala } i\text{-esima}\}$ ($i = 1, 2, 3$). Tali eventi sono necessari e a 2 a 2 incompatibili, ovvero

$$A_1 \cup A_2 \cup A_3 = S, \qquad A_i \cap A_j = \emptyset \text{ per } i \neq j.$$

Pertanto, si ha $P(A_1) + P(A_2) + P(A_3) = 1$. Inoltre, per le ipotesi viste risulta

$$P(A_1) = 2P(A_2), \qquad P(A_2) = 2P(A_3),$$

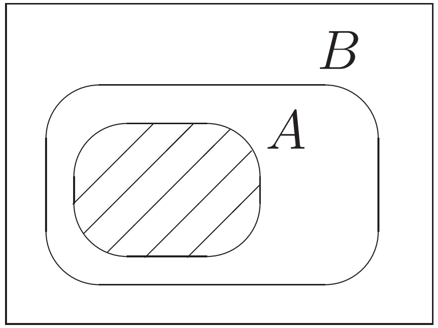
e quindi $P(A_1) = 4P(A_3)$. Ne segue $4P(A_3)+2P(A_3)+P(A_3) = 1$, ossia $7P(A_3) = 1$. In conclusione, si ricava

$$P(A_1) = \frac{4}{7}, \qquad P(A_2) = \frac{2}{7}, \qquad P(A_3) = \frac{1}{7}, \qquad \sum_{i=1}^{3} P(A_i) = 1.$$

Proposizione. Se $A \subset B$, allora

$$P(A) \leq P(B)$$

Dimostrazione. Essendo $A \subset B$, l'evento B può essere espresso come $B = A \cup (\overline{A} \cap B)$, con $A \cap (\overline{A} \cap B) = \emptyset$.



$A \subset B$

Dalla proprietà di additività finita segue
$P(B) = P(A \cup (\overline{A} \cap B)) = P(A) + P(\overline{A} \cap B)$,
da cui si ha $P(B) \geq P(A)$, essendo $P(\overline{A} \cap B) \geq 0$.

Esempio. Nel lancio di un dado, la probabilità che si ottenga 1 è minore della probabilità che si ottenga un numero dispari, poiché $A = \{1\} \subset \{1, 3, 5\} = B$.

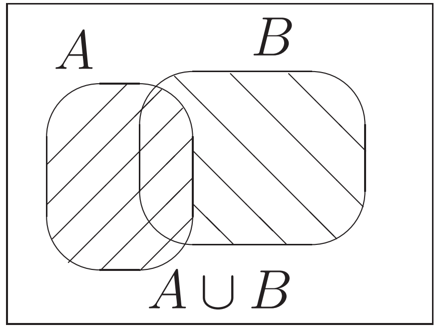
Esempio. Nell'esperimento che consiste del lancio di n monete, la probabilità di $A = \{$esce testa 1 volta$\}$ è minore della probabilità di $B = \{$esce testa almeno 1 volta$\}$, essendo $A \subset B$. Infatti risulta

$$P(A) = \frac{n}{2^n} \leq \frac{2^n - 1}{2^n} = P(B).$$

Proposizione. Dati due eventi $A, B \in \mathcal{F}$ si ha

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

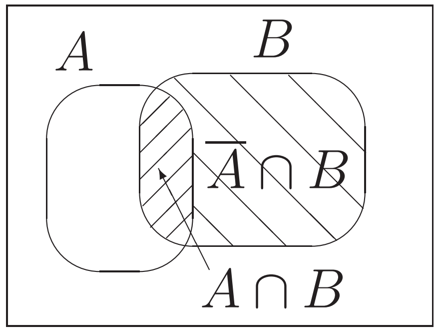
Dimostrazione. Notiamo che $A \cup B$ può essere espresso come unione di due eventi incompatibili A e $\overline{A} \cap B$. Grazie alla proprietà di additività finita otteniamo



$$P(A \cup B) = P(A \cup (\overline{A} \cap B)) = P(A) + P(\overline{A} \cap B).$$

Inoltre, essendo $B = (A \cap B) \cup (\overline{A} \cap B)$, con $A \cap B$ e $\overline{A} \cap B$ eventi incompatibili, applicando nuovamente la proprietà di additività finita abbiamo

$$P(B) = P(A \cap B) + P(\overline{A} \cap B)$$



o, equivalentemente,

$$P(\overline{A} \cap B) = P(B) - P(A \cap B).$$

Sostituendo $P(\overline{A} \cap B)$ nella prima uguaglianza si ha la tesi.

Esercizio. Uno studente deve sostenere due test. Con probabilità 0,5 supererà il primo test; con probabilità 0,4 supererà il secondo test; con probabilità 0,3 li supererà entrambi. Quanto vale la probabilità che non supererà nessuno dei due test?

Soluzione. Sia B_i l'evento che lo studente superi il test i-esimo, $i = 1, 2$. La probabilità che superi almeno un test sarà quindi pari a

$$P(B_1 \cup B_2) = P(B_1) + P(B_2) - P(B_1 \cap B_2) = 0{,}5 + 0{,}4 - 0{,}3 = 0{,}6.$$

La probabilità che lo studente non supererà nessuno dei due test è dunque

$$P(\overline{B_1} \cap \overline{B_2}) = P(\overline{B_1 \cup B_2}) = 1 - P(B_1 \cup B_2) = 1 - 0{,}6 = 0{,}4.$$

Notiamo che la relazione $P(\overline{B_1} \cap \overline{B_2}) = P(\overline{B_1 \cup B_2})$ segue dalla formula di De Morgan.

Esercizio. Dimostrare le seguenti relazioni, per eventi A_1 e A_2 qualsiasi:

$$P(A_1 \cup A_2) \leq P(A_1) + P(A_2), \qquad P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1.$$

Esercizio. Sapendo che A_1 e A_2 sono eventi tali che $P(A_1) = P(A_2) = 0{,}9$ mostrare che $P(A_1 \cap A_2) \geq 0{,}8$ e $P(\overline{A}_1 \cup \overline{A}_2) \leq 0{,}2$.

Esercizio. In un *database* vi sono 450 cartelle di utenti di un sistema di servizio, di cui 135 hanno priorità di $\mathbf{1}^\circ$ tipo e 315 hanno priorità di $\mathbf{2}^\circ$ tipo. Inoltre, 195 dei 450 utenti hanno stipulato un'assicurazione triennale sul servizio. La percentuale di utenti con priorità di $\mathbf{2}^\circ$ tipo e dotati di assicurazione triennale è pari al 20% del totale.
(a) Se si sceglie una cartella a caso, qual è la probabilità che si riferisca ad un utente avente priorità di $\mathbf{2}^\circ$ tipo e privo di assicurazione triennale?
(b) Se si sceglie una cartella a caso, qual è la probabilità che si riferisca ad un utente avente priorità di $\mathbf{1}^\circ$ tipo oppure non dotato di assicurazione triennale?

Soluzione. (a) Per gli eventi $A = \{$la cartella scelta a caso è di un utente assicurato$\}$ e $B = \{$la cartella scelta a caso è di utente avente priorità di $\mathbf{1}^\circ$ tipo$\}$ risulta:

$$P(B) = \frac{135}{450} = 0{,}3 \qquad P(\overline{B}) = \frac{315}{450} = 0{,}7 \qquad P(A) = \frac{195}{450} = 0{,}4\overline{3}.$$

$$P(A \cap \overline{B}) = 0{,}2 \quad \Rightarrow \quad P(\overline{A} \cap \overline{B}) = P(\overline{B}) - P(A \cap \overline{B}) = 0{,}7 - 0{,}2 = 0{,}5$$

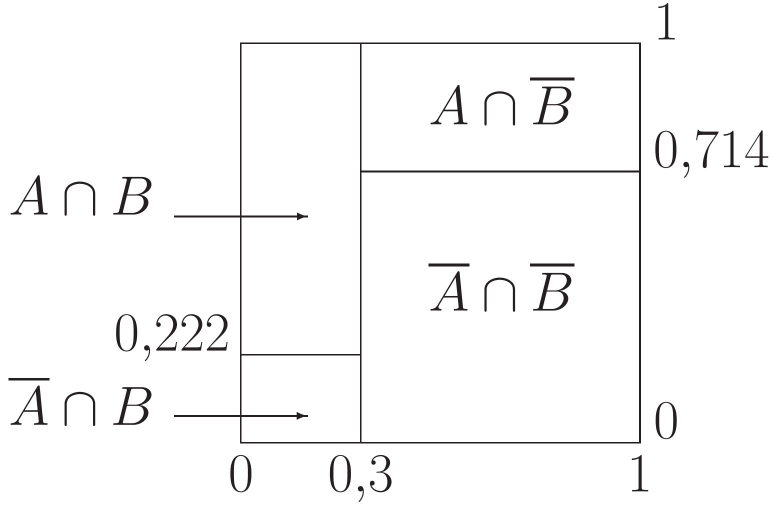
avendo ricordato che $P(\overline{B}) = P(A \cap \overline{B}) + P(\overline{A} \cap \overline{B})$.
(b) Si ha $P(\overline{A} \cup B) = P(\overline{A}) + P(B) - P(\overline{A} \cap B) = 1 - 0{,}4\overline{3} + 0{,}3 - 0{,}0\overline{6} = 0{,}2$ poiché $P(\overline{A}) = 1 - P(A) = 1 - 0{,}4\overline{3}$ e $P(\overline{A} \cap B) = P(\overline{A}) - P(\overline{A} \cap \overline{B}) = 1 - 0{,}4\overline{3} - 0{,}5 = 0{,}0\overline{6}$.

Abbiamo già notato in precedenza che, dati 2 eventi A e B qualsiasi, allora gli eventi $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, $\overline{A} \cap \overline{B}$ costituiscono una collezione di eventi necessari e a 2 a 2 incompatibili. Pertanto risulta conveniente racchiudere le probabilità di tali eventi in una tabella del seguente tipo, anche con riferimento all'esercizio precedente, dove le somme sulle righe e sulle colonne danno i risultati presenti ai margini delle tabelle:

	B	$\overline{B}$	
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	$P(S)$

	B	$\overline{B}$	
A	$0{,}2\overline{3}$	$0{,}20$	$0{,}4\overline{3}$
$\overline{A}$	$0{,}0\overline{6}$	$0{,}50$	$0{,}5\overline{6}$
	$0{,}30$	$0{,}70$	1

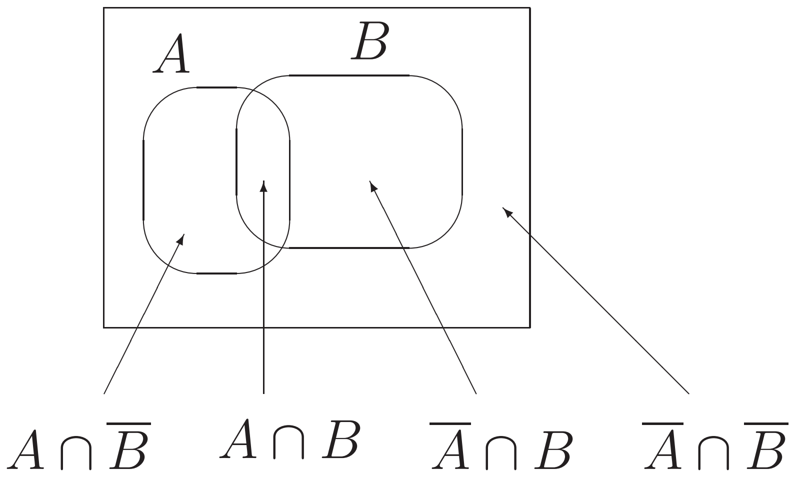
Una rappresentazione geometrica dello spazio campionario S può essere anche data da un quadrato di lato 1, in modo che le aree delle regioni corrispondano alle probabilità degli eventi coinvolti.



Notiamo che sommando i termini presenti nella tabella precedente si ritrovano le seguenti relazioni:

$$P(A \cap B) + P(A \cap \overline{B}) = P(A), \qquad P(\overline{A} \cap B) + P(\overline{A} \cap \overline{B}) = P(\overline{A}),$$

$$P(A \cap B) + P(\overline{A} \cap B) = P(B), \qquad P(A \cap \overline{B}) + P(\overline{A} \cap \overline{B}) = P(\overline{B}),$$



Risulta anche

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A \cap \overline{B}) + P(A \cap B) + P(\overline{A} \cap B) \end{aligned}$$

Proposizione. Dati tre eventi $A, B, C \in \mathcal{F}$ si ha

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

($\star\star$) Dimostrazione. Ricordando che $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$, si ha

$$P(A \cup B \cup C) = P((A \cup B) \cup C) = P(A \cup B) + P(C) - P((A \cup B) \cap C),$$

e ancora

$$P(A \cup B \cup C) = P(A) + P(B) - P(A \cap B) + P(C) - P((A \cup B) \cap C).$$

Per la legge distributiva si ha

$$P(A \cup B \cup C) = P(A) + P(B) - P(A \cap B) + P(C) - P((A \cap C) \cup (B \cap C)),$$

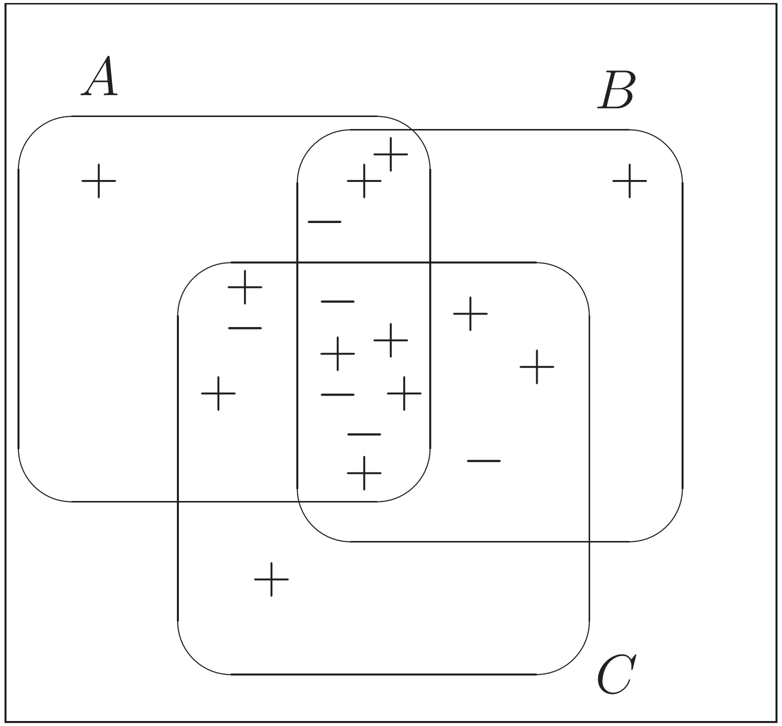
da cui segue

$$P(A \cup B \cup C) = P(A) + P(B) - P(A \cap B) + P(C) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

ossia la tesi.

La seguente figura illustra l'interpretazione grafica della formula

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$



Proposizione. Principio di inclusione/esclusione

La probabilità dell'unione di n eventi $A_1, A_2, \ldots, A_n$ può esprimersi al seguente modo:

per $n = 2$: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$;

per $n = 3$:
$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) \\ &\quad - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ &\quad + P(A_1 \cap A_2 \cap A_3); \end{aligned}$$

in generale:

$$\begin{aligned} P(A_1 \cup A_2 \cup \ldots \cup A_n) \ &= \ \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \\ &- \sum_{i<j<k<r} P(A_i \cap A_j \cap A_k \cap A_r) \\ &+ \ \ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n). \end{aligned}$$

La seguente formula esprime in modo compatto il principio di inclusione/esclusione:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1<i_2<\ldots<i_r} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r}).$$

Il numero di termini a 2° membro è $\displaystyle\sum_{r=1}^{n} \binom{n}{r} = 2^n - 1$ in quanto la sommatoria $\displaystyle\sum_{i_1<i_2<\ldots<i_r} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r})$ è calcolata per tutti gli $\binom{n}{r}$ possibili modi di scegliere gli r indici $i_1 < i_2 < \ldots < i_r$ dall'insieme $\{1, 2, \ldots, n\}$.

Infatti, i termini a 2° membro sono

$2^n - 1 = 3$ per $n = 2$:
$P(A_1) + P(A_2) - P(A_1 \cap A_2)$;

$2^n - 1 = 7$ per $n = 3$:
$P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$.

2.6 Spazi campionari con esiti equiprobabili

In certi esperimenti, per motivi di simmetria, è naturale assumere che gli esiti di uno spazio campionario finito $S = \{1, 2, \ldots, N\}$ siano equiprobabili. Allora si ipotizza che

$$P(\{1\}) = P(\{2\}) = \ldots = P(\{N\})$$

il che implica

$$P(\{i\}) = \frac{1}{N}, \qquad i = 1, 2, \ldots, N,$$

essendo

$$1 = P(S) = P(\{1\} \cup \{2\} \cup \ldots \cup \{N\}) = P(\{1\}) + P(\{2\}) + \ldots + P(\{N\}).$$

Quindi, per la proprietà di additività avremo che per ogni evento A risulta

$$P(A) = \frac{|A|}{|S|} = \frac{\text{numero di elementi di } A}{\text{numero di elementi di } S} \qquad \text{(definizione classica di probabilità)}$$

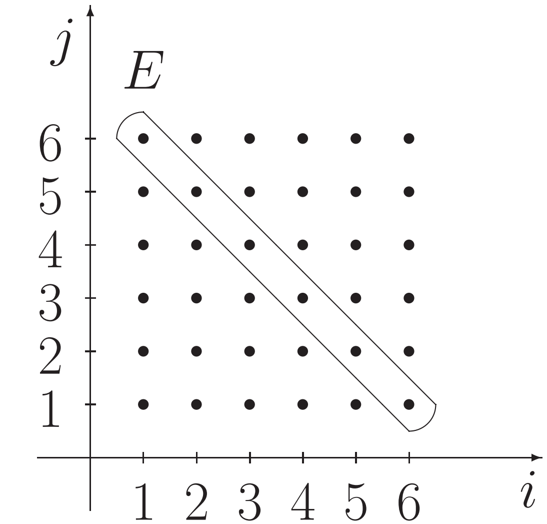
Se assumiamo che gli esiti di un esperimento siano equiprobabili, allora la probabilità di ogni evento A è uguale alla proporzione degli esiti dello spazio campionario contenuti in A (visto come rapporto di casi favorevoli su casi possibili).

Esempio. Se si lanciano 2 dadi, qual è la probabilità che la somma dei risultati sia uguale a 7?

Soluzione. Assumendo che i 36 possibili esiti siano equiprobabili, poiché ci sono 6 possibili esiti che danno come somma 7, con

$E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$,

la probabilità richiesta è $P(E) = \frac{|E|}{|S|} = \frac{6}{36} = \frac{1}{6} = 0{,}1667.$



Esempio. Se estraiamo 3 biglie a caso da un'urna che contiene 6 biglie bianche e 5 nere, qual è la probabilità che una sia bianca e le altre due nere?

Soluzione. Poiché gli esiti dello spazio campionario sono equiprobabili, e sono $|S| = (11)_3 = 11 \cdot 10 \cdot 9 = 990$, tenendo conto dell'ordine di estrazione la probabilità richiesta è espressa mediante numeri di disposizioni:

$$\frac{(6 \cdot 5 \cdot 4) + (5 \cdot 6 \cdot 4) + (5 \cdot 4 \cdot 6)}{11 \cdot 10 \cdot 9} = \frac{3 \cdot 120}{990} = \frac{4}{11} = 0{,}3636.$$

Esempio. Se estraiamo 3 biglie a caso da un'urna che contiene 6 biglie bianche e 5 nere, qual è la probabilità che una sia bianca e le altre due nere?

Soluzione. Se non teniamo conto dell'ordine di estrazione, la probabilità richiesta è espressa mediante numeri di combinazioni:

$$\frac{\binom{6}{1}\binom{5}{2}}{\binom{11}{3}} = \frac{6 \cdot 10}{165} = \frac{4}{11} = 0{,}3636.$$

Esempio. Una commissione di 5 persone viene estratta da un gruppo composto di 6 uomini e 9 donne. Se la selezione avviene in modo casuale, qual è la probabilità che la commissione consti di 3 uomini e 2 donne oppure di 2 uomini e 3 donne?

Soluzione. Ognuna delle $\binom{15}{5}$ combinazioni è estratta in modo equiprobabile, quindi

$$\frac{\binom{6}{3}\binom{9}{2}}{\binom{15}{5}} + \frac{\binom{6}{2}\binom{9}{3}}{\binom{15}{5}} = \frac{20 \cdot 36}{3003} + \frac{15 \cdot 84}{3003} = \frac{240}{1001} + \frac{420}{1001} = \frac{660}{1001} = 0{,}6593.$$

Esempio. Un'urna contiene n biglie, di cui una è bianca. Se estraiamo k biglie una alla volta, in modo tale che a ogni estrazione la probabilità di estrarre una qualunque delle biglie rimanenti sia la stessa, qual è la probabilità che la biglia bianca sia estratta?

Soluzione. Ognuno dei $\binom{n}{k}$ possibili insiemi è estratto in modo equiprobabile, quindi ponendo $B = \{\text{si estrae la biglia bianca}\}$ si ha

$$P(B) = \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\dfrac{(n-1)!}{(k-1)!\,(n-k)!}}{\dfrac{n!}{k!\,(n-k)!}} = \frac{\dfrac{(n-1)!}{(k-1)!\,(n-k)!}}{\dfrac{n\,(n-1)!}{k\,(k-1)!\,(n-k)!}} = \frac{k}{n}$$

Alternativamente, sia $A_i = \{\text{la biglia bianca si estrae nell'}i\text{-esima estrazione}\}$, $i = 1, 2, \ldots, k$. Risulta $P(A_i) = 1/n$ perché ognuna delle n biglie ha uguale probabilità di essere scelta all'i-esima estrazione. Poiché gli eventi A_i sono incompatibili, usando la proprietà di additività finita si ha quindi

$$P(B) = P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i) = \sum_{i=1}^{k} \frac{1}{n} = \frac{k}{n}$$

Esempio. Da un lotto di $N = 200$ pezzi, costituito da 40 pezzi difettosi e 160 buoni, si estrae a caso un campione di $n = 10$ pezzi.
(a) Qual è la probabilità che il campione estratto abbia la stessa frazione di pezzi difettosi del lotto originario?
(b) Qual è la probabilità che il campione estratto non contenga pezzi difettosi?
(c) Cosa cambia se si raddoppia il numero di pezzi estratti?

Soluzione. (a) Ognuno degli $\binom{N}{n}$ possibili insiemi è estratto in modo equiprobabile. La frazione di pezzi difettosi nel lotto è $p = 40/200 = 1/5 = 20\%$. Quindi, per l'evento $A = \{\text{il } 20\% \text{ dei pezzi estratti è difettoso}\}$ risulta

$$P(A) = \frac{\binom{40}{2}\binom{160}{8}}{\binom{200}{10}} = 0{,}3098$$

(in questo caso l'evento A corrisponde ad avere 2 pezzi difettosi e 8 pezzi buoni) cosicché è più probabile che un campione estratto a caso da quel lotto non riproduca esattamente le caratteristiche dell'intero lotto.

(b) La probabilità che il campione non contenga pezzi difettosi è

$$P(E) = \frac{\binom{40}{0}\binom{160}{10}}{\binom{200}{10}} = 0{,}1013.$$

(c) Raddoppiando il numero di elementi del campione diminuisce la probabilità che questo riproduca la stessa frazione ($p = 20\%$) di pezzi difettosi, essendo ora

$$P(A) = \frac{\binom{40}{4}\binom{160}{16}}{\binom{200}{20}} = 0{,}2299$$

(in questo caso l'evento A corrisponde ad avere 4 pezzi difettosi e 16 pezzi buoni) però diminuisce sensibilmente la probabilità che il campione estratto non contenga pezzi difettosi, poiché

$$P(E) = \frac{\binom{40}{0}\binom{160}{20}}{\binom{200}{20}} = 0{,}0089.$$

($\star\star$) Esempio. Supponiamo che $n+m$ biglie, di cui n siano rosse e m blu, vengano disposte in fila in modo casuale, così che ognuna delle $(n+m)!$ possibili disposizioni sia equiprobabile. Se siamo interessati solo alla successione dei colori delle biglie in fila, si provi che tutti i possibili risultati rimangono equiprobabili.

Soluzione. Consideriamo ognuna delle $(n+m)!$ possibili disposizioni ordinate delle biglie: se permutiamo tra loro le biglie rosse e facciamo lo stesso con le blu, la successione dei colori non cambia. Ne segue che ogni successione dei colori corrisponde a $n!\,m!$ differenti disposizioni ordinate delle $n+m$ biglie, così che ogni distribuzione di colori ha probabilità di verificarsi

$$\frac{n!\,m!}{(n+m)!} = \frac{1}{\binom{n+m}{n}}.$$

Ad esempio se $n = m = 2$, delle $4! = 24$ possibili disposizioni ordinate delle biglie, ci saranno $2!\,2! = 4$ distribuzioni che danno la stessa distribuzione di colori; ad esempio

$$r_1, b_1, r_2, b_2, \qquad r_1, b_2, r_2, b_1, \qquad r_2, b_1, r_1, b_2, \qquad r_2, b_2, r_1, b_1.$$

Le 6 distribuzioni di colori distinte saranno: $bbrr$, $brbr$, $brrb$, $rbbr$, $rbrb$, $rrbb$.

Esempio. In un gruppo di n persone, qual è la probabilità che almeno due persone compiano gli anni lo stesso giorno? (Assumendo che ogni persona abbia la stessa possibilità di compiere gli anni in uno qualsiasi dei 365 giorni di un anno.)

Soluzione. Detto A l'evento che almeno due persone compiono gli anni lo stesso giorno, l'evento complementare $\overline{A}$ si realizza quando le n persone compiono gli anni in giorni tutti diversi. Quindi, usando la definizione classica di probabilità si ha

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{|\overline{A}|}{|S|} = 1 - \frac{(365)_n}{365^n} = 1 - \frac{(365)(364)\cdots(365-n+1)}{365^n}.$$

Ad esempio:

– per $n = 10$: $P(A) = 0{,}117$

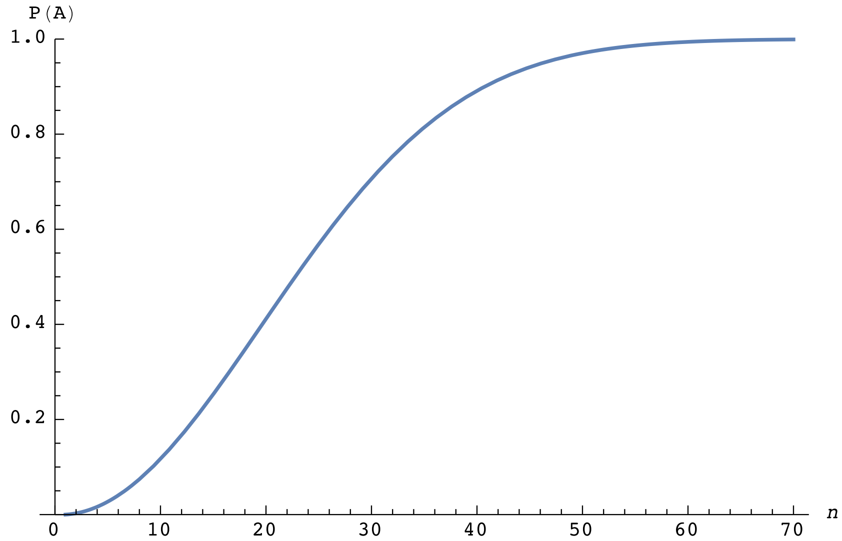
– per $n = 15$: $P(A) = 0{,}253$

– per $n = 23$: $P(A) = 0{,}507$

– per $n = 30$: $P(A) = 0{,}706$

– per $n = 40$: $P(A) = 0{,}891$

– per $n = 50$: $P(A) = 0{,}970$



Esempio. Nel lancio di una moneta ripetuto 3 volte, un giocatore scommette sulla fuoriuscita di testa nei primi 2 lanci, un secondo giocatore scommette sul fatto che i tre lanci diano lo stesso risultato, un terzo giocatore scommette sul fatto che esce almeno una volta testa negli ultimi 2 lanci. (i) Quale giocatore ha probabilità più alta di vincita? (ii) Qual è la probabilità che almeno un giocatore vinca?

Soluzione. Nello spazio campionario vi sono $2^3 = 8$ sequenze di t (testa) e c (croce);
$A = \{\text{vince il primo giocatore}\} = \{tt\omega : \omega \in \{c,t\}\}$; $P(A) = 2/8 = 1/4$;
$B = \{\text{vince il secondo giocatore}\} = \{\omega\omega\omega : \omega \in \{c,t\}\}$; $P(B) = 2/8 = 1/4$;
$C = \{\text{vince il terzo giocatore}\} = \{\omega ct, \omega tc, \omega tt : \omega \in \{c,t\}\}$; $P(C) = 6/8 = 3/4$.
(i) Il terzo giocatore ha probabilità più alta di vincita.
(ii) $P(\{\text{almeno un giocatore vince}\}) = P(A \cup B \cup C)$
$= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

$$= \frac{1}{4} + \frac{1}{4} + \frac{3}{4} - \frac{1}{8} - \frac{1}{4} - \frac{1}{8} + \frac{1}{8} = \frac{7}{8}$$

essendo $A \cap B = \{ttt\}$; $A \cap C = \{ttc, ttt\}$; $B \cap C = \{ttt\}$; $A \cap B \cap C = \{ttt\}$.
Notiamo che $P(\overline{A \cup B \cup C}) = P(\overline{A} \cap \overline{B} \cap \overline{C}) = P(\{tcc\}) = \frac{1}{8}$.

Si può giungere alla soluzione anche rappresentando lo spazio campionario e gli eventi dell'esempio mediante una tabella (dove il simbolo $\star$ mostra che la sequenza di risultati della riga appartiene all'evento riferito alla colonna) e tenendo conto che ogni sequenza dello spazio campionario S ha probabilità 1/8:

S	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$	$A \cup B \cup C$
ccc		$\star$						$\star$
cct			$\star$					$\star$
ctc			$\star$					$\star$
ctt			$\star$					$\star$
tcc								
tct			$\star$					$\star$
ttc	$\star$		$\star$		$\star$			$\star$
ttt	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$
$P(\cdot)$	1/4	1/4	3/4	1/8	1/4	1/8	1/8	7/8

Esempio. Nel gioco del lotto si estraggono a caso 5 numeri da un'urna contenente 90 numeri compresi tra 1 e 90. (Le cinquine sono equiprobabili). Se si scelgono k numeri distinti compresi tra 1 e 90 qual è la probabilità che questi saranno tra i 5 estratti?

Soluzione. Poiché l'ordine di estrazione non è rilevante, lo spazio campionario è l'insieme delle $\binom{90}{5}$ combinazioni semplici di 90 oggetti in 5 classi.

L'evento $A_k = \{\text{i } k \text{ numeri scelti sono tra i 5 estratti}\}$ è costituito dalle sequenze di S che presentano all'interno i k numeri scelti; pertanto la sua cardinalità è pari al numero di modi di prendere tutti i k numeri scelti, che è $\binom{k}{k} = 1$, per il numero $\binom{90-k}{5-k}$ di modi di prendere i restanti $5-k$ numeri dai $90-k$ numeri ulteriori. Si ha quindi:

$P(A_k)$	$P(A_1)$	$P(A_2)$	$P(A_3)$	$P(A_4)$	$P(A_5)$
$\dfrac{\binom{90-k}{5-k}}{\binom{90}{5}}$	$\dfrac{\binom{89}{4}}{\binom{90}{5}} = \dfrac{1}{18}$	$\dfrac{\binom{88}{3}}{\binom{90}{5}} = \dfrac{2}{801}$	$\dfrac{\binom{87}{2}}{\binom{90}{5}} = \dfrac{1}{11.748}$	$\dfrac{\binom{86}{1}}{\binom{90}{5}} = \dfrac{1}{511.038}$	$\dfrac{\binom{85}{0}}{\binom{90}{5}} = \dfrac{1}{43.949.268}$

Esempio. Da un'urna contenente 3 biglie numerate da 1 a 3 si estraggono casualmente in sequenza le 3 biglie. Diciamo che si ha una concordanza nell'i-esima estrazione se in tale estrazione viene estratta la biglia avente numero i, per $i = 1, 2, 3$. Calcolare la probabilità degli eventi $C_k = \{\text{si hanno in totale } k \text{ concordanze}\}$, per $k = 0, 1, 2, 3$.

Soluzione. I possibili risultati dell'esperimento sono le $3! = 6$ permutazioni dei numeri $1, 2, 3$. Poichè gli esiti sono equiprobabili, ognuna di tali sequenze si verifica con probabilità $1/6$. Dalla seguente tabella si ottengono le probabilità richieste:

esito delle estrazioni	numero di concordanze
1 2 3	3
1 3 2	1
2 1 **3**	1
2 3 1	0
3 1 2	0
3 **2** 1	1

$$P(C_0) = \frac{2}{6} \qquad P(C_1) = \frac{3}{6}$$

$$P(C_2) = 0 \qquad P(C_3) = \frac{1}{6}$$

(C_0, C_1, C_2, C_3 sono necessari e a 2 a 2 incompatibili.)

La probabilità di avere almeno una concordanza è

$$P(C_1 \cup C_2 \cup C_3) = \tfrac{3}{6} + 0 + \tfrac{1}{6} = \tfrac{2}{3} \quad = \quad P(\overline{C}_0) = 1 - P(C_0) = 1 - \tfrac{2}{6} = \tfrac{4}{6} = \tfrac{2}{3}.$$

Esempio. Da un'urna contenente 4 biglie numerate da 1 a 4 si estraggono casualmente in sequenza le 4 biglie. Diciamo che si ha una concordanza nell'i-esima estrazione se in tale estrazione viene estratta la biglia avente numero i, per $i = 1, 2, 3.4$. Calcolare la probabilità (i) di avere concordanza nell'i-esima estrazione, per $i = 1, 2, 3, 4$; (ii) di almeno una concordanza; (iii) di non avere concordanze.

Soluzione. I possibili risultati dell'esperimento sono le $4! = 4 \times 3 \times 2 \times 1 = 24$ sequenze ottenute permutando i numeri $1, 2, 3, 4$, ciascuna avente probabilità $1/24$.

(i) Poniamo $E_i = \{\text{si ha concordanza nell'}i\text{-esima estrazione}\}$ per $i = 1, 2, 3, 4$. Ad esempio E_1 è costituito dalle sequenze del tipo $(1, \omega_2, \omega_3, \omega_4)$, per ogni permutazione $(\omega_2, \omega_3, \omega_4)$ delle cifre $2, 3, 4$. Quindi l'evento E_1 ha cardinalità $3! = 6$ e pertanto

$$P(E_1) = \frac{|E_1|}{|S|} = \frac{6}{24} = \frac{1}{4}.$$

Analogamente, risulta

$$P(E_2) = P(E_3) = P(E_4) = \frac{1}{4}.$$

Infatti, ad esempio, E_2 è costituito dalle $3! = 6$ sequenze del tipo $(\omega_1, 2, \omega_3, \omega_4)$, per ogni permutazione $(\omega_1, \omega_3, \omega_4)$ delle cifre $1, 3, 4$.

(ii) La probabilità di avere almeno una concordanza è

$$P\left(\bigcup_{i=1}^{4} E_i\right) = \sum_{i=1}^{4} P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - P(E_1 \cap E_2 \cap E_3 \cap E_4).$$

Notiamo che $E_1 \cap E_2$ è costituito dalle sequenze del tipo $(1, 2, \omega_3, \omega_4)$, per ogni permutazione (ω_3, ω_4) delle cifre $3, 4$. Quindi

$$P(E_1 \cap E_2) = \frac{|E_1 \cap E_2|}{|S|} = \frac{2}{24} = \frac{1}{12},$$

e analogamente $P(E_i \cap E_j) = \frac{1}{12}$ per tutte le $\binom{4}{2} = \frac{4 \cdot 3}{2} = 6$ coppie (i, j) per cui $i < j$:

$$\sum_{i<j} P(E_i \cap E_j) = 6 \times \frac{1}{12} = \frac{1}{2}.$$

Inoltre, $E_1 \cap E_2 \cap E_3$ è costituito dalla sequenza $(1, 2, 3, 4)$, e quindi

$$\sum_{i<j<k} P(E_i \cap E_j \cap E_k) = P(E_1E_2E_3) + P(E_1E_2E_4) + P(E_1E_3E_4) + P(E_2E_3E_4)$$

$$= 4 \times \frac{1}{24} = \frac{1}{6}.$$

Inoltre $E_1 \cap E_2 \cap E_3 \cap E_4$ è costituito da un'unica sequenza, e pertanto

$$P(E_1 \cap E_2 \cap E_3 \cap E_4) = \frac{1}{24}.$$

Quindi, la probabilità di avere almeno una concordanza è

$$P\left(\bigcup_{i=1}^{4} E_i\right) = \sum_{i=1}^{4} P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - P(E_1 \cap E_2 \cap E_3 \cap E_4)$$

$$= \sum_{i=1}^{4} \frac{1}{4} - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} = \frac{15}{24} = 0{,}625.$$

(iii) Di conseguenza, la probabilità di non avere concordanze è

$$P\left(\bigcap_{i=1}^{4} \overline{E}_i\right) = P\left(\overline{\bigcup_{i=1}^{4} E_i}\right) = 1 - P\left(\bigcup_{i=1}^{4} E_i\right) = 1 - \frac{15}{24} = \frac{9}{24} = 0{,}375$$

avendo fatto uno della formula di De Morgan.

($\star\star$) Esempio. Il problema delle concordanze. Da un'urna che contiene biglie numerate da 1 a n si estraggono casualmente in sequenza le n biglie. Diciamo che si ha una concordanza nell'i-esima estrazione se in tale estrazione viene estratta la biglia avente numero i, per $i = 1, 2, \ldots, n$. Calcolare la probabilità che non si abbia concordanza in nessuna estrazione.

Soluzione. Per $i = 1, 2, \ldots, n$ sia $E_i = \{\text{si ha concordanza nell'}i\text{-esima estrazione}\}$. Pertanto, la probabilità che si abbia almeno una concordanza è

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1<i_2<\ldots<i_r} P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_r}).$$

Rappresentiamo l'esito dell'esperimento come un vettore $(\omega_1, \omega_2, \ldots, \omega_n)$, dove ω_i denota il numero estratto all'estrazione i-esima. Lo spazio campionario è quindi l'insieme delle $n!$ permutazioni

$$S = \{(\omega_1, \omega_2, \ldots, \omega_n) \colon \omega_i \in \{1, 2, \ldots, n\},\ \omega_i \neq \omega_j \text{ per } i \neq j\}.$$

L'evento E_i si verifica se nell'estrazione i-esima si ha concordanza (e non è escluso che si possa avere concordanza in altre estrazioni). Ciò può accadere in $(n-1)!$ modi, perché nell'estrazione i-esima l'esito deve essere il numero i (per avere concordanza), mentre nelle altre estrazioni ci sono $(n-1)\,(n-2)\,\ldots\,2\cdot 1$ modi per individuare gli esiti dell'esperimento. Si ha pertanto

$$P(E_i)=\frac{(n-1)!}{n!}=\frac{1}{n}\quad \text{e analogamente}\quad P(E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_r})=\frac{(n-r)!}{n!}.$$

Poiché vi sono $\binom{n}{r}$ termini in $\sum\limits_{i_1<i_2<\ldots<i_r} P(E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_r})$, risulta quindi

$$\sum_{i_1<i_2<\ldots<i_r} P(E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_r})=\binom{n}{r}\frac{(n-r)!}{n!}=\frac{n!}{r!\,(n-r)!}\cdot\frac{(n-r)!}{n!}=\frac{1}{r!}$$

da cui:

$$P\left(\bigcup_{i=1}^{n}E_i\right)=\sum_{r=1}^{n}(-1)^{r+1}\sum_{i_1<i_2<\ldots<i_r} P(E_{i_1}\cap E_{i_2}\cap\ldots\cap E_{i_r})=\sum_{r=1}^{n}(-1)^{r+1}\frac{1}{r!}.$$

La probabilità che non si abbia concordanza in nessuna estrazione è quindi

$$p_n = P\left(\overline{\bigcup_{i=1}^{n} E_i}\right) = 1 - P\left(\bigcup_{i=1}^{n} E_i\right) = 1 - \sum_{r=1}^{n} \frac{(-1)^{r+1}}{r!} = 1 + \sum_{r=1}^{n} \frac{(-1)^r}{r!} = \sum_{r=0}^{n} \frac{(-1)^r}{r!}$$

Notiamo che per n grande la probabilità

$$p_n = \sum_{r=0}^{n} \frac{(-1)^r}{r!}$$

può essere approssimata da

$$e^{-1} = \sum_{r=0}^{+\infty} \frac{(-1)^r}{r!} \approx 0{,}367879$$

(Ricordiamo che $\boxed{e^x = \sum_{r=0}^{+\infty} \frac{x^r}{r!}}$ per $x \in \mathbb{R}$)

n	p_n
1	0
2	0,5
3	0,333333
4	0,375
5	0,366667
6	0,368056
7	0,367857
8	0,367882
9	0,367879
10	0,367879

Esempio. Nell'esperimento che consiste nel lanciare n volte una moneta non truccata, indicando con t la fuoriuscita di testa e con c di croce, lo spazio campionario è

$$S = \{\underline{\omega} := (\omega_1, \omega_2, \ldots, \omega_n) \colon \omega_i \in \{t, c\} \text{ per } 1 \le i \le n\},$$

e le sue 2^n sequenze sono equiprobabili. Per $1 \le k \le n$, calcolare le probabilità di
$T_k = \{\text{al } k\text{-esimo lancio esce testa}\} = \{\underline{\omega} \colon \omega_k = t; \omega_i \in \{t, c\} \text{ per } i \ne k\}$
$A_k = \{\text{nei primi } k \text{ lanci esce testa}\} = \{\underline{\omega} \colon \omega_1 = \ldots = \omega_k = t; \omega_i \in \{t, c\} \text{ per } i > k\}$
$A'_k = \{\text{nei primi } k \text{ lanci esce croce}\} = \{\underline{\omega} \colon \omega_1 = \ldots = \omega_k = c; \omega_i \in \{t, c\} \text{ per } i > k\}$
$B_k = \{\text{esce testa } k \text{ volte}\} = \{\underline{\omega} \colon \sum_{i=1}^{n} I[\omega_i = t] = k\}$ $(I[A] = 1 \text{ se } A \text{ è vera, } 0 \text{ senò})$.

Soluzione. Trattandosi di uno spazio campionario con esiti equiprobabili, si ha

$$P(T_k) = \frac{2^{n-1}}{2^n} = \frac{1}{2}, \qquad P(A_k) = P(A'_k) = \frac{2^{n-k}}{2^n} = \frac{1}{2^k}, \qquad P(B_k) = \frac{\binom{n}{k}}{2^n}.$$

Esercizio. Riferendosi all'esempio precedente, mostrare che gli eventi $B_0, B_1, \ldots, B_n$ sono necessari e a 2 a 2 incompatibili, ed inoltre che

$$P(T_1 \cap A_k) = \frac{1}{2^k}, \qquad P(T_1 \cup A_k) = \frac{1}{2},$$

$$P(T_1 \cap B_k) = \frac{\binom{n-1}{k-1}}{2^n}, \qquad P(T_1 \cup B_k) = \frac{1}{2} + \frac{\binom{n-1}{k}}{2^n},$$

$$P(A_k \cap A'_k) = 0, \qquad P(A_k \cup A'_k) = \frac{1}{2^{k-1}},$$

$$P(A_k \cap B_k) = \frac{1}{2^n}, \qquad P(A_k \cup B_k) = \frac{1}{2^k} + \frac{\binom{n}{k}}{2^n} - \frac{1}{2^n},$$

$$P(A'_{k-1} \cup T_k) = \frac{1}{2} + \frac{1}{2^k}, \qquad P(A'_{k-1} \cap T_k) = \frac{1}{2^k}.$$

2.7 La probabilità come funzione di insieme continua

Una successione di eventi $\{E_n, n \geq 1\}$ è detta *successione crescente* se

$$E_1 \subset E_2 \subset \ldots \subset E_n \subset E_{n+1} \subset \ldots$$

mentre è detta *successione decrescente* se

$$E_1 \supset E_2 \supset \ldots \supset E_n \supset E_{n+1} \supset \ldots$$

Se $\{E_n, n \geq 1\}$ è una successione crescente di eventi, allora introduciamo il limite della successione, che denotiamo con $\lim_{n\to\infty} E_n$, come l'evento così definito:

$$\lim_{n\to\infty} E_n = \bigcup_{i=1}^{\infty} E_i$$

Similmente, se $\{E_n, n \geq 1\}$ è una successione decrescente di eventi, definiamo $\lim_{n\to\infty} E_n$ come

$$\lim_{n\to\infty} E_n = \bigcap_{i=1}^{\infty} E_i$$

Proposizione. (Proprietà di continuità della probabilità) Se $\{E_n, n \geq 1\}$ è una successione di eventi, crescente o decrescente, allora

$$\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n).$$

Esempio. Nel lancio di una moneta non truccata, ripetuto indefinitamente, studiare le seguenti successioni di eventi e i loro limiti, se esistenti (per $n = 1, 2, \ldots$)

(i) $A_n = \{$esce testa almeno una volta nei primi n lanci$\}$,

(ii) $B_n = \{$nei primi n lanci esce sempre testa$\}$,

(iii) $C_n = \{$esce una sola volta testa nei primi n lanci$\}$.

Soluzione. (i) È facile vedere che $\{A_n, n \geq 1\}$ è una successione crescente di eventi, essendo $A_1 \subset A_2 \subset A_3 \subset \ldots$, e risulta $P(A_n) = 1 - \frac{1}{2^n}$. Quindi si ha

$$\lim_{n\to\infty} A_n = \bigcup_{i=1}^{\infty} A_i = A = \{\text{esce almeno una volta testa nella sequenza dei lanci}\},$$

$$P(A) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} \left(1 - \frac{1}{2^n}\right) = 1.$$

(ii) La successione $\{B_n, n \geq 1\}$ è decrescente, poiché $B_1 \supset B_2 \supset B_3 \supset \ldots$, con $P(B_n) = \frac{1}{2^n}$. Quindi si ha:

$$\lim_{n\to\infty} B_n = \bigcap_{i=1}^{\infty} B_i = B = \{\text{esce testa in tutti i lanci}\},$$

$$P(B) = P(\lim_{n\to\infty} B_n) = \lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} \frac{1}{2^n} = 0.$$

(iii) È facile verificare che $\{C_n, n \geq 1\}$ è una successione non monotona. Infatti, se esce una sola volta testa nei primi n lanci non è detto che esca una sola volta testa nei primi $n+1$ lanci, e viceversa. Risulta comunque

$$P(C_n) = \frac{n}{2^n}.$$

(⋆⋆) Esempio. Nel lancio di una moneta non truccata ripetuto indefinitamente calcolare

$$P(E^{\mathrm{p}}) = P(\{\text{esce testa per la prima volta in un lancio pari}\}),$$

$$P(E^{\mathrm{d}}) = P(\{\text{esce testa per la prima volta in un lancio dispari}\}).$$

Soluzione. Consideriamo la successione di eventi

$$F_k = \{\text{esce testa per la prima volta nel lancio } k\text{-esimo}\}, \qquad k \geq 1.$$

Gli eventi $F_1, F_2, \ldots$ sono incompatibili e, per quanto visto in precedenza, si ha

$$\begin{aligned} P(F_1) &= P(T_1) = \frac{1}{2}, \\ P(F_k) &= P(\overline{T_1} \cap \overline{T_2} \cap \cdots \cap \overline{T_{k-1}} \cap T_k) = P(A'_{k-1} \cap T_k) = \frac{1}{2^k}, \qquad k \geq 2. \end{aligned}$$

Notiamo che

$$E^{\mathrm{p}} = \bigcup_{k=1}^{\infty} F_{2k} = \{\text{esce testa per la prima volta in un lancio pari}\},$$

$$E^{\mathrm{d}} = \bigcup_{k=1}^{\infty} F_{2k-1} = \{\text{esce testa per la prima volta in un lancio dispari}\}.$$

Dalla proprietà di additività numerabile e ricordando che $P\left(F_k\right) = \dfrac{1}{2^k}$ segue:

$$P(E^{\mathrm{p}}) = P\left(\bigcup_{k=1}^{+\infty} F_{2k}\right) = \sum_{k=1}^{+\infty} P\left(F_{2k}\right) = \sum_{k=1}^{+\infty} \frac{1}{2^{2k}} = \sum_{k=1}^{+\infty} \left(\frac{1}{4}\right)^k,$$

$$P(E^{\mathrm{d}}) = P\left(\bigcup_{k=1}^{+\infty} F_{2k-1}\right) = \sum_{k=1}^{+\infty} P\left(F_{2k-1}\right) = \sum_{k=1}^{+\infty} \frac{1}{2^{2k-1}} = 2\sum_{k=1}^{+\infty} \left(\frac{1}{4}\right)^k.$$

Poiché $\boxed{\displaystyle\sum_{k=1}^{\infty} x^k = \frac{x}{1-x} \quad (-1 < x < 1)}$ si ha $P(E^{\mathrm{p}}) = \dfrac{\frac{1}{4}}{1-\frac{1}{4}} = \dfrac{1}{3}$ e $P(E^{\mathrm{d}}) = \dfrac{2}{3}$.

È utile sottolineare che sebbene sia

$$P(E^{\mathrm{p}}) + P(E^{\mathrm{d}}) = 1, \qquad E^{\mathrm{p}} \cap E^{\mathrm{d}} = \emptyset,$$

gli eventi E^{p} ed E^{d} non sono tra loro complementari. Infatti introducendo l'evento

$$E_0 = \{\text{non esce mai testa}\},$$

si ha

$$E^{\mathrm{p}} \cup E^{\mathrm{d}} \cup E_0 = S, \qquad \text{con } E^{\mathrm{p}}, E^{\mathrm{d}}, E_0 \text{ incompatibili.}$$

Inoltre, ponendo $A'_k = \{\text{nei primi } k \text{ lanci esce sempre croce}\}$ per $k \geq 1$, la successione $\{A'_1, A'_2, \ldots\}$ è decrescente ed ha come limite l'evento

$$\lim_{k\to\infty} A'_k = \bigcap_{k=1}^{\infty} A'_k = E_0.$$

Pertanto, per la continuità della probabilità, e notando che $P(A'_k) = \dfrac{1}{2^k}$, si ha

$$P(E_0) = P\left(\lim_{k\to\infty} A'_k\right) = \lim_{k\to\infty} P(A'_k) = \lim_{k\to\infty} \frac{1}{2^k} = 0.$$

Esercizi per casa

2.1) Nel lancio di due dadi non truccati, sia $A = \{$il primo dado dà esito 6$\}$ e $B = \{$il secondo dado dà esito diverso da 6$\}$.
(i) Definire lo spazio campionario e determinarne la cardinalità.
(ii) Calcolare $P(A)$, $P(B)$, $P(A \cup B)$.

2.2) Un esperimento consiste nel lanciare n volte una moneta equa. Sia $A = \{$negli n lanci si ottiene 2 volte testa$\}$ e $B = \{$negli n lanci si ottiene sempre croce oppure una sola volta testa$\}$.
(i) Definire lo spazio campionario e determinarne la cardinalità.
(ii) Calcolare $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup \overline{B})$.

2.3) Un esperimento consiste nell'estrarre a caso k biglie da un'urna contenente n biglie numerate da 1 ad n. Sia $A = \{$la biglia numero 1 viene estratta$\}$ e $B = \{$le biglie numerate da 1 a k vengono estratte$\}$. Calcolare $P(A)$, $P(B)$, $P(A \cap B)$ e $P(A \cup B)$.

2.4) Da un elenco di 10 impiegati, di cui 6 laureati e 4 diplomati, se ne selezionano 3 a caso.
(i) Calcolare la probabilità che almeno uno degli impiegati selezionati sia diplomato.
(ii) Calcolare la probabilità che gli impiegati selezionati siano tutti diplomati.
(iii) Calcolare la probabilità che tra i 3 impiegati selezionati vi sia un solo diplomato.

2.5) Da un'urna contenente 12 biglie (di cui 3 bianche, 4 rosse e 5 blu), se ne estraggono 3 (senza reinserimento). Calcolare la probabilità che tra le 3 estratte almeno uno dei 3 colori non sia presente.

2.6) Calcolare le seguenti somme:

$$\text{(i)} \sum_{n=0}^{+\infty} \frac{(-1)^{n+2}}{n!} \qquad \text{(ii)} \sum_{n=1}^{+\infty} \frac{2^{n-2}}{n!} \qquad \text{(iii)} \sum_{n=0}^{+\infty} \left(\frac{2}{3}\right)^{n-1} \qquad \text{(iv)} \sum_{n=1}^{+\infty} 3^{-2n-1}$$

2.7) Un giocatore di poker ha le seguenti carte: $(A, A, A, K, 2)$; calcolare la probabilità di fare full e la probabilità di fare poker nei seguenti casi:
(i) se cambia il 2;
(ii) se cambia il K e il 2.

2.8) Stabilire quale dei seguenti eventi ha probabilità maggiore:
(i) esce 6 almeno una volta in 4 lanci di un dado;
(ii) esce un doppio 6 almeno una volta in 24 lanci di una coppia di dadi.

2.9) Supponiamo che l'invio di un messaggio su di una linea di comunicazione avvenga con trasmissione asincrona nel 70% dei casi, e con trasmissione sincrona nel 30% dei casi. Inoltre, la trasmissione avviene senza errori nel 90% dei casi, e con errori nel 10% dei casi. La probabilità che una trasmissione sia asincrona e non dia errori è pari a 0,65. Quanto vale la probabilità che una trasmissione sia sincrona e presenti errori?

2.10) Da un'urna contenente 3 biglie numerate da 1 a 3 si estraggono a caso 3 biglie (con reinserimento). Diciamo che si ha una concordanza nell'i-esima estrazione se in tale estrazione viene estratta la biglia avente numero i, per $i = 1, 2, 3$. Calcolare la probabilità che si abbiano in totale k concordanze, per $k = 0, 1, 2, 3$.

CAPITOLO 3 – Probabilità condizionata e indipendenza

3.1 Introduzione

In questo capitolo introduciamo la

probabilità condizionata,

uno dei concetti più importanti della teoria della probabilità.

L'importanza del concetto è duplice.

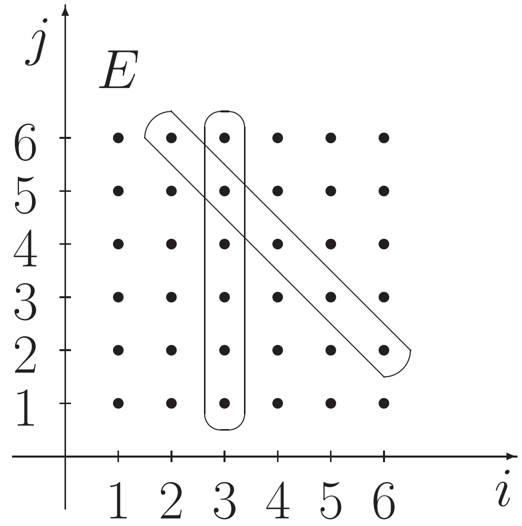
Innanzitutto si è spesso interessati a calcolare la probabilità di un evento disponendo di qualche informazione parziale.

Inoltre le probabilità condizionate sono spesso utilizzate per calcolare più facilmente le probabilità richieste.

3.1 Probabilità condizionata

Supponiamo di lanciare 2 dadi e che ognuno dei 36 possibili esiti sia *equiprobabile* con probabilità $\frac{1}{36}$. Supponiamo inoltre di osservare che l'esito del primo lancio è 3. Allora, data questa informazione, qual è la probabilità che la somma dei due dadi sia uguale a 8? Dato che il primo dado vale 3, vi sono al più 6 possibili esiti dell'esperimento: $(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)$.

Dato che all'inizio questi esiti erano equiprobabili, essi devono avere anche ora la stessa probabilità. Sapendo che il primo dado vale 3, la probabilità (condizionata) di ognuno dei 6 esiti è uguale a $\frac{1}{6}$, mentre la probabilità (condizionata) degli altri 30 punti dello spazio campionario è pari a 0. Pertanto la probabilità richiesta vale $\frac{1}{6}$, perché uno solo dei 6 risultati dà somma 8: $(3,5)$.



In assenza d'informazione sul primo lancio la probabilità di avere somma 8 vale $\frac{5}{36}$.

Definizione. Se $P(F) > 0$, la probabilità condizionata di E dato F è data da

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

Tale definizione è giustificata dalle seguenti considerazioni:

Per esperimenti dotati di spazio campionario finito e con esiti equiprobabili, abbiamo visto che per ogni evento A risulta: $P(A) = |A|/|S|$.

Pertanto, volendo esprimere in tale ambito la probabilità condizionata di E dato F, siamo condotti ad usare il rapporto di casi favorevoli al verificarsi di E (sapendo che si è verificato F) su casi possibili (gli elementi di F), cosicché:

$$P(E|F) = \frac{|E \cap F|}{|F|} = \frac{|E \cap F|/|S|}{|F|/|S|} = \frac{P(E \cap F)}{P(F)}.$$

Questa osservazione, valida quando gli esiti sono equiprobabili, giustifica la definizione di probabilità condizionata data in precedenza, che vale in generale.

Esempio. Nell'esperimento del lancio di un dado non truccato calcolare le probabilità condizionate di $A = \{1, 2\}$ dati gli eventi $B_1 = \{4, 5, 6\}$, $B_2 = \{1, 5, 6\}$, $B_3 = \{1, 2, 6\}$.
Soluzione. Risulta

$$P(A) = \frac{2}{6} = \frac{1}{3}, \qquad P(B_i) = \frac{3}{6} = \frac{1}{2}, \quad i = 1, 2, 3$$

e quindi

$$P(A|B_1) = \frac{P(A \cap B_1)}{P(B_1)} = \frac{P(\emptyset)}{1/2} = 0,$$
$$P(A|B_2) = \frac{P(A \cap B_2)}{P(B_2)} = \frac{P(\{1\})}{1/2} = \frac{1/6}{1/2} = \frac{1}{3},$$
$$P(A|B_3) = \frac{P(A \cap B_3)}{P(B_3)} = \frac{P(\{1, 2\})}{1/2} = \frac{1/3}{1/2} = \frac{2}{3}.$$

Pertanto, sebbene gli eventi B_1, B_2, B_3 siano equiprobabili, la probabilità condizionata di A dato B_k cambia al variare di k, ed in particolare risulta $P(A|B_2) = P(A)$.

Esempio. Nell'esperimento del lancio di una moneta ripetuto 2 volte, supponendo che i quattro punti dello spazio campionario $S = \{cc, ct, tc, tt\}$ siano equiprobabili, qual è la probabilità condizionata che venga testa in entrambi i lanci sapendo che esce testa (a) nel primo lancio; (b) in almeno un lancio?

Soluzione. Sia $E = \{tt\} = \{$testa in entrambi i lanci$\}$, e $F = \{tc, tt\} = \{$testa al primo lancio$\}$. La probabilità cercata in (a) è allora

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(\{tt\})}{P(\{tc, tt\})} = \frac{1/4}{2/4} = \frac{1}{2}.$$

Sia $A = \{ct, tc, tt\}$; la soluzione di (b) è

$$P(E|A) = \frac{P(E \cap A)}{P(A)} = \frac{P(\{tt\})}{P(\{ct, tc, tt\})} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Notiamo che $P(E|F)$ e $P(E|A)$ sono diversi da $P(E) = \dfrac{1}{4}$.

Esempio. Nell'esperimento del lancio di una moneta ripetuto n volte, supponendo che i 2^n punti dello spazio campionario S siano equiprobabili, qual è la probabilità condizionata che venga testa in ogni lancio sapendo che esce testa (a) nel primo lancio; (b) in almeno un lancio?

Soluzione. Sia $E = \{\text{testa in ogni lancio}\}$ e $F = \{\text{testa al primo lancio}\}$. La probabilità cercata in (a) è allora

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E)}{P(F)} = \frac{1/2^n}{1/2} = \frac{1}{2^{n-1}}.$$

Sia $A = \{\text{testa in almeno un lancio}\}$. Poiché $P(A) = 1 - P(\overline{A})$, con $P(\overline{A}) = 1/2^n$, si ha

$$P(E|A) = \frac{P(E \cap A)}{P(A)} = \frac{P(E)}{1 - P(\overline{A})} = \frac{1/2^n}{1 - 1/2^n} = \frac{1}{2^n - 1}.$$

Notiamo che le probabilità condizionate $P(E|F)$ e $P(E|A)$ sono diverse da $P(E) = \dfrac{1}{2^n}$.

Esempio. Uno studente sta svolgendo un test da consegnare dopo un'ora. Si suppone che la probabilità che lo studente termini il test in meno di t ore sia uguale a $t/2$, per ogni $t \in [0,1]$. Qual è la probabilità che lo studente usufruisca dell'intera ora? Qual è la probabilità condizionata che lo studente usufruisca dell'intera ora

(a) sapendo che egli è ancora al lavoro dopo 3/4 d'ora?

(b) sapendo che egli è ancora al lavoro dopo t ore? $(0 \le t \le 1)$

Soluzione. Sia $L_t = \{\text{lo studente conclude l'esame in meno di } t \text{ ore}\}$, con $P(L_t) = t/2$ per $0 \le t \le 1$. Posto $F = \{\text{lo studente utilizza l'intera ora}\}$, si ha $F = \overline{L_1}$ e pertanto

$$P(F) = P(\overline{L_1}) = 1 - P(L_1) = 1 - \frac{1}{2} = \frac{1}{2}.$$

(a) Poiché $F \subset \overline{L_{3/4}}$, si ha $F \cap \overline{L_{3/4}} = F$ e quindi la probabilità cercata è

$$P(F|\overline{L_{3/4}}) = \frac{P(F \cap \overline{L_{3/4}})}{P(\overline{L_{3/4}})} = \frac{P(F)}{1 - P(L_{3/4})} = \frac{\frac{1}{2}}{1 - \frac{3}{8}} = \frac{\frac{1}{2}}{\frac{5}{8}} = \frac{4}{5} = 0{,}8.$$

(b) Analogamente, poiché $F \subset \overline{L_t}$, si ha $F \cap \overline{L_t} = F$ e quindi

$$P(F|\overline{L_t}) = \frac{P(F \cap \overline{L_t})}{P(\overline{L_t})} = \frac{P(F)}{1 - P(L_t)} = \frac{\frac{1}{2}}{1 - \frac{t}{2}} = \frac{1}{2 - t}, \qquad 0 \le t \le 1.$$

Esempio. Nel gioco del bridge le 52 carte sono distribuite equamente a 4 giocatori (Est, Ovest, Nord e Sud). Se Nord e Sud hanno in tutto 8 picche, qual è la probabilità p_k che Est abbia k delle 5 carte di picche rimanenti? $(k = 0, 1, 2, 3, 4, 5)$

Soluzione. Si tratta di probabilità condizionata, ma è più semplice procedere con lo spazio campionario ridotto. Dato che Nord-Sud hanno in tutto 8 picche tra le loro 26 carte, vi sono esattamente 5 picche nelle rimanenti 26 carte di Est-Ovest. Per la equiprobabilità delle distribuzioni delle carte, le probabilità cercate sono date da

$$p_0 = \frac{\binom{5}{0}\binom{21}{13}}{\binom{26}{13}} = p_5 = \frac{\binom{5}{5}\binom{21}{8}}{\binom{26}{13}} = \frac{1 \cdot 203\,490}{10\,400\,600} = \frac{9}{460} \approx 0{,}0196$$

$$p_1 = \frac{\binom{5}{1}\binom{21}{12}}{\binom{26}{13}} = p_4 = \frac{\binom{5}{4}\binom{21}{9}}{\binom{26}{13}} = \frac{5 \cdot 293\,930}{10\,400\,600} = \frac{13}{92} \approx 0{,}141$$

$$p_2 = \frac{\binom{5}{2}\binom{21}{11}}{\binom{26}{13}} = p_3 = \frac{\binom{5}{3}\binom{21}{10}}{\binom{26}{13}} = \frac{10 \cdot 352\,716}{10\,400\,600} = \frac{39}{115} \approx 0{,}339$$

Esempio. Da un'urna contenente r biglie (di cui b sono blu) si estraggono in sequenza n biglie a caso, senza reinserimento ($n \leq r$).
(i) Qual è la probabilità che la prima biglia estratta sia blu?
(ii) Qual è la probabilità condizionata che la prima biglia estratta sia blu, sapendo che k delle n biglie estratte sono blu?

Soluzione. Facciamo riferimento ad uno spazio campionario con esiti equiprobabili.
(i) Ponendo $E = \{\text{la prima biglia estratta è blu}\}$, si ha

$$P(E) = \frac{b\,(r-1)\,(r-2)\,\cdots\,(r-n+1)}{r\,(r-1)\,(r-2)\,\cdots\,(r-n+1)} = \frac{b}{r}.$$

(ii) Ponendo $B_k = \{\text{vengono estratte } k \text{ biglie blu}\}$, $0 \leq k \leq n$, la probabilità condizionata può essere calcolata come rapporto tra il numero di sequenze di lunghezza n aventi una biglia blu al primo posto e $k-1$ biglie blu nei rimanenti $n-1$ posti diviso il numero di sequenze di lunghezza n aventi k biglie blu all'interno. Pertanto si ha

$$P(E|B_k) = \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{(n-1)!}{(k-1)!(n-k)!} \cdot \frac{k!(n-k)!}{n!} = \frac{k}{n}.$$

Moltiplicando ambo i membri dell'uguaglianza $P(E|F) = \dfrac{P(E \cap F)}{P(F)}$ per $P(F)$, si ha

$$P(E \cap F) = P(F)\, P(E|F), \qquad \text{se } P(F) > 0.$$

Analogamente, da $P(F|E) = \dfrac{P(E \cap F)}{P(E)}$ si ha

$$P(E \cap F) = P(E)\, P(F|E), \qquad \text{se } P(E) > 0.$$

Una generalizzazione delle formule precedenti è mostrata qui di seguito, ed è anche nota come *legge delle probabilità composte.*

Proposizione. (Regola del prodotto) Se $P(E_1 \cap \ldots \cap E_{n-1}) > 0$, allora

$$P(E_1 \cap \ldots \cap E_n) = P(E_1)\, P(E_2|E_1)\, P(E_3|E_2 \cap E_1) \ldots P(E_n|E_1 \cap \ldots \cap E_{n-1})$$

Dimostrazione. Per la definizione di probabilità condizionata, dal 2$^{\text{o}}$ membro si ha

$$P(E_1)\frac{P(E_1 \cap E_2)}{P(E_1)}\frac{P(E_1 \cap E_2 \cap E_3)}{P(E_1 \cap E_2)} \ldots \frac{P(E_1 \cap E_2 \cap \ldots \cap E_n)}{P(E_1 \cap E_2 \cap \ldots \cap E_{n-1})} = P(E_1 \cap \ldots \cap E_n)$$

con le probabilità a denominatore strettamente positive perché $P(E_1 \cap \ldots \cap E_{n-1}) > 0$.

Esempio. Da un'urna contenente n biglie numerate da 1 a n si estraggono 3 biglie a caso (senza reinserimento). Affermando che vi è concordanza all'estrazione k-esima se in tale estrazione fuoriesce la biglia avente numero k, calcolare la probabilità

(a) di avere 3 concordanze,

(b) di avere concordanza solo nelle prime 2 estrazioni.

Soluzione. Posto $A_k = \{\text{si ha concordanza all'estrazione } k\text{-esima}\}$, dalla legge delle probabilità composte segue che la probabilità richiesta in (a) è data da

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2};$$

analogamente, la probabilità richiesta in (b) è

$$P(A_1 \cap A_2 \cap \overline{A_3}) = P(A_1)\, P(A_2|A_1)\, P(\overline{A_3}|A_1 \cap A_2) = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{n-3}{n-2}.$$

Notiamo che

$$P(\overline{A_3}|A_1 \cap A_2) = 1 - P(A_3|A_1 \cap A_2) = 1 - \frac{1}{n-2} = \frac{n-3}{n-2}.$$

Esempio. Nel gioco del lotto si estraggono a caso 5 numeri da un'urna contenente 90 numeri compresi tra 1 e 90. (Le cinquine estraibili sono equiprobabili).
(i) Se si sceglie un numero compreso tra 1 e 90 qual è la probabilità che questo sarà tra i 5 estratti?
(ii) E se si scelgono 2 numeri distinti?
(iii) E se, più in generale, se ne scelgono k $(1 \leq k \leq 5)$?

Soluzione. (i) Definiamo l'evento $A_1 = \{$il numero scelto è tra i 5 estratti$\}$; evidentemente risulta

$$P(A_1) = \frac{5}{90} = \frac{1}{18}.$$

(ii) Definiamo l'evento $A_2 = \{$i 2 numeri scelti sono tra i 5 estratti$\}$; questo si può esprimere come intersezione degli eventi $B_1 = \{$il primo numero scelto è tra i 5 estratti$\}$ e $B_2 = \{$il secondo numero scelto è tra i 5 estratti$\}$. Quindi, usando la regola del prodotto e ricordando che le estrazioni sono senza reinserimento, si ha

$$P(A_2) = P(B_1 \cap B_2) = P(B_1)\, P(B_2|B_1) = \frac{5}{90} \cdot \frac{4}{89} = \frac{2}{801} \approx 0{,}0025.$$

(iii) Se si scelgono k numeri distinti ($1 \le k \le 5$), definiamo l'evento $A_k = \{\text{i } k$ numeri scelti sono tra i 5 estratti$\}$. Questo si può esprimere in termini degli eventi $B_i = \{$l'i-esimo numero scelto è tra i 5 estratti$\}$, $1 \le i \le k$:

$$A_k = B_1 \cap B_2 \cap \ldots \cap B_k, \qquad 1 \le k \le 5.$$

Dalla regola del prodotto segue

$$P(A_k) = P(B_1 \cap B_2 \cap \ldots \cap B_k) = P(B_1)\, P(B_2|B_1)\, \ldots\, P(B_k|B_1 \cap \ldots \cap B_{k-1}).$$

Poiché le estrazioni sono senza reinserimento, si ottiene:

$$\begin{aligned}
P(A_3) &= P(B_1)\, P(B_2|B_1)\, P(B_3|B_1B_2) = \frac{5}{90}\,\frac{4}{89}\,\frac{3}{88} = \frac{1}{11.748},\\
P(A_4) &= P(B_1)\, P(B_2|B_1)\, P(B_3|B_1B_2)\, P(B_4|B_1B_2B_3) = \frac{5}{90}\,\frac{4}{89}\,\frac{3}{88}\,\frac{2}{87} = \frac{1}{511.038},\\
P(A_5) &= P(B_1)\, P(B_2|B_1)\, \cdots\, P(B_5|B_1B_2B_3B_4) = \frac{5}{90}\,\frac{4}{89}\,\frac{3}{88}\,\frac{2}{87}\,\frac{1}{86} = \frac{1}{43.949.268}.
\end{aligned}$$

Esempio. Da un'urna contenente r biglie (di cui b sono blu) si estraggono in sequenza n biglie a caso, senza reinserimento ($n \leq r$). Sapendo che k delle biglie estratte sono blu,

(a) qual è la probabilità condizionata che la prima biglia estratta sia blu?

(b) qual è la probabilità condizionata che le prime 2 biglie estratte siano blu?

Soluzione. Poniamo $E = \{$la prima biglia estratta è blu$\}$, $F = \{$la seconda biglia estratta è blu$\}$, e inoltre $B_k = \{$vengono estratte k biglie blu$\}$, $0 \leq k \leq n$. Si ha allora

$$P(E|B_k) = \frac{k}{n}, \qquad P(E \cap F|B_k) = P(E|B_k)\, P(F|E \cap B_k) = \frac{k}{n} \cdot \frac{k-1}{n-1}.$$

Infatti, sapendo che si è verificato l'evento B_k, la sequenza delle n biglie estratte è costituita da k biglie blu e $n-k$ biglie rosse.

(a) In tal caso vi sono k casi favorevoli, su n, affinché la prima biglia sia blu.

(b) Analogamente, vi sono $k-1$ casi favorevoli, su $n-1$, affinché la seconda biglia sia blu sapendo che la prima biglia è blu. Dalla legge delle probabilità composte segue il risultato.

Esercizio. Un'azienda commercializza la propria produzione per il 60% in Italia e per il 40% all'estero. Il 30% dei prodotti vengono venduti *online*, e la quota rimanente nei punti vendita. La metà dei prodotti venduti *online* vengono venduti in Italia.
(a) Qual è la probabilità che un prodotto sia venduto in un punto vendita in Italia?
(b) Se un prodotto è venduto all'estero qual è la probabilità che sia venduto online?

Soluzione. Posto $A = \{$un prodotto è venduto in Italia$\}$, $\overline{A} = \{$un prodotto è venduto all'estero$\}$ e $B = \{$un prodotto è venduto *online*$\}$, $\overline{B} = \{$un prodotto è venduto nei punti vendita$\}$, dalle informazioni note segue:

$$P(A) = 0{,}60 \qquad P(\overline{A}) = 0{,}40 \qquad P(B) = 0{,}30 \qquad P(\overline{B}) = 0{,}70$$

$$P(A|B) = 0{,}50 \qquad \Rightarrow \qquad P(\overline{A}|B) = 0{,}50.$$

(a) Per ricavare $P(A \cap \overline{B})$, notiamo che da $P(A) = P(A \cap B) + P(A \cap \overline{B})$ segue $P(A \cap \overline{B}) = P(A) - P(A \cap B) = P(A) - P(A|B)P(B) = 0{,}60 - 0{,}50 \cdot 0{,}30 = 0{,}45.$
(b) Calcoliamo $P(B|\overline{A})$ tenendo presente che

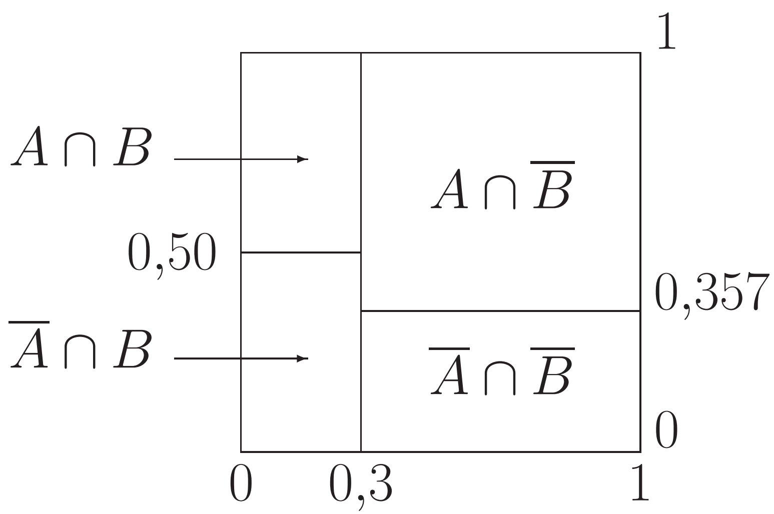
$$P(B|\overline{A}) = \frac{P(\overline{A} \cap B)}{P(\overline{A})} = \frac{P(\overline{A}|B)\,P(B)}{P(\overline{A})} = \frac{0{,}50 \cdot 0{,}30}{0{,}40} = 0{,}375.$$

Poiché gli eventi $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, $\overline{A} \cap \overline{B}$ sono necessari e a 2 a 2 incompatibili, è possibile riportare le probabilità dell'esercizio nella seguente tabella, dove le somme sulle righe e sulle colonne danno i risultati ai margini delle tabelle:

	B	$\overline{B}$	
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
	$P(B)$	$P(\overline{B})$	$P(S)$

	B	$\overline{B}$	
A	0,15	0,45	0,60
$\overline{A}$	0,15	0,25	0,40
	0,30	0,70	1

Ricordiamo che una rappresentazione geometrica dello spazio campionario S può essere anche data da un quadrato di lato 1, in modo che le aree delle regioni corrispondano alle probabilità degli eventi coinvolti.

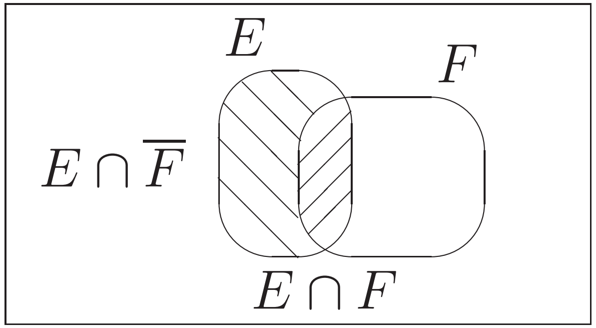


3.3 La formula delle alternative e la formula di Bayes

Proposizione. (Formula delle alternative) Sia F tale che $0 < P(F) < 1$. Se E è un evento qualsiasi risulta

$$P(E) = P(E|F)\,P(F) + P(E|\overline{F})\,P(\overline{F}).$$

Dimostrazione. L'evento E si può esprimere come $E = (E \cap F) \cup (E \cap \overline{F})$, con $E \cap F$ e $E \cap \overline{F}$ eventi incompatibili. Infatti, se un evento elementare appartiene all'evento E, esso inoltre appartiene o all'evento F o al suo complementare $\overline{F}$, e quindi appartiene o all'evento $E \cap F$ oppure a $E \cap \overline{F}$.



Usando la proprietà di additività finita e la regola del prodotto segue:

$$\begin{aligned} P(E) &= P(E \cap F) + P(E \cap \overline{F}) && (E \cap F \text{ e } E \cap \overline{F} \text{ sono incompatibili}) \\ &= P(E|F)\,P(F) + P(E|\overline{F})\,P(\overline{F}), && \text{da cui segue la tesi.} \end{aligned}$$

La formula delle alternative permette di determinare la probabilità di un evento condizionandolo prima alla realizzazione o meno di un altro evento.

Esempio. Da un'urna contenente 5 biglie bianche e 1 biglia rossa, 6 giocatori estraggono a turno 1 biglia a caso, senza reinserimento. Qual è la probabilità che il giocatore k-esimo estragga la biglia rossa?

Soluzione. Posto $A_k = \{\text{il giocatore } k\text{-esimo estrae la biglia rossa}\}$, risulta

$$P(A_1) = \frac{1}{6}, \qquad P(A_2) = P(A_2|A_1)P(A_1) + P(A_2|\overline{A_1})P(\overline{A_1}) = 0 \cdot \frac{1}{6} + \frac{1}{5} \cdot \frac{5}{6} = \frac{1}{6}.$$

Analogamente, si ottiene $P(A_k) = \dfrac{1}{6}$ per $k = 1, 2, \ldots, 6$. Notiamo che risulta

$$P(A_3|\overline{A_1} \cap \overline{A_2}) = \frac{1}{4}, \qquad P(A_4|\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}) = \frac{1}{3},$$

$$P(A_5|\overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap \overline{A_4}) = \frac{1}{2}, \qquad P(A_6|\overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap \overline{A_4} \cap \overline{A_5}) = 1.$$

Osserviamo inoltre che gli eventi $A_1, A_2, \ldots, A_6$ sono necessari e a 2 a 2 incompatibili.

Esempio. Una compagnia assicuratrice suddivide le persone in due categorie: quelle propense a incidenti (il 30%) e quelle che non lo sono (il 70%). Le statistiche mostrano che le persone propense a incidenti hanno probabilità 0,4 di avere un incidente in un anno, mentre per le altre vale 0,2.
(a) Qual è la probabilità che un nuovo assicurato abbia un incidente entro un anno?
(b) Se un nuovo assicurato ha un incidente entro un anno, qual è la probabilità che si tratti di una persona propensa agli incidenti?

Soluzione. Definiamo gli eventi $F = \{\text{una persona è propensa a incidenti}\}$ ed $E = \{\text{un nuovo assicurato ha un incidente entro un anno}\}$. Per le ipotesi fatte risulta:

$$P(F) = 0{,}3 \qquad P(\overline{F}) = 0{,}7 \qquad P(E|F) = 0{,}4 \qquad P(E|\overline{F}) = 0{,}2.$$

(a) Si ha: $P(E) = P(E|F)\,P(F) + P(E|\overline{F})\,P(\overline{F}) = 0{,}4 \cdot 0{,}3 + 0{,}2 \cdot 0{,}7 = 0{,}26.$
(b) Risulta:

$$P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{P(E|F)\,P(F)}{P(E)} = \frac{0{,}4 \cdot 0{,}3}{0{,}26} = \frac{6}{13} \approx 0{,}461.$$

Proposizione. (Formula delle alternative, con n alternative) Se gli eventi $F_1, F_2, \ldots, F_n$ sono a due a due incompatibili, necessari, e ciascuno con probabilità positiva, e se E è un evento qualsiasi, allora risulta

$$P(E) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i).$$

Dimostrazione. Scrivendo

$$E = E \cap S = E \cap \left(\bigcup_{i=1}^{n} F_i \right) = \bigcup_{i=1}^{n} (E \cap F_i) \qquad \text{(con } E \text{ evento qualsiasi)}$$

e osservando che gli eventi $E \cap F_i$, $i = 1, 2, \ldots, n$ sono a due a due incompatibili, per la proprietà di additività finita e per la regola del prodotto si ha infine

$$P(E) = P\left(\bigcup_{i=1}^{n} (E \cap F_i) \right) = \sum_{i=1}^{n} P(E \cap F_i) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i).$$

Nella formula delle alternative

$$P(E) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i)$$

la probabilità di E viene espressa come media ponderata delle $P(E|F_i)$, dove il peso di ciascun termine è uguale alla probabilità dell'evento F_i, rispetto al quale si condiziona.

Dalle ipotesi che gli eventi $F_1, F_2, \ldots, F_n$ sono a due a due incompatibili e necessari segue che in un esperimento si realizza uno e uno solo degli eventi $F_1, F_2, \ldots, F_n$, che evidentemente costituiscono una partizione dello spazio campionario, e quindi

$$\sum_{i=1}^{n} P(F_i) = P\left(\bigcup_{i=1}^{n} F_i\right) = P(S) = 1,$$

per la proprietà di additività finita.

Esempio. Un'urna contiene 3 monete; la 1^a è non truccata, la 2^a mostra testa con probabilità p, mentre la 3^a dà testa con probabilità $1-p$, con $0<p<1$. Se si sceglie una moneta a caso qual è la probabilità che lanciata mostri testa? Se la moneta lanciata mostra testa, qual è la probabilità che si tratti della 2^a?

Soluzione. Definiamo gli eventi $T=\{\text{esce testa}\}$ e $F_j=\{\text{si sceglie la moneta } j\text{-esima}\}$, $j=1,2,3$. Segue che F_1, F_2, F_3 sono necessari e a 2 a 2 incompatibili, con

$$P(F_j)=\frac{1}{3}\qquad (j=1,2,3),$$

e inoltre

$$P(T|F_1)=0{,}5\qquad P(T|F_2)=p\qquad P(T|F_3)=1-p.$$

La probabilità di avere testa è quindi, per la formula delle alternative,

$$P(T)=\sum_{j=1}^{3}P(T|F_j)\,P(F_j)=0{,}5\cdot\frac{1}{3}+p\cdot\frac{1}{3}+(1-p)\cdot\frac{1}{3}=1{,}5\cdot\frac{1}{3}=\frac{1}{2}=0{,}5.$$

Pertanto,

$$P(F_2|T)=\frac{P(T\cap F_2)}{P(T)}=\frac{P(T|F_2)P(F_2)}{P(T)}=\frac{p\,(1/3)}{1/2}=\frac{2}{3}\,p.$$

Esempio. Un vettore booleano di lunghezza 5 contiene 2 bit pari a **1** e 3 bit pari a **0**. Un algoritmo esamina i bit del vettore uno per volta, e si ferma in corrispondenza del primo bit pari a **1**. Qual è la probabilità che l'algoritmo si fermi al passo k-esimo ($1 \leq k \leq 4$)? Qual è la probabilità che il bit successivo al primo **1** sia pari a **0**?

Soluzione. Lo spazio campionario è costituito da $|S| = \binom{5}{2} = 10$ vettori booleani. Ponendo $F_k = \{$l'algoritmo si ferma al passo k-esimo$\}$ ($1 \leq k \leq 4$), tale evento è costituito da tutte le sequenze di S aventi **0** nei primi $k-1$ bit, **1** nel bit k-esimo, e che negli ultimi $5-k$ bit contengono un solo bit pari a **1**, pertanto:

$$P(F_1) = \frac{4}{10}, \qquad P(F_2) = \frac{3}{10}, \qquad P(F_3) = \frac{2}{10}, \qquad P(F_4) = \frac{1}{10}.$$

Posto $E = \{$il bit successivo al primo $\mathbf{1}$ è pari a $\mathbf{0}\}$, si ha

$$P(E|F_1) = \frac{3}{4}, \qquad P(E|F_2) = \frac{2}{3}, \qquad P(E|F_3) = \frac{1}{2}, \qquad P(E|F_4) = 0,$$

e quindi $P(E) = \displaystyle\sum_{k=1}^{4} P(E|F_k)\, P(F_k) = \frac{3}{4} \cdot \frac{4}{10} + \frac{2}{3} \cdot \frac{3}{10} + \frac{1}{2} \cdot \frac{2}{10} + 0 \cdot \frac{1}{10} = \frac{6}{10}.$

Se nell'esempio precedente il vettore booleano ha lunghezza n, allora $|S| = \binom{n}{2}$ e l'evento F_k è costituito dalle sequenze di S aventi **0** nei primi $k-1$ bit, **1** nel bit k-esimo, e che negli ultimi $n-k$ bit contengono un solo bit pari a **1**. Pertanto:

$$P(F_k) = \frac{n-k}{\binom{n}{2}} = \frac{2(n-k)}{n(n-1)}, \qquad P(E|F_k) = \frac{n-k-1}{n-k}, \qquad 1 \le k \le n-1,$$

e quindi

$$P(E) = \sum_{k=1}^{n-1} P(E|F_k)\, P(F_k) = \sum_{k=1}^{n-1} \frac{n-k-1}{n-k} \cdot \frac{2(n-k)}{n(n-1)} = \frac{2}{n(n-1)} \sum_{j=0}^{n-2} j,$$

avendo posto $j = n-k-1$. Ricordando che $\sum_{j=1}^{N} j = \frac{N(N+1)}{2}$, si ha

$$P(E) = \frac{2}{n(n-1)} \cdot \frac{(n-1)(n-2)}{2} = \frac{n-2}{n}.$$

Notiamo che: $\displaystyle\sum_{k=1}^{n-1} P(F_k) = \sum_{k=1}^{n-1} \frac{2(n-k)}{n(n-1)} = \frac{2}{n(n-1)} \sum_{j=1}^{n-1} j = \frac{2}{n(n-1)} \frac{n(n-1)}{2} = 1.$

La formula delle alternative $P(E) = \sum_{i=1}^{n} P(E|F_i)\, P(F_i)$ permette di determinare la probabilità di un evento condizionandolo prima alla realizzazione di uno, e uno solo, degli n eventi $F_1, F_2, \ldots, F_n$. Supponiamo ora che E si sia verificato e di voler determinare quali degli eventi alternativi $F_1, F_2, \ldots, F_n$ si sia anch'esso verificato.

Proposizione. (Formula di Bayes) Se E è un evento avente probabilità positiva, e $F_1, F_2, \ldots, F_n$ sono eventi a due a due incompatibili, ciascuno avente probabilità positiva, e necessari, allora

$$P(F_j|E) = \frac{P(E|F_j)\, P(F_j)}{P(E)} = \frac{P(E|F_j)\, P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\, P(F_i)} \qquad (j = 1, 2, \ldots, n).$$

Dimostrazione. Dalla definizione di probabilità condizionata, dalla regola del prodotto e dalla formula delle alternative segue immediatamente

$$P(F_j|E) = \frac{P(E \cap F_j)}{P(E)} = \frac{P(E|F_j)\, P(F_j)}{P(E)} = \frac{P(E|F_j)\, P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\, P(F_i)} \qquad (j = 1, 2, \ldots, n).$$

Verifichiamo che le probabilità della formula di Bayes sommano all'unità; infatti risulta

$$\sum_{j=1}^{n} P(F_j|E) = \sum_{j=1}^{n} \frac{P(E|F_j)\,P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\,P(F_i)} = \frac{\sum_{j=1}^{n} P(E|F_j)\,P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\,P(F_i)} = 1.$$

Esempio. In un gioco vi sono 3 carte di forma identica, la 1^a con entrambe le facce di colore rosso $\boxed{\mathbf{R}|\mathbf{R}}$, la 2^a con entrambe le facce di colore nero $\boxed{\mathbf{N}|\mathbf{N}}$, la 3^a con una faccia rossa e una nera $\boxed{\mathbf{R}|\mathbf{N}}$. Si sceglie a caso una carta e la si appoggia sul tavolo; se la faccia superiore della carta è rossa, qual è la probabilità che l'altra sia nera?

Soluzione. Indichiamo con F_1, F_2 e F_3 gli eventi riferiti alle scelta di ognuna delle 3 carte, con $P(F_1) = P(F_2) = P(F_3) = 1/3$; poniamo $R = \{$la faccia superiore della carta scelta è rossa$\}$. Dalla formula di Bayes segue

$$P(F_3|R) = \frac{P(R|F_3)\,P(F_3)}{\sum_{i=1}^{3} P(R|F_i)\,P(F_i)} = \frac{\dfrac{1}{2}\cdot\dfrac{1}{3}}{1\cdot\dfrac{1}{3}+0\cdot\dfrac{1}{3}+\dfrac{1}{2}\cdot\dfrac{1}{3}} = \frac{\dfrac{1}{6}}{\dfrac{1}{2}} = \frac{2}{6} = \frac{1}{3}.$$

Notiamo che tale risultato si può ottenere anche come rapporto di casi favorevoli su casi possibili, in quanto una sola delle tre facce rosse ha una faccia nera sul retro.

La formula di Bayes

$$P(F_j|E) = \frac{P(E|F_j)\,P(F_j)}{P(E)} = \frac{P(E|F_j)\,P(F_j)}{\sum_{i=1}^{n} P(E|F_i)\,P(F_i)} \qquad (j = 1, 2, \ldots, n)$$

quando $n = 2$ si riduce alla seguente formula, con $F_1 = F$ e $F_2 = \overline{F}$:

$$P(F|E) = \frac{P(E|F)\,P(F)}{P(E)} = \frac{P(E|F)\,P(F)}{P(E|F)\,P(F) + P(E|\overline{F})\,P(\overline{F})}.$$

Analogamente si ha:

$$P(\overline{F}|E) = \frac{P(E|\overline{F})\,P(\overline{F})}{P(E)} = \frac{P(E|\overline{F})\,P(\overline{F})}{P(E|F)\,P(F) + P(E|\overline{F})\,P(\overline{F})}.$$

Notiamo facilmente che risulta

$$P(F|E) + P(\overline{F}|E) = 1.$$

Esempio. Un sistema di gestione della posta elettronica riceve un messaggio, che si suppone sia *spam* con probabilità 0,7 e *regolare* con probabilità 0,3. Il sistema effettua un controllo su ogni messaggio ricevuto; se riceve un messaggio *spam* lo valuta come tale con probabilità 0,9 (e lo valuta erroreamente come *regolare* con probabilità 0,1) mentre se riceve un messaggio *regolare* lo valuta come tale con probabilità 0,8 (e lo valuta erroreamente come *spam* con probabilità 0,2).

(i) Calcolare la probabilità che il sistema valuti come *spam* il messaggio ricevuto.

(ii) Se il sistema ha valutato come *spam* il messaggio ricevuto, qual è la probabilità che invece sia *regolare*?

(iii) Se il sistema ha valutato come *regolare* il messaggio ricevuto, qual è la probabilità che invece sia *spam*?

Soluzione. Per le ipotesi sull'evento $F = \{\text{il messaggio è } spam\}$ si ha

$$P(F) = 0{,}7 \qquad P(\overline{F}) = 0{,}3.$$

Inoltre, posto $E = \{$il sistema valuta come *spam* il messaggio ricevuto$\}$, risulta

$$P(E\,|\,F) = 0{,}9 \qquad P(\overline{E}\,|\,F) = 0{,}1 \qquad P(\overline{E}\,|\,\overline{F}) = 0{,}8 \qquad P(E\,|\,\overline{F}) = 0{,}2.$$

(i) Pertanto, dalla formula delle alternative segue

$$P(E) = P(E\,|\,F)\,P(F) + P(E\,|\,\overline{F})\,P(\overline{F}) = 0{,}9\cdot 0{,}7 + 0{,}2\cdot 0{,}3 = 0{,}63 + 0{,}06 = 0{,}69.$$

Dalla formula di Bayes si ricavano le probabilità condizionate richieste:

(ii)

$$P(\overline{F}\,|\,E) = \frac{P(E\,|\,\overline{F})\,P(\overline{F})}{P(E)} = \frac{0{,}2\cdot 0{,}3}{0{,}69} = \frac{0{,}06}{0{,}69} = 0{,}0870$$

(iii)

$$P(F\,|\,\overline{E}) = \frac{P(\overline{E}\,|\,F)\,P(F)}{P(\overline{E})} = \frac{0{,}1\cdot 0{,}7}{0{,}31} = \frac{0{,}07}{0{,}31} = 0{,}2258$$

essendo $P(\overline{E}) = 1 - P(E) = 1 - 0{,}69 = 0{,}31.$

Esempio. Si lanciano a caso n monete non truccate; per ogni moneta che mostra testa si inserisce una biglia nera in un'urna, mentre per ogni croce si inserisce una biglia bianca. Se poi si estrae a caso una biglia dall'urna, qual è la probabilità che sia nera? Se la biglia estratta è nera, qual è la probabilità che nell'urna vi erano k biglie nere?

Soluzione. Sia $E_k = \{\text{esce testa } k \text{ volte}\} = \{\text{nell'urna vi sono } k \text{ biglie nere e } n-k \text{ bianche}\}$, $k = 0, 1, \ldots, n$. Tali eventi sono a due a due incompatibili, sono necessari, e $P(E_k) = \frac{\binom{n}{k}}{2^n} > 0$. Sia $A = \{\text{la biglia è nera}\}$; per la formula delle alternative si ha

$$P(A) = \sum_{k=0}^{n} P(A|E_k)\, P(E_k) = \sum_{k=1}^{n} \frac{k}{n} \cdot \frac{\binom{n}{k}}{2^n} = \frac{1}{2^n} \sum_{k=1}^{n} \binom{n-1}{k-1}$$

essendo

$$\frac{k}{n} \cdot \binom{n}{k} = \frac{k}{n} \cdot \frac{n!}{k!\,(n-k)!} = \frac{k}{n} \cdot \frac{n(n-1)!}{k(k-1)!\,(n-k)!} = \frac{(n-1)!}{(k-1)!\,(n-k)!} = \binom{n-1}{k-1}.$$

Ponendo $r = k - 1$, dal teorema del binomio segue

$$P(A) = \frac{1}{2^n} \sum_{r=0}^{n-1} \binom{n-1}{r} = \frac{1}{2^n} \cdot 2^{n-1} = 2^{-1} = \frac{1}{2}.$$

Per ricavare $P(E_k|A)$ facciamo uso della formula di Bayes:

$$P(E_k|A) = \frac{P(A|E_k)\,P(E_k)}{\displaystyle\sum_{i=0}^{n} P(A|E_i)\,P(E_i)} = \frac{\dfrac{k}{n}\cdot\dfrac{\binom{n}{k}}{2^n}}{\dfrac{1}{2}} = \begin{cases} 0 & \text{se } k = 0 \\ \dfrac{\binom{n-1}{k-1}}{2^{n-1}} & \text{se } k = 1, 2, \ldots, n. \end{cases}$$

È facile verificare che la somma delle probabilità $P(E_k|A)$ è unitaria:

$$\sum_{k=0}^{n} P(E_k|A) = \sum_{k=1}^{n} \frac{\binom{n-1}{k-1}}{2^{n-1}} = \frac{1}{2^{n-1}} \sum_{r=0}^{n-1} \binom{n-1}{r} = \frac{1}{2^{n-1}} \cdot 2^{n-1} = 1.$$

3.4 Eventi indipendenti

La probabilità condizionata di E dato F non è generalmente uguale a $P(E)$. In altri termini, la conoscenza della realizzazione dell'evento F modifica in generale la possibilità del realizzarsi o meno di E.

Se $P(E|F) = P(E)$ diciamo che E è indipendente da F. Cioè, E è indipendente da F se la conoscenza della realizzazione di F non cambia la probabilità che si realizzi E.

Se $P(F) > 0$, dalla formula $P(E|F) = \dfrac{P(E \cap F)}{P(F)}$ segue che E è indipendente da F se

$$\frac{P(E \cap F)}{P(F)} = P(E)$$

ossia

$$P(E \cap F) = P(E)\, P(F).$$

Tale formula sussiste anche se si scambia il ruolo di E ed F, pertanto se $P(E) > 0$ e $P(F) > 0$, l'evento E è indipendente da F se F è indipendente da E e viceversa.

La seguente definizione include anche i casi in cui $P(E) = 0$ oppure $P(F) = 0$.

Definizione. Due eventi E ed F si dicono *indipendenti* se vale

$$P(E \cap F) = P(E)\,P(F).$$

Due eventi che non sono indipendenti si dicono *dipendenti*.

Esempio. Uno studente deve sottoporsi a due test. Con probabilità 0,5 supererà il primo test; con probabilità 0,4 supererà il secondo test; con probabilità 0,3 li supererà entrambi. Gli eventi relativi al superamento dei due test sono indipendenti?

Soluzione. Sia B_i l'evento che lo studente superi il test i-esimo, $i = 1, 2$. Risulta

$$P(B_1 \cap B_2) = 0{,}3 \neq 0{,}2 = 0{,}5 \cdot 0{,}4 = P(B_1)\,P(B_2),$$

quindi gli eventi B_1 e B_2 sono dipendenti.

Proposizione. Se A e B eventi tali che $P(A) > 0$ e $P(B) > 0$, allora le seguenti uguaglianze sono equivalenti:

(i) $P(A \cap B) = P(A)\, P(B)$,

(ii) $P(A|B) = P(A)$,

(iii) $P(B|A) = P(B)$.

Dimostrazione. (i) $\Rightarrow$ (ii):

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\, P(B)}{P(B)} = P(A).$$

(ii) $\Rightarrow$ (iii):

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)\, P(B)}{P(A)} = \frac{P(A)\, P(B)}{P(A)} = P(B).$$

(iii) $\Rightarrow$ (i):

$$P(A \cap B) = P(A)\, P(B|A) = P(A)\, P(B).$$

Esercizio. Nell'esperimento che consiste nel lancio di n monete non truccate, sia $T_1 = \{$esce testa al primo lancio$\}$, $U = \{$esce lo stesso risultato negli n lanci$\}$, $A = \{$esce testa almeno 1 volta$\}$. Mostrare che T_1 e U sono indipendenti, ed inoltre che T_1 e A non sono indipendenti. Mostrare che A e U sono indipendenti se, e solo se, $n = 1$.

Nota. Se $P(A) = 0$ oppure $P(A) = 1$, allora l'evento A è indipendente da qualsiasi altro evento B.

Esercizio. Mostrare quanto segue:
(i) $P(A|B) > P(A)$ se e solo se $P(B|A) > P(B)$ se e solo se $P(A\cap B) > P(A)\,P(B)$;
(ii) $P(A|B) < P(A)$ se e solo se $P(B|A) < P(B)$ se e solo se $P(A\cap B) < P(A)\,P(B)$.
Nel caso (i), gli eventi A e B si dicono *positivamente correlati*. Intuitivamente, il verificarsi di uno dei due eventi implica che l'altro è più probabile. Nel caso (ii), A e B si dicono *negativamente correlati*. In questo caso, il verificarsi di uno dei due eventi implica che l'altro è meno probabile.

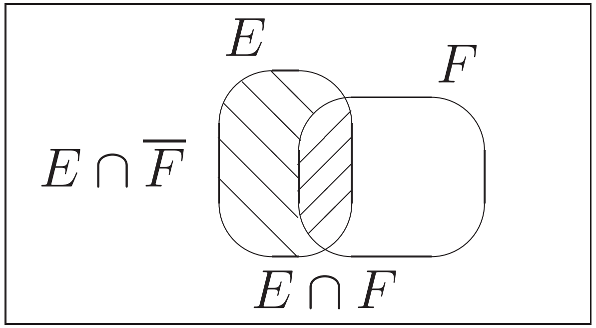
Esercizio. Si effettuano tre estrazioni a caso da un'urna contenente 2 biglie bianche e 3 biglie nere.
(i) Posto $A = \{$almeno una delle tre estratte è bianca$\}$ e $B = \{$ogni estratta ha colore diverso dalla precedente$\}$, calcolare $P(A)$, $P(B)$ e $P(A \cup B)$.
(ii) Stabilire se A e B sono eventi necessari e incompatibili.
(iii) Determinare $P(A|B)$ e $P(B|A)$. Cosa si può dire sull'indipendenza di A e B?
Soluzione. Notiamo che $|S| = (5)_3 = 5 \cdot 4 \cdot 3 = 60$. Pertanto si ha
(i) $P(A) = 1 - P(\overline{A}) = 1 - \dfrac{(3)_3}{(5)_3} = 1 - \dfrac{3 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3} = \dfrac{9}{10}$,
$P(B) = \dfrac{2 \cdot 3 \cdot 1 + 3 \cdot 2 \cdot 2}{5 \cdot 4 \cdot 3} = \dfrac{18}{60} = \dfrac{3}{10}$, $P(A \cup B) = P(A) = \dfrac{9}{10}$, essendo $B \subset A$.
(ii) A e B non sono necessari e non sono incompatibili.
(iii) Si ha $P(A|B) = 1$ e $P(B|A) = \dfrac{1}{3}$, quindi A e B non sono indipendenti bensì positivamente correlati.

Nota. Se per gli eventi A e B risulta $A \subset B$, allora sussiste indipendenza tra i 2 eventi se e solo se $P(A) = 0$ oppure $P(B) = 1$. Negli altri casi, si ha che A e B sono positivamente correlati.

Proposizione. Se E ed F sono eventi indipendenti, allora E ed $\overline{F}$ sono indipendenti.

Dimostrazione. Poichè risulta $E = (E \cap F) \cup (E \cap \overline{F})$, con $E \cap F$ ed $E \cap \overline{F}$ eventi incompatibili, dalla proprietà di additività finita segue

$$P(E) = P(E \cap F) + P(E \cap \overline{F}).$$



Poiché per ipotesi E ed F sono indipendenti, si ha

$$P(E) = P(E)\,P(F) + P(E \cap \overline{F}),$$

ossia

$$P(E \cap \overline{F}) = P(E) - P(E)\,P(F) = P(E)\,[1 - P(F)] = P(E)\,P(\overline{F}).$$

Quindi E ed $\overline{F}$ sono indipendenti.

Notiamo pertanto che se E è indipendente da F, la probabilità che E si realizzi non è modificata dalla realizzazione o meno di F.

Inoltre, se E ed F sono indipendenti, tali sono anche $\overline{E}$ ed F, e gli eventi $\overline{E}$ ed $\overline{F}$.

Esempio. Consideriamo i seguenti eventi nel lancio di due dadi non truccati: $E = \{$la somma dei dadi è 7$\}$, $F = \{$il primo dado dà 4$\}$, $G = \{$il secondo dado dà 3$\}$. Esaminare l'indipendenza delle coppie di eventi E ed F, E e G, E ed $F \cap G$.

Soluzione. L'evento E è indipendente da F ed anche da G, in quanto

$$P(E \cap F) = P(\{(4,3)\}) = \frac{1}{36} \quad = \quad \frac{1}{6} \cdot \frac{1}{6} = P(E)\, P(F),$$

$$P(E \cap G) = P(\{(4,3)\}) = \frac{1}{36} \quad = \quad \frac{1}{6} \cdot \frac{1}{6} = P(E)\, P(G).$$

Inoltre l'evento E non è indipendente da $F \cap G$, poiché

$$P(E|F \cap G) = 1 \quad \neq \quad P(E) = \frac{1}{6}.$$

Questo esempio mostra che se E è indipendente da F e da G, allora non è detto che E sia indipendente da $F \cap G$. Pertanto, appare ragionevole definire l'indipendenza di tre eventi non limitandosi a richiedere l'indipendenza delle 3 possibili coppie, ma imponendo anche una condizione che coinvolga complessivamente i 3 eventi.

Definizione. Tre eventi E, F, G si dicono indipendenti se

$$P(E \cap F \cap G) = P(E)\, P(F)\, P(G)$$

$$P(E \cap F) = P(E)\, P(F)$$

$$P(E \cap G) = P(E)\, P(G)$$

$$P(F \cap G) = P(F)\, P(G)$$

Si noti che se E, F, G sono indipendenti, allora E è indipendente da ogni evento formato a partire da F e G. Ad esempio E è indipendente da $F \cup G$. Infatti si ha

$$\begin{aligned} P(E \cap (F \cup G)) &= P((E \cap F) \cup (E \cap G)) \\ &= P(E \cap F) + P(E \cap G) - P(E \cap F \cap G) \\ &= P(E)\, P(F) + P(E)\, P(G) - P(E)\, P(F \cap G) \\ &= P(E)\, [P(F) + P(G) - P(F \cap G)] \\ &= P(E)\, P(F \cup G). \end{aligned}$$

Esempio. Nel lancio di due monete non truccate, con $S = \{cc, ct, tc, tt\}$, stabilire se i seguenti eventi sono indipendenti:

$$T_1 = \{tc, tt\}, \qquad T_2 = \{ct, tt\}, \qquad U = \{cc, tt\}.$$

Soluzione. Si ricava facilmente che

$$\begin{aligned}
\frac{1}{4} = P(\{tt\}) = P(T_1 \cap T_2) \quad &= \quad P(T_1)\,P(T_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \\
\frac{1}{4} = P(\{tt\}) = P(T_1 \cap U) \quad &= \quad P(T_1)\,P(U) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \\
\frac{1}{4} = P(\{tt\}) = P(T_2 \cap U) \quad &= \quad P(T_2)\,P(U) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.
\end{aligned}$$

Gli eventi T_1, T_2, U sono dunque indipendenti a coppie. Ciononostante, essendo

$$\frac{1}{4} = P(\{tt\}) = P(T_1 \cap T_2 \cap U) \quad \neq \quad P(T_1)\,P(T_2)\,P(U) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8},$$

i tre eventi non sono indipendenti.

Esempio. Nell'eperimento che consiste nel generare a caso una sequenza booleana di lunghezza 3, e quindi $|S| = 2^3 = 8$, stabilire se i seguenti eventi sono indipendenti:

$$A = \{000, 001, 010, 100\}, \quad B = \{000, 001, 100, 101\}, \quad C = \{011, 100, 110, 111\}.$$

Soluzione. Lo spazio campionario è costituito da 8 sequenze equiprobabili, quindi

$$\begin{aligned}
\frac{1}{8} = P(\{100\}) = P(A \cap B \cap C) \quad &= \quad P(A)\,P(B)\,P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}, \\
\frac{3}{8} = P(\{000, 001, 100\}) = P(A \cap B) \quad &\neq \quad P(A)\,P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \\
\frac{1}{8} = P(\{100\}) = P(A \cap C) \quad &\neq \quad P(A)\,P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}, \\
\frac{1}{8} = P(\{100\}) = P(B \cap C) \quad &\neq \quad P(B)\,P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.
\end{aligned}$$

Gli eventi A, B ,C non sono indipendenti poiché non sono indipendenti a coppie.

Definizione. Gli eventi $E_1, E_2, \ldots, E_n$ si dicono indipendenti se per ogni sottoinsieme $E_{i_1}, E_{i_2}, \ldots, E_{i_r}$, $2 \leq r \leq n$, di questi eventi si ha

$$P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_r}) = P(E_{i_1})\, P(E_{i_2})\, \ldots\, P(E_{i_r}).$$

Il numero di uguaglianze coinvolte nell'indipendenza di n eventi è pari al numero di sottoinsiemi di cardinalità $r = 2, 3, \ldots, n$ di un insieme di n elementi, ossia:

$$\sum_{r=2}^{n} \binom{n}{r} = 2^n - \binom{n}{1} - \binom{n}{0} = 2^n - n - 1.$$

Definizione. Gli eventi di una successione $E_1, E_2, \ldots$ si dicono indipendenti se ogni sottoinsieme finito di essi $E_{i_1}, E_{i_2}, \ldots, E_{i_r}$, $r \geq 2$, è formato da eventi indipendenti.

Notiamo che se gli eventi $E_1, E_2, \ldots$ sono indipendenti, allora ogni sequenza di eventi ottenuta sostituendo qualche evento E_i con l'evento complementare $\overline{E}_i$ è ancora costituita da eventi indipendenti.

Alcuni esperimenti talvolta consistono nell'effettuare una successione di sub-esperimenti identici, ripetuti nelle stesse condizioni, ai quali si dà il nome di *prove*.

Ad esempio, nel lancio ripetuto di una moneta, ogni lancio corrisponde ad una prova.

Talora è ragionevole supporre che gli esiti di ogni gruppo di prove non abbiano effetti sugli esiti delle altre prove. In tal caso diciamo che le prove sono indipendenti. Ne segue che la successione

$$E_1, E_2, \ldots, E_n, \ldots$$

costituisce un insieme di eventi indipendenti, dove E_i dipende esclusivamente dall'esito dell'i-esima prova.

Esempio. Nel lancio di 3 monete non truccate, consideriamo i seguenti eventi: $T_1 = \{\text{testa al primo lancio}\}$, $B = \{\text{esce almeno una volta testa negli ultimi 2 lanci}\}$ e $U = \{\text{nei 3 lanci si ha lo stesso risultato}\}$. Studiare l'indipendenza di T_1, B e U.

Soluzione. Lo spazio campionario è costituito da 8 sequenze equiprobabili, quindi

$$P(T_1) = \frac{4}{8} = \frac{1}{2}, \qquad P(B) = 1 - P(\overline{B}) = 1 - \frac{2}{8} = \frac{6}{8} = \frac{3}{4}, \qquad P(U) = \frac{2}{8} = \frac{1}{4},$$

essendo $\overline{B} = \{\text{non esce testa negli ultimi 2 lanci}\} = \{ccc, tcc\}$. Inoltre si ha

$$P(T_1 \cap B) = P(\{tct, ttc, ttt\}) = \frac{3}{8} \quad = \quad P(T_1)\, P(B) = \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8},$$

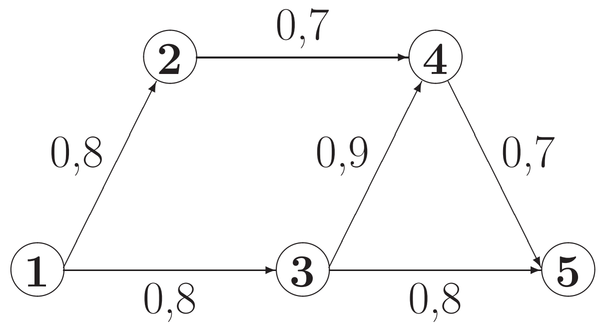
$$P(T_1 \cap U) = P(\{ttt\}) = \frac{1}{8} \quad = \quad P(T_1)\, P(U) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8},$$

$$P(B \cap U) = P(\{ttt\}) = \frac{1}{8} \quad \neq \quad P(B)\, P(U) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16},$$

$$P(T_1 \cap B \cap U) = P(\{ttt\}) = \frac{1}{8} \quad \neq \quad P(T_1)\, P(B)\, P(U) = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{32}.$$

Quindi, T_1 e B sono indipendenti, T_1 e U sono indipendenti, B e U non sono indipendenti, e infine T_1, B e U non sono indipendenti. Notiamo che T_1 e B si riferiscono a esperimenti distinti, al contrario di T_1 e U.

Esempio. Si consideri un sistema costituito da 5 unità collegate mediante archi orientati (vedi figura, dove su ogni arco è indicata la probabilità che esso sia attivo, indipendentemente dagli altri archi). Calcolare la probabilità che vi sia almeno un percorso da **1** a **5** con tutti gli archi attivi.



Soluzione. Sia $A = \{$gli archi del percorso **1245** sono attivi$\}$, $B = \{$gli archi del percorso **1345** sono attivi$\}$ e $C = \{$gli archi del percorso **135** sono attivi$\}$. Per la formula di inclusione/esclusione e l'indipendenza degli archi, la probabilità richiesta è

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

con $P(A) = 0{,}8 \cdot 0{,}7 \cdot 0{,}7 = 0{,}392$; $P(B) = 0{,}8 \cdot 0{,}9 \cdot 0{,}7 = 0{,}504$; $P(C) = 0{,}8 \cdot 0{,}8 = 0{,}64$; $P(A \cap B) = 0{,}8 \cdot 0{,}7 \cdot 0{,}7 \cdot 0{,}8 \cdot 0{,}9 = 0{,}2822$; $P(A \cap C) = 0{,}8 \cdot 0{,}7 \cdot 0{,}7 \cdot 0{,}8 \cdot 0{,}8 = 0{,}2509$; $P(B \cap C) = 0{,}8 \cdot 0{,}9 \cdot 0{,}7 \cdot 0{,}8 = 0{,}4032$; $P(A \cap B \cap C) = 0{,}8 \cdot 0{,}7 \cdot 0{,}7 \cdot 0{,}8 \cdot 0{,}9 \cdot 0{,}8 = 0{,}2258$; ne segue: $P(A \cup B \cup C) = 0{,}8255$.

Esempio. Con riferimento all'esempio precedente, trovare l'errore nel seguente testo:

$$\begin{aligned} P(A \cup B \cup C) &= 1 - P(\overline{A} \cap \overline{B} \cap \overline{C}) \\ &= 1 - P(\overline{A}) \cdot P(\overline{B}) \cdot P(\overline{C}) \\ &= 1 - (1 - 0{,}392) \cdot (1 - 0{,}504) \cdot (1 - 0{,}64) \\ &= 1 - 0{,}1086 = 0{,}8914. \end{aligned}$$

Soluzione. Gli eventi A, B, C non sono indipendenti, quindi la seconda uguaglianza è falsa. Invero, gli eventi A e C sono indipendenti, mentre gli eventi A e B, e gli eventi B e C, non sono indipendenti.

Esempio. In una sequenza infinita di prove indipendenti ogni prova ha 2 esiti: successo con probabilità p e insuccesso con probabilità $1-p$. Qual è la probabilità che
(a) vi sia almeno un successo nelle prime n prove;
(b) vi siano esattamente k successi nelle prime n prove, con $0 \leq k \leq n$;
(c) tutte le prove abbiano successo?

Soluzione. (a) Ponendo $E_i = \{\text{si ha successo alla prova } i\text{-esima}\}$, $i = 1, 2, \ldots$, con $P(E_i) = p$, la probabilità di avere almeno un successo nelle prime n prove è data da

$$
\begin{aligned}
P\left(\bigcup_{i=1}^{n} E_i\right) &= P\left(\overline{\bigcap_{i=1}^{n} \overline{E}_i}\right) && \text{(dalla formula di De Morgan)} \\
&= 1 - P\left(\bigcap_{i=1}^{n} \overline{E}_i\right) && \text{(passando all'evento complementare)} \\
&= 1 - \prod_{i=1}^{n} P(\overline{E}_i) && \text{(facendo uso dell'indipendenza)} \\
&= 1 - (1-p)^n && \text{(essendo } P(\overline{E}_i) = 1-p\text{).}
\end{aligned}
$$

Notiamo che $P\left(\bigcap_{i=1}^{n} \overline{E}_i\right)$ è la probabilità di avere tutti insuccessi nelle prime n prove.

(b) Consideriamo le sequenze di prove che nei primi n esiti hanno k successi e $n-k$ insuccessi. Ognuna di tali sequenze si realizza, per l'indipendenza, con probabilità $p^k(1-p)^{n-k}$. Ad esempio, se i k successi si hanno nelle prime k prove si ha

$$\begin{aligned} P(E_1 \dots E_k \overline{E}_{k+1} \dots \overline{E}_n) &= P(E_1) \cdots P(E_k) P(\overline{E}_{k+1}) \cdots P(\overline{E}_n) \\ &= p \cdot p \cdots p \cdot (1-p) \cdot (1-p) \cdots (1-p) = p^k (1-p)^{n-k}. \end{aligned}$$

Vi sono $\binom{n}{k}$ sequenze di questo tipo poiché vi sono $\frac{n!}{k!(n-k)!}$ permutazioni distinte di k successi e $n-k$ insuccessi, quindi la probabilità richiesta è $\binom{n}{k} p^k(1-p)^{n-k}$.

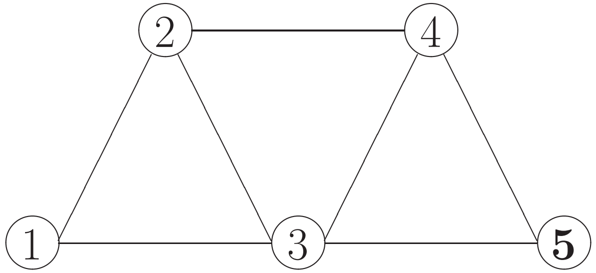
(c) La probabilità di avere n successi nelle prime n prove è, per l'indipendenza,

$$P\left(\bigcap_{i=1}^{n} E_i\right) = \prod_{i=1}^{n} P(E_i) = p^n.$$

Quindi, per la proprietà di continuità della probabilità, la probabilità che tutte le prove abbiano successo è data da

$$P\left(\bigcap_{i=1}^{\infty} E_i\right) = P\left(\lim_{n\to\infty} \bigcap_{i=1}^{n} E_i\right) = \lim_{n\to\infty} P\left(\bigcap_{i=1}^{n} E_i\right) = \lim_{n\to\infty} p^n = \begin{cases} 0 & \text{se } p < 1 \\ 1 & \text{se } p = 1. \end{cases}$$

Esempio. Consideriamo il grafo in figura, costituito da 5 nodi e da 7 archi. I primi 4 nodi sono colorati di bianco, il nodo 5 è colorato di nero. Si attiva un protocollo (*a conoscenza zero*) in cui Alice invia a Bob informazioni attraverso n prove di Bernoulli ripetute, in cui ad ogni prova, indipendentemente dalle altre, trasmette un arco scelto a caso, mostrando solo il colore dei due nodi connessi all'arco scelto. Se Alice dichiara che il grafo non ha nodi neri, qual è la probabilità che Bob non scopra che ciò è falso? Quante prove occorre ripetere per garantire che tale probabilità sia minore di 10^{-k}?



Soluzione. Detto F l'evento d'interesse, esso si realizza quando in ciascuna delle n prove si trasmette uno dei 5 archi che non sono connessi al nodo di colore nero. Quindi, per l'indipendenza delle prove, si ha

$$P(F) = \left(\frac{5}{7}\right)^n < 10^{-k} \quad \Leftrightarrow \quad n \log \frac{5}{7} < -k \log 10 \quad \Leftrightarrow \quad n > k \, \frac{\log 10}{|\log(5/7)|} = k \; 6{,}84.$$

Esempio. In un gioco del lancio di 2 dadi si vince se esce almeno una volta 6.
(a) Qual è la probabilità p di vincita di un gioco?
(b) Qual la probabilità che in 5 giochi si vinca 2 volte? (c) si vinca almeno una volta?

Soluzione. (a) Posto $E_i = \{\text{il dado } i\text{-esimo dà } 6\}$, $i = 1, 2$, la probabilità di vincita di un gioco è $p = P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$. Per l'indipendenza dei lanci risulta $p = P(E_1) + P(E_2) - P(E_1)P(E_2) = \frac{1}{6} + \frac{1}{6} - \frac{1}{6} \cdot \frac{1}{6} = \frac{11}{36} = 0{,}30\overline{5}$.
(b) Poichè i 5 giochi sono indipendenti, la probabilità che in 5 giochi si vinca 2 volte (e quindi si perda 3 volte) è

$$\binom{5}{2} p^2 (1-p)^3 = 10 \cdot \left(\frac{11}{36}\right)^2 \left(\frac{25}{36}\right)^3 = 0{,}3127.$$

(c) Esprimendo la probabilità di vincere almeno una volta in termini dell'evento complementare si ha

$$1 - (1-p)^5 = 1 - \left(\frac{25}{36}\right)^5 = 0{,}8385.$$

Esempio. Risolvere i quesiti dell'esempio precedente supponendo che il primo dado è stato truccato, nel senso che il 5 è stato modificato in 6.

Soluzione. (a) Posto $E_i = \{\text{il dado } i\text{-esimo dà } 6\}$, $i = 1, 2$, la probabilità di vincita di un gioco è $p = P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$. Per l'indipendenza dei lanci risulta $p = P(E_1) + P(E_2) - P(E_1)P(E_2) = \frac{2}{6} + \frac{1}{6} - \frac{2}{6} \cdot \frac{1}{6} = \frac{16}{36} = \frac{4}{9} = 0,\overline{4}$.

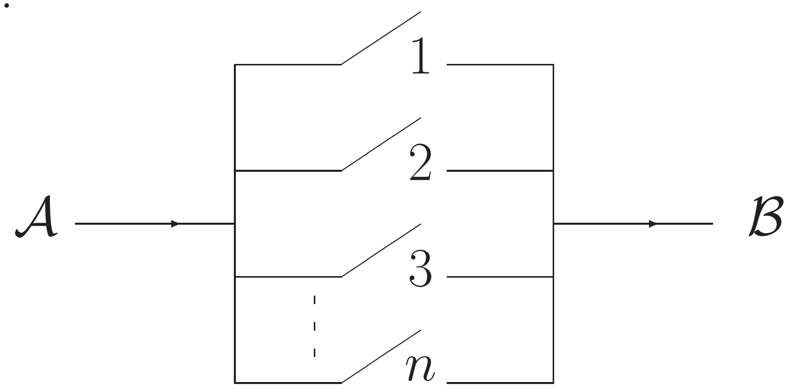
(b) Analogamente all'esempio precedente, la probabilità che in un 5 giochi si vinca 2 volte (e quindi si perda 3 volte) è

$$\binom{5}{2} p^2 (1-p)^3 = 10 \cdot \left(\frac{4}{9}\right)^2 \left(\frac{5}{9}\right)^3 = 0{,}3387$$

(c) e la probabilità di vincere almeno una volta è

$$1 - (1-p)^5 = 1 - \left(\frac{5}{9}\right)^5 = 0{,}9471.$$

Esempio. Un sistema formato da n componenti è detto in parallelo se esso funziona quando almeno uno dei suoi componenti funziona.
Il componente i-esimo, indipendentemente dagli altri, funziona con probabilità p_i, $i = 1, 2, \dots, n$.
Qual è la probabilità che il sistema funzioni?



Soluzione.
Sia $E_i = \{\text{il componente } i\text{-esimo funziona}\}$. Allora

$$
\begin{aligned}
P(F_p) = P(\text{il sistema funziona}) &= P\left(\bigcup_{i=1}^{n} E_i\right) = 1 - P\left(\bigcap_{i=1}^{n} \overline{E_i}\right) \\
&= 1 - \prod_{i=1}^{n} P\left(\overline{E_i}\right) = 1 - \prod_{i=1}^{n} (1 - p_i)
\end{aligned}
$$

per l'indipendenza.

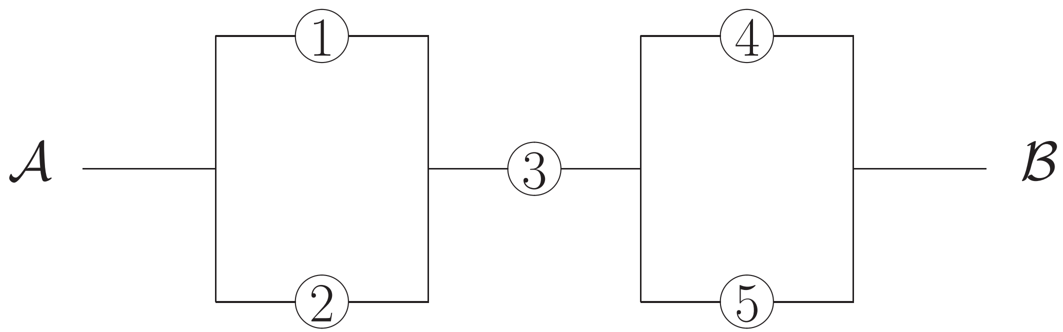
Esempio. Un sistema formato da n componenti è detto in serie se esso funziona quando tutti i suoi componenti funzionano. Il componente i-esimo, indipendentemente dagli altri, funziona con probabilità p_i, $i = 1, 2, \ldots, n$. Qual è la probabilità che il sistema funzioni?

Soluzione. Sia $E_i = \{\text{il componente } i\text{-esimo funziona}\}$. Allora

$$P(F_s) = P(\text{il sistema funziona}) = P\left(\bigcap_{i=1}^{n} E_i\right) = \prod_{i=1}^{n} P\left(E_i\right) = \prod_{i=1}^{n} p_i$$

per l'indipendenza.

Esempio. Un sistema è formato da 5 componenti come in figura. Il componente i-esimo, indipendentemente dagli altri, funziona con probabilità p_i, $i = 1, 2, \ldots, 5$. Qual è la probabilità che il sistema funzioni? E se $p_i = p$ per ogni i? Confrontare tale probabilità con quelle dei sistemi in serie e in parallelo con 3 componenti.



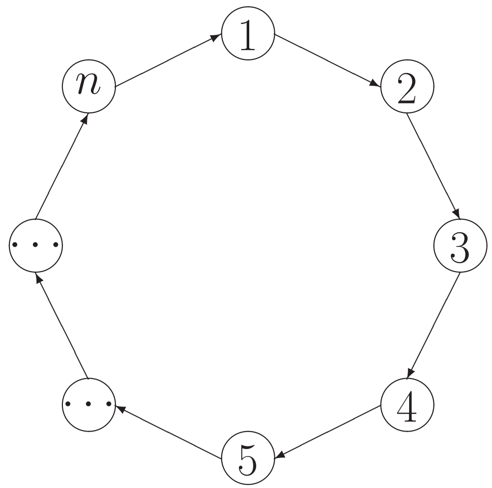
Soluzione. Sia $E_i = \{\text{il componente } i\text{-esimo funziona}\}$. Allora, per l'indipendenza,

$$\begin{aligned} P(F_m) &:= P(\text{il sistema funziona}) = P\left((E_1 \cup E_2) \cap E_3 \cap (E_4 \cup E_5)\right) \\ &= P(E_1 \cup E_2) P(E_3) P(E_4 \cup E_5) = (p_1 + p_2 - p_1 p_2)\, p_3\, (p_4 + p_5 - p_4 p_5). \end{aligned}$$

Pertanto, se $p_i = p$ per ogni i, risulta $P(F_m) = p^3 (2-p)^2$, $0 \le p \le 1$. Confrontando tale probabilità con con quelle dei sistemi in serie e in parallelo con 3 componenti si ha

$$P(F_s) = 1 - (1-p)^3 \le P(F_m) \le p^3 = P(F_p), \qquad 0 \le p \le 1.$$

Esempio. Si consideri un sistema costituito da n unità numerate da 1 a n (vedi figura). Ciascuna unità, indipendentemente dalle altre, sceglie a caso un numero da 1 a n e lo invia all'unità successiva in senso orario. Calcolare la probabilità che almeno un'unità riceva dall'unità precedente un numero identico al proprio, e valutarla quando $n \to \infty$ ricordando che $\boxed{\lim_{n\to\infty}\left(1+\frac{\alpha}{n}\right)^n = e^\alpha}$



Soluzione. Per $C_i = \{\text{l'unità } i\text{-esima riceve il numero } i\}$, si ha $P(C_i) = \frac{1}{n}$ e quindi per l'indipendenza risulta

$$P\left(\bigcup_{i=1}^{n} C_i\right) = 1 - P\left(\bigcap_{i=1}^{n} \overline{C}_i\right) = 1 - \prod_{i=1}^{n} P\left(\overline{C}_i\right) = 1 - \left(1 - \frac{1}{n}\right)^n.$$

Pertanto,

$$\lim_{n\to+\infty} P\left(\bigcup_{i=1}^{n} C_i\right) = 1 - \lim_{n\to+\infty}\left(1-\frac{1}{n}\right)^n = 1 - e^{-1} = 0{,}6321.$$

Esempio. Nell'esperimento che consiste nel lanciare n volte una moneta truccata, indicando con t la fuoriuscita di testa e con c di croce, lo spazio campionario è

$$S = \{(\omega_1, \omega_2, \ldots, \omega_n) \colon \omega_i \in \{t, c\} \text{ per } 1 \leq i \leq n\},$$

le cui 2^n sequenze in generale non sono equiprobabili. Supponendo che per $0 < p < 1$

$$P(T_k) = P(\{\text{al } k\text{-esimo lancio esce testa}\}) = p, \qquad 1 \leq k \leq n,$$

e che i lanci siano indipendenti, calcolare le probabilità degli eventi $A_k = \{\text{nei primi } k$ lanci esce testa$\}$, $A'_k = \{\text{nei primi } k \text{ lanci esce croce}\}$, $B_k = \{\text{esce testa } k \text{ volte}\}$.

Soluzione. Per l'indipendenza si ha

$$P(A_k) = P(T_1 \cap T_2 \cap \cdots \cap T_k) = P(T_1)\, P(T_2) \cdots P(T_k) = p^k,$$

e anche $P(A'_k) = P(\overline{T}_1 \cap \overline{T}_2 \cap \cdots \cap \overline{T}_k) = (1-p)^k$ essendo $P(\overline{T}_k) = 1 - p$. Inoltre:

$$P(B_k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad 0 \leq k \leq n.$$

Nel caso di moneta non truccata, essendo $p = \frac{1}{2}$, si ha

$$P(A_k) = P(A'_k) = \frac{1}{2^k}, \qquad P(B_k) = \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \binom{n}{k} \frac{1}{2^n}.$$

Esercizio. In un corso universitario il 42% degli studenti effettua tirocinio esterno e non svolge attività in Erasmus. Inoltre, tra gli studenti che svolgono attività in Erasmus il 40% effettua tirocinio interno. Infine, l'82% degli studenti effettua tirocinio interno oppure non svolge attività in Erasmus. Se si sceglie uno studente a caso, qual è la probabilità che effettui tirocinio esterno?, che effettui tirocinio esterno e svolga attività in Erasmus?, che effettui tirocinio interno e non svolga attività in Erasmus? Sussiste indipendenza tra l'effettuare tirocinio esterno e svolgere attività in Erasmus?

Soluzione. Poniamo $A = \{\text{si effettua tirocinio esterno}\}$, $\overline{A} = \{\text{si effettua tirocinio interno}\}$ e $B = \{\text{si svolge attività in Erasmus}\}$, $\overline{B} = \{\text{non si svolge attività in Erasmus}\}$. Per quanto risulta noto possiamo scrivere:

$$P(A \cap \overline{B}) = 0{,}42 \qquad P(\overline{A}|B) = 0{,}40 \qquad P(\overline{A} \cup \overline{B}) = 0{,}82.$$

Notiamo che $P(\overline{A} \cup \overline{B}) = P(\overline{A}) + P(A \cap \overline{B})$, e quindi

$$P(\overline{A}) = P(\overline{A} \cup \overline{B}) - P(A \cap \overline{B}) = 0{,}82 - 0{,}42 = 0{,}40 \quad \Rightarrow \quad P(A) = 0{,}60.$$

Partanto, dalla relazione $P(A) = P(A \cap B) + P(A \cap \overline{B})$ si trae

$$P(A \cap B) = P(A) - P(A \cap \overline{B}) = 0{,}60 - 0{,}42 = 0{,}18.$$

Poiché si desidera calcolare $P(\overline{A} \cap \overline{B})$ notiamo che

$$P(\overline{A} \cap B) = P(\overline{A}|B)\, P(B) = 0{,}40 \cdot P(B) = 0{,}40 \cdot [P(A \cap B) + P(\overline{A} \cap B)]$$

e quindi

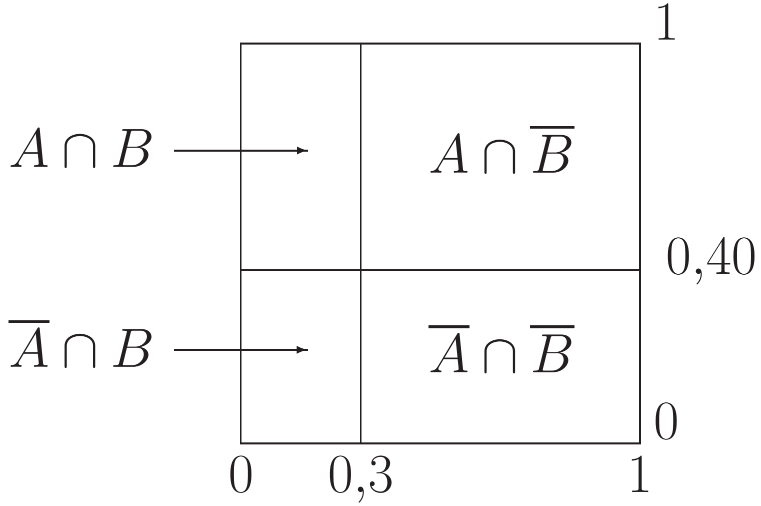
$$P(\overline{A}\cap B) - 0{,}40 \cdot P(\overline{A}\cap B) = 0{,}40 \cdot P(A\cap B) \quad \Rightarrow \quad 0{,}60 \cdot P(\overline{A}\cap B) = 0{,}40 \cdot P(A\cap B).$$

Si ricava

$$P(\overline{A} \cap B) = \frac{2}{3}\, P(A \cap B) = \frac{2}{3}\, 0{,}18 = 0{,}12$$

da cui segue $P(\overline{A} \cap \overline{B}) = P(\overline{A}) - P(\overline{A} \cap B) = 0{,}40 - 0{,}12 = 0{,}28$ e quindi si trae facilmente che A e B sono indipendenti.

	B	$\overline{B}$	
A	0,18	0,42	0,60
$\overline{A}$	0,12	0,28	0,40
	0,30	0,70	1



3.5 $P(\cdot|F)$ è una probabilità

Vedremo ora che la probabilità condizionata soddisfa i tre assiomi della probabilità.

Proposizione. Sia F un evento tale che $P(F) > 0$; allora la probabilità condizionata da F è una funzione $P(\cdot\,|F)\colon \mathcal{F} \to \mathbb{R}$ tale che

(a) $0 \le P(A|F) \le 1$ per ogni evento A;

(b) $P(S|F) = 1$;

(c) per ogni successione di eventi a due a due incompatibili $A_1, A_2, \ldots$,

$$P\left(\bigcup_{i=1}^{\infty} A_i|F\right) = \sum_{i=1}^{\infty} P(A_i|F).$$

Dimostrazione. (a) Occorre mostrare che $0 \le P(A \cap F)/P(F) \le 1$. La prima disuguaglianza è ovvia, mentre la seconda discende dal fatto che $A \cap F \subset F$, da cui segue $P(A \cap F) \le P(F)$.

(b) Si ha

$$P(S|F) = \frac{P(S \cap F)}{P(F)} = \frac{P(F)}{P(F)} = 1.$$

(c) Risulta

$$P\left(\bigcup_{i=1}^{\infty} A_i \Big| F\right) = P\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap F\right) \frac{1}{P(F)} = P\left(\bigcup_{i=1}^{\infty} (A_i \cap F)\right) \frac{1}{P(F)}.$$

Essendo gli eventi $A_i \cap F$ incompatibili, per la proprietà di additività numerabile si ha

$$P\left(\bigcup_{i=1}^{\infty} A_i \Big| F\right) = \sum_{i=1}^{\infty} P(A_i \cap F) \frac{1}{P(F)} = \sum_{i=1}^{\infty} P(A_i|F),$$

e ciò conclude la dimostrazione.

Nota. Poiché la probabilità condizionata $P(\cdot\,|F)$ soddisfa gli assiomi della probabilità, ad essa si applicano le regole dimostrate in precedenza; ad esempio

$$P(A \cup B|F) = P(A|F) + P(B|F) - P(A \cap B|F), \qquad P(\overline{A}|F) = 1 - P(A|F).$$

Esempio. Stabilire se $\widetilde{P}(E|F) := \frac{P(E \cup F)}{P(F)}$, $\widetilde{Q}(E|F) := \frac{P(E \cap \overline{F})}{P(\overline{F})}$ e $\widetilde{R}(E|F) := \frac{P(E \cap \overline{F})}{P(F)}$ soddisfano i 3 assiomi della probabilità.

Soluzione. $\widetilde{P}(E|F)$ e $\widetilde{R}(E|F)$ non sono probabilità, $\widetilde{Q}(E|F)$ è una probabilità.

Proposizione. Siano F e B eventi tali che $P(F) > 0$, $P(B \cap F) > 0$ e $P(\overline{B} \cap F) > 0$; allora

$$P(A|F) = P(A|B \cap F)\, P(B|F) + P(A|\overline{B} \cap F)\, P(\overline{B}|F).$$

Dimostrazione. Dalle proprietà della probabilità condizionata segue

$$\begin{aligned} P(A|F) &= \frac{P(A \cap F)}{P(F)} = \frac{P(A \cap B \cap F)}{P(F)} + \frac{P(A \cap \overline{B} \cap F)}{P(F)} \\ &= \frac{P(A|B \cap F)\, P(B \cap F)}{P(F)} + \frac{P(A|\overline{B} \cap F)\, P(\overline{B} \cap F)}{P(F)} \\ &= P(A|B \cap F)\, P(B|F) + P(A|\overline{B} \cap F)\, P(\overline{B}|F). \end{aligned}$$

Notiamo che la formula appena dimostrata è un'estensione di quella delle alternative

$$P(A) = P(A|B)\, P(B) + P(A|\overline{B})\, P(\overline{B}).$$

Esempio. Un'urna contiene 6 biglie, di cui 4 sono azzurre e le 2 rimanenti sono rosse. Se ne estraggono a caso 3 in sequenza, senza reinserimento. Se la prima estratta è azzurra, qual è la probabilità che la terza estratta sia azzurra?

Soluzione. Definiamo gli eventi $A_i = \{$l'i-esima biglia estratta è azzurra$\}$, $i = 1, 2, 3$. Poiché si ha

$$P(A_2|A_1) = \frac{3}{5}, \qquad P(\overline{A_2}|A_1) = 1 - P(A_2|A_1) = \frac{2}{5}$$

e

$$P(A_3|A_1 \cap A_2) = \frac{2}{4}, \qquad P(A_3|A_1 \cap \overline{A_2}) = \frac{3}{4},$$

si ottiene

$$\begin{aligned} P(A_3|A_1) &= P(A_3|A_1 \cap A_2)\, P(A_2|A_1) + P(A_3|A_1 \cap \overline{A_2})\, P(\overline{A_2}|A_1) \\ &= \frac{2}{4} \cdot \frac{3}{5} + \frac{3}{4} \cdot \frac{2}{5} = \frac{12}{20} = \frac{3}{5}. \end{aligned}$$

Notiamo che risulta $P(A_3|A_1) = P(A_2|A_1) = \frac{3}{5}$, in analogia con il fatto che $P(A_3) = P(A_2) = \frac{4}{6} = \frac{2}{3}$.

Un concetto importante della teoria della probabilità è quello della indipendenza condizionata di eventi.

Diciamo che due eventi A e B sono *condizionatamente indipendenti* dato F se

$$P(A|B \cap F) = P(A|F)$$

o, equivalentemente, se

$$P(A \cap B|F) = P(A|F)\, P(B|F).$$

La nozione di indipendenza condizionata si può facilmente estendere a più di 2 eventi.

L'indipendenza di eventi non implica l'indipendenza condizionata, e viceversa, ossia

$$P(A \cap B) = P(A)\, P(B) \quad \not\Rightarrow \quad P(A \cap B|F) = P(A|F)\, P(B|F),$$

$$P(A \cap B|F) = P(A|F)\, P(B|F) \quad \not\Rightarrow \quad P(A \cap B) = P(A)\, P(B).$$

Esempio. Un'urna contiene 3 monete, che danno testa con probabilità $\frac{1}{2}$, $\frac{1}{4}$, $\frac{3}{4}$.
(i) Se si lancia una moneta scelta a caso, qual è la probabilità che esca testa?
(ii) Se la moneta scelta dà testa al primo lancio, qual è la probabilità che la stessa dia testa anche al secondo lancio?

Soluzione. Definiamo gli eventi $F_j = \{\text{si sceglie la moneta } j\text{-esima}\}$, $j = 1, 2, 3$ e $T_i = \{\text{esce testa al lancio } i\text{-esimo}\}$, $i = 1, 2$. (i) Per la formula delle alternative si ha:

$$P(T_1) = \sum_{j=1}^{3} P(T_1|F_j)\, P(F_j) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{1}{3} + \frac{3}{4} \cdot \frac{1}{3} = \frac{3}{2} \cdot \frac{1}{3} = \frac{1}{2}.$$

(ii) Se la moneta scelta dà testa al primo lancio, la probabilità che la stessa dia testa anche al secondo lancio è data da

$$P(T_2|T_1) = \frac{P(T_1 \cap T_2)}{P(T_1)}$$

dove

$$P(T_1 \cap T_2) = \sum_{j=1}^{3} P(T_1 \cap T_2|F_j)\, P(F_j).$$

È ragionevole assumere che i risultati di lanci distinti della medesima moneta siano condizionatamente indipendenti, ossia T_1 e T_2 sono condizionatamente indipendenti dato F_j, per ogni $j = 1, 2, 3$, quindi

$$P(T_1 \cap T_2 | F_1) = P(T_1|F_1)\, P(T_2|F_1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$
$$P(T_1 \cap T_2 | F_2) = P(T_1|F_2)\, P(T_2|F_2) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$$
$$P(T_1 \cap T_2 | F_3) = P(T_1|F_3)\, P(T_2|F_3) = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}.$$

Pertanto si ha

$$P(T_1 \cap T_2) = \sum_{j=1}^{3} P(T_1 \cap T_2|F_j)\, P(F_j) = \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{16} \cdot \frac{1}{3} + \frac{9}{16} \cdot \frac{1}{3} = \frac{7}{8} \cdot \frac{1}{3} = \frac{7}{24} = 0{,}2916.$$

Infine, ricordando che $P(T_1) = \dfrac{1}{2}$ si ottiene

$$P(T_2|T_1) = \frac{P(T_1 \cap T_2)}{P(T_1)} = \frac{7/24}{1/2} = \frac{7}{12} = 0{,}5833.$$

Notiamo che $P(T_2|T_1) \neq P(T_2) = \dfrac{1}{2}$, quindi T_1 e T_2 non sono indipendenti.

Esempio. In un laboratorio di analisi l'esame del sangue è efficace al 95% nell'individuare la presenza di una certa malattia. L'esame tuttavia rileva anche dei "falsi positivi" nell'1% delle persone sane. (Con probabilità 0,01 l'esame indica erroneamente come malata una persona sana). Se lo 0,5% della popolazione è affetto dalla malattia, qual è la probabilità che una persona risultata positiva all'esame abbia la malattia?

Soluzione. Definiamo gli eventi $F = \{\text{la persona ha la malattia}\}$ ed $E = \{\text{l'esame dà esito positivo}\}$. Dalle ipotesi segue:

$$P(F) = 0{,}005 \qquad P(\overline{F}) = 0{,}995 \qquad P(E|F) = 0{,}95 \qquad P(E|\overline{F}) = 0{,}01$$

e quindi

$$P(E) = P(E|F)\,P(F) + P(E|\overline{F})\,P(\overline{F}) = 0{,}95 \cdot 0{,}005 + 0{,}01 \cdot 0{,}995 = 0{,}0147.$$

Pertanto la probabilità richiesta è

$$P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{P(E|F)\,P(F)}{P(E)} = \frac{0{,}95 \cdot 0{,}005}{0{,}0147} = \frac{0{,}00475}{0{,}0147} \approx 0{,}323.$$

Esempio. Nell'esempio precedente, supponiamo che una persona si sottoponga 2 volte al test. Se la persona è risultata positiva entrambe le volte, qual è la probabilità che abbia la malattia?

Soluzione. Con riferimento agli eventi $F = \{$la persona scelta a caso ha la malattia$\}$ ed $E_k = \{$il k-esimo esame dà esito positivo$\}$, con $k = 1, 2$, dalle ipotesi fatte segue:

$$P(F) = 0{,}005 \qquad P(\overline{F}) = 0{,}995 \qquad P(E_k|F) = 0{,}95 \qquad P(E_k|\overline{F}) = 0{,}01.$$

Inoltre assumiamo che i 2 test siano *condizionatamente indipendenti*, ossia:

$$P(E_1 \cap E_2|F) = P(E_1|F)\, P(E_2|F) = (0{,}95)^2 = 0{,}9025$$
$$P(E_1 \cap E_2|\overline{F}) = P(E_1|\overline{F})\, P(E_2|\overline{F}) = (0{,}01)^2 = 0{,}0001$$

e quindi

$$\begin{aligned} P(E_1 \cap E_2) &= P(E_1 \cap E_2|F)\, P(F) + P(E_1 \cap E_2|\overline{F})\, P(\overline{F}) \\ &= 0{,}9025 \cdot 0{,}005 + 0{,}0001 \cdot 0{,}995 = 0{,}004612. \end{aligned}$$

Pertanto la probabilità richiesta è

$$P(F|E_1 \cap E_2) = \frac{P(E_1 \cap E_2|F)\, P(F)}{P(E_1 \cap E_2)} = \frac{0{,}9025 \cdot 0{,}005}{0{,}004612} = \frac{0{,}0045125}{0{,}004612} = 0{,}9784.$$

Esempio. Un sistema elettronico è composto da tre componenti $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$ che funzionano secondo le seguenti regole:

[R.1] $\mathcal{C}_1$ funziona con probabilità $1/2$;

[R.2] Se $\mathcal{C}_i$ funziona, allora $\mathcal{C}_{i+1}$ funziona con probabilità $(i+1)/6, \quad i = 1, 2;$

[R.3] Se $\mathcal{C}_i$ non funziona, allora $\mathcal{C}_{i+1}$ funziona con probabilità $i/6, \quad i = 1, 2.$

(i) Calcolare la probabilità che il componente $\mathcal{C}_i$ funzioni, $i = 2, 3$.

(ii) Se il componente $\mathcal{C}_{i+1}$ funziona, determinare la probabilità che $\mathcal{C}_i$ funzioni, $i = 1, 2$.

Soluzione. Per gli eventi $A_i = \{$il componente $\mathcal{C}_i$ funziona$\}$, $i = 1, 2, 3$, risulta

$$P(A_1) = \frac{1}{2}, \qquad P(A_2 \,|\, A_1) = \frac{2}{6} = \frac{1}{3}, \qquad P(A_3 \,|\, A_2) = \frac{3}{6} = \frac{1}{2},$$

$$P(A_2 \,|\, \overline{A_1}) = \frac{1}{6}, \qquad P(A_3 \,|\, \overline{A_2}) = \frac{2}{6} = \frac{1}{3}.$$

(i) Quindi, dalla formula delle alternative e da $P(\overline{A}_i) = 1 - P(A_i)$ si ricava

$$P(A_2) = P(A_2 \,|\, A_1) \cdot P(A_1) + P(A_2 \,|\, \overline{A_1}) \cdot P(\overline{A_1}) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{4},$$

$$P(A_3) = P(A_3 \,|\, A_2) \cdot P(A_2) + P(A_3 \,|\, \overline{A_2}) \cdot P(\overline{A_2}) = \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{3}{4} = \frac{3}{8}.$$

(ii) Usando la formula di Bayes si ottiene

$$P(A_1 \,|\, A_2) = \frac{P(A_2 \,|\, A_1) \cdot P(A_1)}{P(A_2)} = \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{4}} = \frac{2}{3},$$

$$P(A_2 \,|\, A_3) = \frac{P(A_3 \,|\, A_2) \cdot P(A_2)}{P(A_3)} = \frac{\frac{1}{2} \cdot \frac{1}{4}}{\frac{3}{8}} = \frac{1}{3}.$$

Esempio. Se $P(A) = P(B) \in (0,1)$, in quale caso risulta $P(A \,|\, \overline{B}) = P(\overline{A} \,|\, B)$?

Soluzione. Essendo

$$P(A \,|\, \overline{B}) = \frac{P(A \cap \overline{B})}{P(\overline{B})} = \frac{P(A) - P(A \cap B)}{P(\overline{B})},$$

$$P(\overline{A} \,|\, B) = \frac{P(\overline{A} \cap B)}{P(B)} = \frac{P(B) - P(A \cap B)}{P(B)},$$

dall'ipotesi $P(A) = P(B)$ segue che $P(A \,|\, \overline{B}) = P(\overline{A} \,|\, B)$ se $P(B) = P(\overline{B})$, ossia se

$$P(A) = P(B) = \frac{1}{2}.$$

Esempio. Un esperimento consiste nel generare a caso un vettore booleano di lunghezza n. Sapendo che il vettore contiene almeno un bit pari ad $\mathbf{1}$,
(i) qual è la probabilità che i primi due bit sono pari ad $\mathbf{1}$?
(ii) qual è la probabilità che i primi due bit sono pari a $\mathbf{0}$?

Soluzione. Siano $A = \{$il vettore contiene almeno un bit pari ad $\mathbf{1}\}$, $B = \{$i primi due bit del vettore sono pari ad $\mathbf{1}\}$ e $C = \{$i primi due bit del vettore sono pari ad $\mathbf{0}\}$. Si ha che $|S| = 2^n$, con

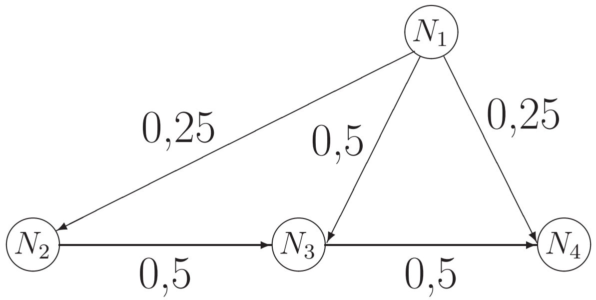
(i) $B \subset A$ e quindi $A \cap B = B$, con $P(\overline{A}) = \dfrac{1}{2^n}$, pertanto

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(A)} = \frac{\dfrac{2^{n-2}}{2^n}}{1 - \dfrac{1}{2^n}} = \frac{2^{n-2}}{2^n - 1}.$$

(ii)

$$P(C \mid A) = \frac{P(A \cap C)}{P(A)} = \frac{P(A|C)\,P(C)}{P(A)} = \frac{\left[1 - \dfrac{1}{2^{n-2}}\right]\dfrac{1}{4}}{1 - \dfrac{1}{2^n}} = \frac{2^{n-2} - 1}{2^n - 1}.$$

Esempio. Si consideri un sistema costituito da 4 unità collegate mediante archi orientati come in figura, dove su ogni arco è indicata la probabilità che esso sia attivo, indipendentemente dagli altri archi.



(i) Calcolare la probabilità che almeno uno dei tre percorsi da N_1 ad N_4 sia attivo.

(ii) Sapendo che almeno uno dei tre percorsi da N_1 ad N_4 è attivo, calcolare la probabilità che sia attivo il percorso $N_1 \to N_3 \to N_4$.

Soluzione. Definiamo gli eventi

$E_1 = \{\text{il percorso } N_1 \to N_4 \text{ è attivo}\}$,

$E_2 = \{\text{il percorso } N_1 \to N_3 \to N_4 \text{ è attivo}\}$,

$E_3 = \{\text{il percorso } N_1 \to N_2 \to N_3 \to N_4 \text{ è attivo}\}$.

Risulta

$$P(E_1) = \frac{1}{4} = 0{,}25; \quad P(E_2) = 0{,}5 \cdot 0{,}5 = \frac{1}{4} = 0{,}25;$$

$$P(E_3) = 0{,}5 \cdot 0{,}5 \cdot 0{,}25 = \frac{1}{16} = 0{,}0625.$$

(i) La probabilità che almeno uno dei tre percorsi da N_1 ad N_4 sia attivo è

$$\begin{aligned} P(E_1 \cup E_2 \cup E_3) &= P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_1 \cap E_3) \\ &\quad -P(E_2 \cap E_3) + P(E_1 \cap E_2 \cap E_3) \\ &= 0{,}25 + 0{,}25 + 0{,}0625 - 0{,}25 \cdot 0{,}25 - 0{,}25 \cdot 0{,}25 \cdot 0{,}5 \cdot 0{,}5 - 0{,}25 \cdot 0{,}25 \cdot 0{,}5 \\ &\quad +0{,}25 \cdot 0{,}25 \cdot 0{,}25 \cdot 0{,}5 = \frac{59}{128} = 0{,}4609. \qquad \text{(per l'indipendenza)} \end{aligned}$$

(ii) Sapendo che almeno uno dei tre percorsi da N_1 ad N_4 è attivo, la probabilità che sia attivo il percorso $N_1 \to N_3 \to N_4$ è

$$P(E_2 \,|\, E_1 \cup E_2 \cup E_3) = \frac{P(E_2 \cap (E_1 \cup E_2 \cup E_3))}{P(E_1 \cup E_2 \cup E_3)}$$

$$= \frac{P(E_2)}{P(E_1 \cup E_2 \cup E_3)} = \frac{0{,}25}{0{,}4609} = \frac{32}{59} = 0{,}5424.$$

Esempio. In un esperimento si generano sequenze di n bit casuali, tali che ogni bit, indipendentemente dagli altri, assume valore **1** oppure **0** con probabilità 1/2. Dati gli eventi $A = \{\text{almeno un bit assume valore } \mathbf{1}\}$, $B_k = \{\text{esattamente } k \text{ bit assumono valore } \mathbf{1}\}$ e $C_k = \{\text{i primi } k \text{ bit assumono valore } \mathbf{1}\}$, $k = 1, 2, \ldots, n$, calcolare

(i) $P(A)$, $P(B_k)$ e $P(C_k)$;

(ii) $P(A \cap B_k)$, $P(A \cap C_k)$ e $P(B_k \cap C_k)$;

(iii) $P(A \cup B_k \,|\, C_k)$ e $P(B_k \,|\, C_k)$.

Soluzione. Sia $k = 1, 2, \ldots, n$. Risulta, con $|S| = 2^n$,

(i)

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{1}{2^n},$$

$$P(B_k) = \frac{\binom{n}{k}}{2^n},$$

$$P(C_k) = \frac{2^{n-k}}{2^n} = \frac{1}{2^k}.$$

(ii) $B_k \subset A$ e quindi $A \cap B_k = B_k$, inoltre $C_k \subset A$ e quindi $A \cap C_k = C_k$, pertanto

$$P(A \cap B_k) = P(B_k) = \frac{\binom{n}{k}}{2^n},$$

$$P(A \cap C_k) = P(C_k) = \frac{1}{2^k},$$

$$P(B_k \cap C_k) = P(A \cap B_k \cap C_k) = \frac{1}{2^n}.$$

(iii) Notiamo che $C_k \subset A \cup B_k$ e quindi si ricava immediatamente

$$P(A \cup B_k \,|\, C_k) = \frac{P((A \cup B_k) \cap C_k)}{P(C_k)} = \frac{P(C_k)}{P(C_k)} = 1.$$

Inoltre si ha

$$P(B_k \,|\, C_k) = \frac{1}{2^{n-k}},$$

in quanto, affinché si realizzi B_k dato C_k, deve risultare che gli ultimi $n-k$ bit della sequenza assumano valore $\mathbf{0}$.

Esempio. Un negozio di elettronica riceve prodotti da tre fornitori F_1, F_2, F_3. Il 50% dei prodotti disponibili è fornito da F_1, il 20% da F_2 ed il 30% da F_3. Tuttavia, il 5% dei prodotti forniti da F_1 presenta dei difetti, mentre per F_2 e F_3 tale percentuale è pari a 3% e 6%, rispettivamente.

(i) Scelto a caso un prodotto, qual è la probabilità che sia difettoso?

(ii) Se un prodotto a caso è difettoso, qual è la probabilità che sia stato fornito da F_1?

Soluzione. Posto $F_i = \{$il prodotto proviene dal fornitore i-esimo$\}$, $i = 1, 2, 3$ e sia $A = \{$il prodotto è difettoso$\}$, si ha

$$P(F_1) = 0{,}5; \qquad P(F_2) = 0{,}2; \qquad P(F_3) = 0{,}3.$$

(i) Se si sceglie a caso un prodotto, la probabilità che sia difettoso è

$$\begin{aligned} P(A) &= P(A \,|\, F_1) \cdot P(F_1) + P(A \,|\, F_2) \cdot P(F_2) + P(A \,|\, F_3) \cdot P(F_3) \\ &= 0{,}05 \cdot 0{,}5 + 0{,}03 \cdot 0{,}2 + 0{,}06 \cdot 0{,}3 = 0{,}025 + 0{,}006 + 0{,}018 = 0{,}049. \end{aligned}$$

(ii) Se un prodotto scelto a caso è difettoso, la probabilità che sia stato fornito da F_1 è

$$P(F_1 \,|\, A) = \frac{P(A \,|\, F_1) P(F_1)}{P(A)} = \frac{0{,}05 \cdot 0{,}5}{0{,}049} = \frac{0{,}025}{0{,}049} = 0{,}5102.$$

Esempio. Si effettuano 3 trasmissioni indipendenti su di un canale di comunicazione rumoroso, per cui ad ogni trasmissione (indipendentemente dalle altre) l'output risulta essere **1** con probabilità 0,6 e **0** con probabilità 0,4. Posto

$A = \{$per almeno una delle 3 trasmissioni l'output è $\mathbf{0}\}$,

$B = \{$nelle prime 2 trasmissioni i valori in output sono diversi$\}$,

(i) stabilire se A e B sono indipendenti,

(ii) calcolare $P(\overline{A}|B)$ e $P(\overline{B}|A)$,

(iii) ricavare $P(\overline{A} \cup B)$.

Soluzione. (i) Sia $U_k = \{$nella trasmissione k-esima l'output è $\mathbf{1}\}$, con $P(U_k) = 0{,}6$ per $k = 1, 2, 3$. Poiché tali eventi sono indipendenti risulta

$$P(A) = P(\overline{U_1} \cup \overline{U_2} \cup \overline{U_3}) = P(\overline{U_1 \cap U_2 \cap U_3}) = 1 - P(U_1 \cap U_2 \cap U_3)$$

$$= 1 - P(U_1) \cdot P(U_2) \cdot P(U_3) = 1 - (0{,}6)^3 = 1 - 0{,}216 = 0{,}784.$$

Essendo $B = (U_1 \cap \overline{U_2}) \cup (\overline{U_1} \cap U_2)$, con $U_1 \cap \overline{U_2}$ e $\overline{U_1} \cap U_2$ incompatibili, si ha

$$P(B) = P(U_1 \cap \overline{U_2}) + P(\overline{U_1} \cap U_2) = 0{,}6 \cdot 0{,}4 + 0{,}4 \cdot 0{,}6 = 2 \cdot 0{,}6 \cdot 0{,}4 = 0{,}48.$$

Poiché $B \subset A$ si ha $A \cap B = B$, e quindi A e B non sono indipendenti, essendo

$$P(A \cap B) = P(B) = 0{,}48 \neq P(A) \cdot P(B).$$

(ii) Gli eventi $\overline{A}$ e B sono incompatibili, e quindi risulta

$$P(\overline{A}|B) = \frac{P(\overline{A} \cap B)}{P(B)} = 0.$$

Inoltre,

$$\begin{aligned} P(\overline{B}|A) &= \frac{P(A \cap \overline{B})}{P(A)} = \frac{P(\overline{U_1} \cap \overline{U_2}) + P(U_1 \cap U_2 \cap \overline{U_3})}{P(A)} \\ &= \frac{P(\overline{U_1}) \cdot P(\overline{U_2}) + P(U_1) \cdot P(U_2) \cdot P(\overline{U_3})}{P(A)} \\ &= \frac{0{,}4 \cdot 0{,}4 + 0{,}6 \cdot 0{,}6 \cdot 0{,}4}{0{,}784} = \frac{0{,}304}{0{,}784} = 0{,}3878. \end{aligned}$$

(iii) Poiché $\overline{A}$ e B sono incompatibili, dalla regola di additività finita e ricordando che $P(\overline{A}) = 1 - P(A) = 1 - 0,784 = 0{,}216$ si ha

$$P(\overline{A} \cup B) = P(\overline{A}) + P(B) = 0{,}216 + 0{,}48 = 0{,}696.$$

Esempio. Una lista contiene alcuni nominativi di dipendenti di un'azienda. Sappiamo che il 30% dei nominativi della lista sono dirigenti, e il restante 70% sono impiegati. Inoltre, il 50% dei nominativi della lista sono laureati, e il restante 50% sono diplomati. Infine, il 40% dei laureati della lista sono dirigenti.

Se si sceglie un nominativo della lista a caso, calcolare la probabilità che

(i) sia un dirigente diplomato,

(ii) sia un impiegato laureato,

(iii) sia un dirigente o un laureato.

(iv) Se si sceglie un nominativo a caso tra gli impiegati, qual è la probabilità che sia un laureato?

Soluzione. Sia $A = \{\text{un nominativo scelto a caso è di dirigente}\}$ e $B = \{\text{un nominativo scelto a caso è di laureato}\}$, cosicché

$$P(A) = 0{,}3 \qquad P(B) = 0{,}5 \qquad P(A|B) = 0{,}4.$$

Pertanto si ha

$$P(A \cap B) = P(A|B) \cdot P(B) = 0{,}4 \cdot 0{,}5 = 0{,}2.$$

(i) La probabilità che un nominativo scelto a caso sia di dirigente diplomato è

$$P(A \cap \overline{B}) = P(A) - P(A \cap B) = 0{,}3 - 0{,}2 = 0{,}1.$$

(ii) La probabilità che un nominativo scelto a caso sia di impiegato laureato è

$$P(\overline{A} \cap B) = P(B) - P(A \cap B) = 0{,}5 - 0{,}2 = 0{,}3.$$

(iii) La probabilità che un nominativo scelto a caso sia di un dirigente o un laureato è

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0{,}3 + 0{,}5 - 0{,}2 = 0{,}6.$$

(iv) Se si sceglie un nominativo a caso tra gli impiegati, la probabilità che sia un laureato è

$$P(B|\overline{A}) = \frac{P(\overline{A} \cap B)}{P(\overline{A})} = \frac{0{,}3}{0{,}7} = 0{,}4286.$$

	B	$\overline{B}$	
A	0,20	0,10	0,30
$\overline{A}$	0,30	0,40	0,70
	0,50	0,50	1

Notiamo che $P(A \cap B) \neq P(A)\,P(B)$ e quindi A e B non sono indipendenti.

Esempio. Da un'urna contenente 3 biglie bianche e 3 biglie nere si estraggono 3 biglie senza reinserimento.

(i) Calcolare la probabilità che tra le 3 estratte vi sia almeno una biglia bianca.

(ii) Se tra le 3 estratte vi è almeno una biglia bianca, qual è la probabilità che tutte le biglie estratte siano bianche?

(iii) Calcolare la probabilità che le 3 biglie estratte abbiano lo stesso colore.

(iv) Calcolare la probabilità che tra le 3 estratte vi sia almeno una biglia bianca sapendo che la prima biglia estratta è nera.

Soluzione. Poniamo $B_k = \{\text{si estraggono } k \text{ biglie bianche}\}$ per $k = 0, 1, 2, 3$. Tali eventi sono necessari e a 2 a 2 incompatibili. Si ha $|S| = \binom{6}{3} = 20$, e quindi

$$P(B_k) = \frac{\binom{3}{k}\binom{3}{3-k}}{20} = \begin{cases} 1/20 & \text{per } k = 0, 3 \\ 9/20 & \text{per } k = 1, 2. \end{cases}$$

(i) La probabilità che tra le 3 estratte vi sia almeno una biglia bianca è

$$P\left(\bigcup_{k=1}^{3} B_k\right) = \sum_{k=1}^{3} P(B_k) = 1 - P(B_0) = 1 - \frac{1}{20} = \frac{19}{20} = 0{,}95.$$

(ii) Se tra le 3 estratte vi è almeno una biglia bianca, la probabilità che tutte le biglie estratte siano bianche è

$$P(B_3|B_1 \cup B_2 \cup B_3) = \frac{P(B_3)}{P(B_1 \cup B_2 \cup B_3)} = \frac{1/20}{19/20} = \frac{1}{19} = 0{,}0526.$$

(iii) La probabilità che le 3 biglie estratte abbiano lo stesso colore è

$$P(B_0 \cup B_3) = P(B_0) + P(B_3) = \frac{1}{20} + \frac{1}{20} = \frac{1}{10} = 0{,}1.$$

(iv) La probabilità che tra le 3 estratte vi sia almeno una biglia bianca sapendo che la prima biglia estratta è nera (corrisponde ad avere almeno una biglia bianca quando si estraggono 2 biglie da un'urna contenente 3 biglie bianche e 2 biglie nere) è data da

$$1 - \frac{\binom{2}{2}\binom{3}{0}}{\binom{5}{2}} = 1 - \frac{1}{10} = 0{,}9.$$

Ponendo $A_k = \{\text{si estrae una bianca nella } k\text{-esima estrazione}\}$, si ha anche

$$P(A_2 \cup A_3|\overline{A}_1) = 1 - P(\overline{A}_2|\overline{A}_1)\, P(\overline{A}_3|\overline{A}_1 \cap \overline{A}_2) = 1 - \frac{2}{5} \cdot \frac{1}{4} = 1 - \frac{1}{10} = 0{,}9.$$

Esempio. Si lanciano indipendentemente 3 monete, di cui le prime 2 sono non truccate mentre la terza è truccata, nel senso che mostra testa con probabilità $1/3$ e croce con probabilità $2/3$. Posto $A = \{$esce testa almeno una volta nei 3 lanci$\}$ e $B = \{$nei primi 2 lanci si ha lo stesso risultato$\}$, stabilire se ognuna delle seguenti affermazioni è vera o falsa, giustificando le risposte:

(i) $P(\overline{A}) + P(\overline{B}) \leq 1$;

(ii) A e B sono indipendenti;

(iii) $P(B|A) = P(A \cup B)$;

(iv) $P(A|B) = 1 - P(A|\overline{B})$.

Soluzione. Sia $T_i = \{$esce testa al lancio i-esimo$\}$ $(i = 1, 2, 3)$, con $P(T_1) = P(T_2) = 1/2$ e $P(T_3) = 1/3$.

(i) Si ha quindi, per l'indipendenza:

$$P(\overline{A}) = P(\overline{T}_1 \cap \overline{T}_2 \cap \overline{T}_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{6}$$

$$P(\overline{B}) = P(T_1 \cap T_2) + P(\overline{T}_1 \cap \overline{T}_2) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2} \cdot \frac{1}{2}.$$

Pertanto, si ricava $P(\overline{A}) + P(\overline{B}) \leq 1$.

(ii) Gli eventi A e B non sono indipendenti in quanto risulta

$$P(A\cap B)=P(T_1\cap T_2)+P(\overline{T}_1\cap \overline{T}_2\cap T_3)=\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{3}=\frac{1}{3}\neq P(A)\,P(B)=\frac{5}{12},$$

essendo $P(A)=5/6$ e $P(B)=1/2$.

(iii) Si ha $P(B|A)\neq P(A\cup B)$ poiché

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{1/3}{5/6}=\frac{2}{5},$$

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)=\frac{5}{6}+\frac{1}{2}-\frac{1}{3}=1.$$

(iv) Risulta $P(A|B)\neq 1-P(A|\overline{B})$, essendo

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{1/3}{1/2}=\frac{2}{3},$$

$$P(A|\overline{B})=\frac{P(A\cap \overline{B})}{P(\overline{B})}=\frac{P(T_1\cap \overline{T}_2)+P(\overline{T}_1\cap T_2)}{P(\overline{B})}=\frac{1/2}{1/2}=1.$$

Esempio. Un esperimento consiste nel costruire a caso sequenze ordinate (ossia disposizioni) di lunghezza 3 senza ripetizioni dall'insieme $\{1, 2, 3, 4, 5\}$.

(i) Descrivere lo spazio campionario e determinarne la cardinalità.

(ii) Studiare le relazioni che riguardano l'indipendenza dei seguenti eventi:

$A = \{$i 3 numeri selezionati sono ordinati in senso crescente$\}$,

$B = \{$i 3 numeri selezionati sono primi$\}$,

$C = \{$i 3 numeri selezionati sono dispari$\}$.

(iii) Stabilire se le seguenti relazioni sono vere o false:

$$P(A \cup B) = P(A) + P(B), \qquad P(\overline{A} \cap \overline{B}) = P(\overline{A}) + P(\overline{B}), \qquad P(A|B) = P(B).$$

Soluzione. (i) Risulta $S = \{(x_1, x_2, x_3), x_i \in \{1, 2, 3, 4, 5\}, x_i \neq x_j, i \neq j\}$ e $|S| = (5)_3 = 5 \cdot 4 \cdot 3 = 60$.

(ii) Si ha

$$P(A) = \frac{\binom{5}{3}}{60} = \frac{10}{60} = \frac{1}{6}, \qquad P(B) = \frac{3!}{60} = \frac{6}{60} = \frac{1}{10}, \qquad P(C) = \frac{3!}{60} = \frac{6}{60} = \frac{1}{10}.$$

Inoltre si ha $A \cap B = \{(2,3,5)\}$, $A \cap C = \{(1,3,5)\}$, $B \cap C = \emptyset$, e pertanto

$$P(A \cap B) = \frac{1}{60}, \qquad P(A \cap C) = \frac{1}{60}, \qquad P(B \cap C) = 0, \qquad P(A \cap B \cap C) = 0.$$

Quindi i tre eventi non sono indipendenti, in quanto risulta

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B), \\ P(A \cap C) &= P(A) \cdot P(C), \\ P(B \cap C) &\neq P(B) \cdot P(C), \\ P(A \cap B \cap C) &\neq P(A) \cdot P(B) \cdot P(C). \end{aligned}$$

(iii) Essendo $P(A \cap B) = \dfrac{1}{60} \neq 0$ la relazione $P(A \cup B) = P(A) + P(B)$ non vale. Essendo

$$\begin{aligned} P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) \\ = 1 - P(A) - P(B) + P(A \cap B) = P(\overline{A}) + P(\overline{B}) - 1 + P(A \cap B), \end{aligned}$$

la relazione $P(\overline{A} \cap \overline{B}) = P(\overline{A}) + P(\overline{B})$ non è valida, in quanto $P(A \cap B) = \dfrac{1}{60} \neq 1$.
Infine

$$P(A \,|\, B) = \frac{P(A \cap B)}{P(B)} = \frac{1/60}{1/10} = \frac{1}{6} \neq P(B).$$

Esempio. Un canale di trasmissione è soggetto ad errore nel senso che ogni volta che si trasmette un bit, indipendentemente dalle altre trasmissioni, questo può essere modificato con la probabilità indicata nella seguente tabella:

bit trasmesso	bit ricevuto	probabilità
0	**0**	0,7
0	**?** (errore di tipo A)	0,2
0	**1** (errore di tipo B)	0,1
1	**0** (errore di tipo B)	0,2
1	**?** (errore di tipo A)	0,2
1	**1**	0,6

(i) Calcolare la probabilità che trasmettendo la sequenza binaria **001**

– si verifichi un solo errore, di tipo A,

– si verifichi un solo errore, di tipo B,

– si verifichi un solo errore, di tipo qualsiasi.

(ii) Se nel trasmettere la sequenza **001** si è verificato un solo errore, di tipo qualsiasi, qual è la probabilità che l'errore si sia verificato nella trasmissione del terzo bit?

Soluzione. (i) Poniamo, per $k = 1, 2, 3$,
$A_k = \{$si verifica un errore di tipo A nel trasmettere il bit k-esimo$\}$,
$B_k = \{$si verifica un errore di tipo B nel trasmettere il bit k-esimo$\}$,
$E_k = \{$si verifica un errore di tipo qualsiasi nel trasmettere il bit k-esimo$\}$.
Risulta:

$$P(A_1) = P(A_2) = P(A_3) = 0{,}2 \qquad P(B_1) = P(B_2) = 0{,}1 \qquad P(B_3) = 0{,}2.$$

Inoltre $E_k = A_k \cup B_k$. Quindi, trasmettendo la sequenza binaria **001**, la probabilità che si verifichi un solo errore, di tipo A, è

$$P(I_A) = P(A_1 \cap \overline{E_2} \cap \overline{E_3}) + P(\overline{E_1} \cap A_2 \cap \overline{E_3}) + P(\overline{E_1} \cap \overline{E_2} \cap A_3)$$
$$= 0{,}2 \cdot 0{,}7 \cdot 0{,}6 + 0{,}7 \cdot 0{,}2 \cdot 0{,}6 + 0{,}7 \cdot 0{,}7 \cdot 0{,}2 = 0{,}266$$

avendo usato l'indipendenza degli eventi. Analogamente, trasmettendo la sequenza binaria **001**, la probabilità che si verifichi un solo errore, di tipo B, è

$$P(I_B) = P(B_1 \cap \overline{E_2} \cap \overline{E_3}) + P(\overline{E_1} \cap B_2 \cap \overline{E_3}) + P(\overline{E_1} \cap \overline{E_2} \cap B_3)$$
$$= 0{,}1 \cdot 0{,}7 \cdot 0{,}6 + 0{,}7 \cdot 0{,}1 \cdot 0{,}6 + 0{,}7 \cdot 0{,}7 \cdot 0{,}2 = 0{,}182.$$

Pertanto, la probabilità che si verifichi un solo errore, di tipo qualsiasi, è dato dalla somma delle probabilità precedenti, trattandosi di eventi incompatibili, quindi:

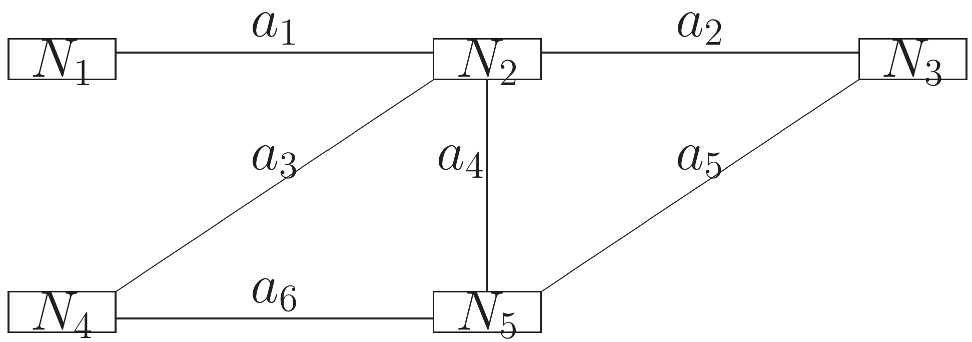
$$P(I_A \cup I_B) = P(I_A) + P(I_B) = 0{,}266 + 0{,}182 = 0{,}448.$$

(ii) Se nel trasmettere la sequenza **001** si è verificato un solo errore, di tipo qualsiasi, la probabilità che l'errore si sia verificato nella trasmissione del terzo bit si ottiene usando la formula di Bayes:

$$P(\overline{E_1} \cap \overline{E_2} \cap E_3 | I_A \cup I_B) = \frac{P(\overline{E_1} \cap \overline{E_2} \cap E_3)}{P(I_A \cup I_B)} = \frac{0{,}7 \cdot 0{,}7 \cdot 0{,}4}{0{,}448} = \frac{0{,}196}{0{,}448} = 0{,}4375$$

avendo notato che $\overline{E_1} \cap \overline{E_2} \cap E_3 \subset I_A \cup I_B$, ossia che l'occorrenza di un solo errore si realizzi nella trasmissione del terzo bit implica che si verifichi un solo errore di tipo qualsiasi.

Esempio. Si consideri l'esperimento che consiste nello scegliere a caso una coppia di archi del seguente grafo:



(i) Determinare la cardinalità dello spazio campionario.

(ii) Calcolare la probabilità degli eventi $C_k = \{$almeno uno dei 2 archi scelti è connesso al nodo $N_k\}$, per $k = 1, 2, \ldots, 5$.

(iii) Cosa si può dire sull'indipendenza di C_1 e C_2, e di C_1 e C_3?

Soluzione. (i) La scelta a caso di una coppia di archi da un insieme di cardinalità 6 si effettua in $\binom{6}{2} = \dfrac{6 \cdot 5}{2} = 15$ modi, quindi lo spazio campionario ha cardinalità 15.

(ii) Essendo $C_k = \{$almeno uno dei 2 archi scelti è connesso al nodo $N_k\}$, si ha

$$P(C_1) = 1 - \frac{\binom{5}{2}}{15} = 1 - \frac{10}{15} = \frac{5}{15} = \frac{1}{3}, \qquad P(C_2) = 1 - \frac{\binom{2}{2}}{15} = \frac{14}{15},$$

$$P(C_3) = 1 - \frac{\binom{4}{2}}{15} = \frac{9}{15} = \frac{3}{5},$$

$$P(C_4) = 1 - \frac{\binom{4}{2}}{15} = \frac{9}{15} = \frac{3}{5}, \qquad P(C_5) = 1 - \frac{\binom{3}{2}}{15} = \frac{12}{15} = \frac{4}{5}.$$

(iii) Si ha $C_1 \subset C_2$ e $C_1 \cap C_3 = \{(a_1, a_2), (a_1, a_5)\}$, pertanto

$$P(C_1 \cap C_2) = \frac{5}{15} = \frac{1}{3} \neq P(C_1)P(C_2) = \frac{1}{3} \cdot \frac{14}{15},$$

$$P(C_1 \cap C_3) = \frac{2}{15} \neq P(C_1)P(C_3) = \frac{1}{3} \cdot \frac{3}{5},$$

quindi non sussiste indipendenza.

Esercizi per casa

3.1) Si consideri l'esperimento che consiste nell'estrarre 2 biglie da un'urna che contiene 1 biglia azzurra, 2 bianche e 3 rosse. Posto $A = \{\text{le 2 biglie estratte non sono rosse}\}$, $B = \{\text{almeno una delle 2 biglie estratte è bianca}\}$, $C = \{\text{una delle 2 biglie estratte è azzurra}\}$, studiare l'indipendenza di tali eventi, nel caso di estrazioni
(i) con reinserimento,
(ii) senza reinserimento.

3.2) Una moneta non truccata viene lanciata 4 volte. Calcolare la probabilità
(i) che esca testa almeno 2 volte,
(ii) che esca testa almeno 2 volte, sapendo che esce testa al primo lancio,
(iii) che esca testa al primo lancio, sapendo che esce testa almeno 2 volte nei 4 lanci.

3.3) Si estraggono le biglie in sequenza da un'urna contenente n biglie numerate da 1 a n. Si ha concordanza all'estrazione k-esima se in tale estrazione esce la biglia numero k, per $k = 1, 2, \dots, n$.
(i) Se nella prima estrazione si verifica concordanza, qual è la probabilità che si abbia concordanza nella seconda estrazione?
(ii) Se nella seconda estrazione si verifica concordanza, qual è la probabilità che si abbia concordanza nella terza estrazione?

3.4) Da un'urna contenente $2n+1$ biglie numerate da $-n$ a n si estraggono due biglie a caso.
(i) Calcolare la probabilità che il prodotto dei due numeri estratti sia positivo, sia nullo, sia negativo.
(ii) Verificare che la somma delle tre probabilità è 1.

3.5) Se si effettua un vaccino antinfluenzale, la probabilità di contrarre l'influenza risulta pari a 0,3, mentre tale probabilità aumenta a 0,5 se si sceglie di non effettuare il vaccino.
(i) Sapendo che 6 individui su 10 si vaccinano, quanto vale la probabilità di contrarre l'influenza?
(ii) Se un individuo contrae l'influenza, qual è la probabilità che si era vaccinato?

3.6) Una lista contiene 9 nominativi (3 di impiegati, 4 di dirigenti, 2 di commerciali). Per errore si cancellano 3 nominativi dalla lista.
(i) Calcolare la probabilità che i 3 nominativi siano tutti della stessa categoria.
(ii) Calcolare la probabilità che tra i 3 nominativi vi sia al più un commerciale.
(iii) Stabilire se gli eventi su considerati sono indipendenti.

CAPITOLO 4 – Variabili aleatorie

4.1 Variabili aleatorie

Problema. Un sistema è costituito da n unità numerate da 1 a n. In ogni unità si sceglie a caso un numero tra 1 e n, indipendentemente dalle altre. Diciamo che si ha concordanza nell'unità j-esima se ivi si sceglie il numero j, per $j = 1, 2, \ldots, n$.

Qual è la probabilità che non vi siano concordanze?
Qual è la probabilità che vi siano k concordanze?

Posto $A_k = \{\text{vi sono } k \text{ concordanze}\}$, gli eventi $A_0, A_1, \ldots, A_n$ sono necessari e a 2 a 2 incompatibili. In virtù dell'indipendenza le probabilità richieste sono:

$$P(A_0) = \left(1 - \frac{1}{n}\right)^n, \qquad P(A_k) = \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k}, \quad k = 1, 2, \ldots, n.$$

Notiamo che $\lim_{n\to\infty} P(A_0) = \lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1} = 0{,}3679.$

Se si definisce $X =$"numero di concordanze", appare più adeguato esprimere le probabilità richieste come $P(X = 0) = P(A_0)$ e $P(X = k) = P(A_k)$.

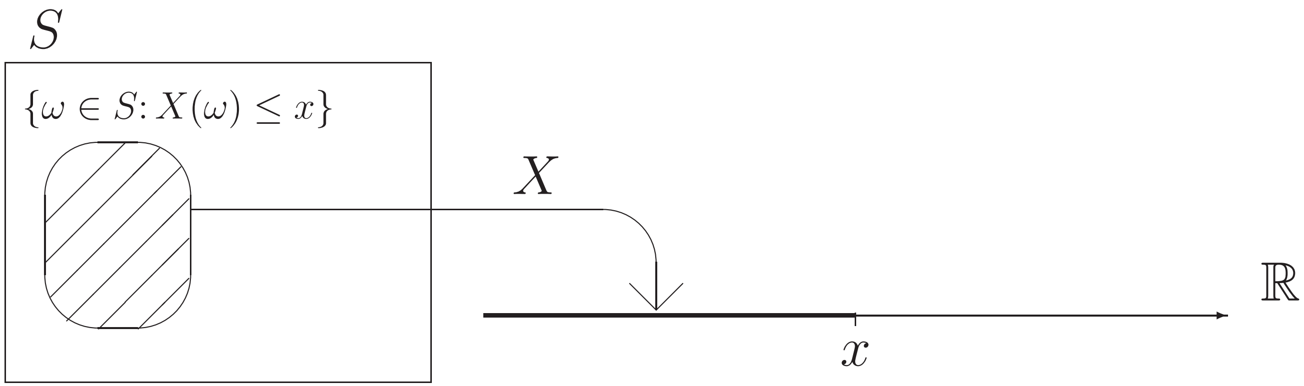
Nella tabella sono riportati i valori delle probabilità di X per varie scelte di n.

	$n=2$	$n=3$	$n=4$	$n=5$
$P(X=0)$	0,25	0,2963	0,3164	0,3277
$P(X=1)$	0,50	0,4444	0,4219	0,4096
$P(X=2)$	0,25	0,2222	0,2109	0,2048
$P(X=3)$		0,0370	0,0469	0,0512
$P(X=4)$			0,0039	0,0064
$P(X=5)$				0,0003

Talora studiando un fenomeno aleatorio l'interesse va riposto su una funzione $X(\omega)$ dell'esito ω dell'esperimento. Ad esempio, lanciando due dadi siamo interessati alla somma dei due valori, oppure lanciando n volte una moneta possiamo riferirci al numero totale di teste. La quantità d'interesse, o più precisamente, queste funzioni a valori reali definite sullo spazio campionario sono note come *variabili aleatorie*.

Definizione. Dato uno spazio di probabilità $(S, \mathcal{F}, P)$, una funzione $X\colon S \to \mathbb{R}$ è detta variabile aleatoria se risulta

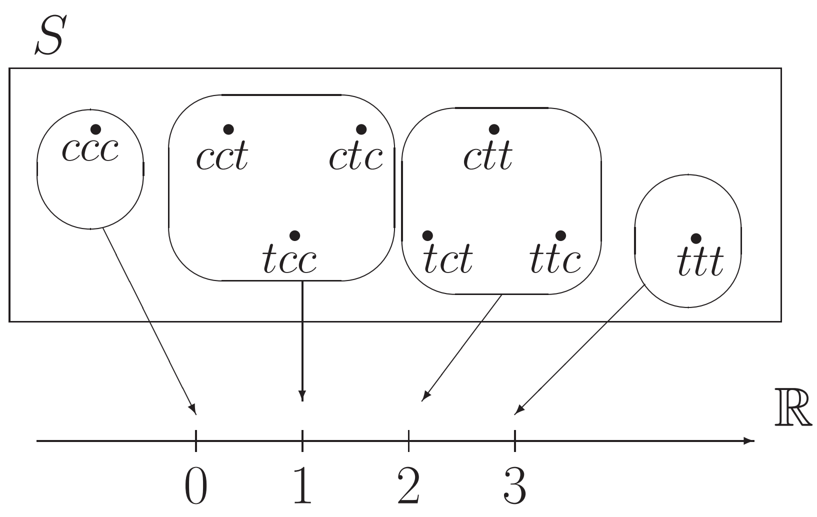
$$\{X \le x\} := \{\omega \in S\colon X(\omega) \le x\} \in \mathcal{F} \qquad \text{per ogni } x \in \mathbb{R}.$$



Poiché il valore di una variabile aleatoria è determinato dall'esito dell'esperimento, possiamo assegnare le probabilità ai possibili valori ottenuti dalla variabile aleatoria.

Esempio. Sia Y il numero di teste che si ottengono nel lancio di 3 monete non truccate. Per $S = \{ccc, cct, ctc, tcc, ctt, tct, ttc, ttt\}$, si ha che $Y : S \to \mathbb{R}$ è una variabile aleatoria che assume i valori $0, 1, 2, 3$ con le seguenti probabilità:

$$
\begin{aligned}
P(Y = 0) &= P(\{ccc\}) = \frac{1}{8} \\
P(Y = 1) &= P(\{cct, ctc, tcc\}) = \frac{3}{8} \\
P(Y = 2) &= P(\{ctt, tct, ttc\}) = \frac{3}{8} \\
P(Y = 3) &= P(\{ttt\}) = \frac{1}{8}
\end{aligned}
$$



Gli eventi $\{Y = k\}_{k=0,1,2,3}$ sono necessari e a 2 a 2 incompatibili. Quindi, poiché Y deve assumere uno e uno solo tra i valori $0, 1, 2, 3$ abbiamo

$$P\left(\bigcup_{k=0}^{3} \{Y = k\}\right) = \sum_{k=0}^{3} P(Y = k) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1.$$

Esempio. Si estraggono 3 biglie a caso senza reinserimento da un'urna contenente venti biglie numerate da 1 a 20. Qual è la probabilità che almeno una tra le biglie estratte abbia un numero maggiore o uguale a 17?

Soluzione. Se denotiamo con X il maggiore tra i 3 numeri estratti, X è una variabile aleatoria che assume i valori $3, 4, \ldots, 20$. Inoltre, supponendo che ognuna delle $\binom{20}{3}$ possibili terne abbia uguale probabilità, si ha

$P(X=i)$	$P(X=17)$	$P(X=18)$	$P(X=19)$	$P(X=20)$
$\dfrac{\binom{i-1}{2}\binom{1}{1}\binom{20-i}{0}}{\binom{20}{3}}$	$\dfrac{\binom{16}{2}}{\binom{20}{3}}=\dfrac{2}{19}$	$\dfrac{\binom{17}{2}}{\binom{20}{3}}=\dfrac{34}{285}$	$\dfrac{\binom{18}{2}}{\binom{20}{3}}=\dfrac{51}{380}$	$\dfrac{\binom{19}{2}}{\binom{20}{3}}=\dfrac{3}{20}$

Infatti il numero di terne che compongono l'evento $\{X=i\}$ è il numero di terne per cui una biglia ha numero i e le altre due hanno numero compreso tra 1 e $i-1$. Si ha

$$P(X \geq 17) = \sum_{i=17}^{20} P(X=i) = \frac{2}{19} + \frac{34}{285} + \frac{51}{380} + \frac{3}{20} = 0{,}508.$$

Esempio. Si estraggono 3 biglie a caso con reinserimento da un'urna contenente venti biglie numerate da 1 a 20. Qual è la probabilità che almeno una tra le biglie estratte abbia un numero maggiore o uguale a 17?

Soluzione. Sia Y la variabile aleatoria che descrive quante delle 3 biglie estratte abbiano un numero maggiore o uguale a 17. Tale variabile assume i valori 0, 1, 2, 3 e descrive il numero di successi in $n = 3$ prove indipendenti, dove la probabilità p di successo in ogni prova è la probabilità di estrarre un numero maggiore o uguale a 17:

$$p = \frac{4}{20} = \frac{1}{5}.$$

Si ha quindi una distribuzione di probabilità di tipo binomiale

$$P(Y = k) = \binom{3}{k} \left(\frac{1}{5}\right)^k \left(\frac{4}{5}\right)^{3-k}, \qquad k = 0, 1, 2, 3.$$

Pertanto risulta $P(Y = 0) = \left(\frac{4}{5}\right)^3$. Ne segue che la probabilità richiesta è

$$P(Y \geq 1) = \sum_{k=1}^{3} P(Y = k) \;=\; 1 - P(Y = 0) = 1 - \left(\frac{4}{5}\right)^3 = 1 - 0{,}512 = 0{,}488.$$

Esempio. Una moneta truccata dà testa con probabilità p in ogni lancio. La lanciamo ripetutamente fino a che non appaia testa per la prima volta oppure si siano fatti n lanci. Sia X la variabile aleatoria che descrive il numero totale di volte che lanciamo la moneta, e che assume valori $1, 2, \ldots, n$. Se C_i e $T_i = \overline{C_i}$ rappresentano fuoriuscita di croce e testa al lancio i-esimo, facendo uso dell'indipendenza nei lanci si ha

$$\begin{aligned}
P(X = 1) &= P(T_1) = p \\
P(X = k) &= P(C_1 \cap C_2 \cap \ldots \cap C_{k-1} \cap T_k) = (1-p)^{k-1}\, p \qquad (k = 2, 3, \ldots, n-1) \\
P(X = n) &= P(C_1 \cap C_2 \cap \ldots \cap C_n) + P(C_1 \cap C_2 \cap \ldots \cap C_{n-1} \cap T_n) \\
&= P(C_1 \cap C_2 \cap \ldots \cap C_{n-1}) = P(C_1)P(C_2) \cdots P(C_{n-1}) = (1-p)^{n-1},
\end{aligned}$$

$$\begin{aligned}
\sum_{k=1}^{n} P(X = k) &= \sum_{k=1}^{n-1} (1-p)^{k-1}\, p + (1-p)^{n-1} \\
&= p \sum_{r=0}^{n-2} (1-p)^r + (1-p)^{n-1} = p \left[\frac{1 - (1-p)^{n-1}}{1 - (1-p)}\right] + (1-p)^{n-1} = 1,
\end{aligned}$$

avendo posto $r = k - 1$ e avendo ricordato che $\sum_{j=0}^{m} c^j = \dfrac{1 - c^{m+1}}{1 - c}$ per $c \neq 1$.

4.2 Funzioni di distribuzione

Definizione. Data una variabile aleatoria X, la funzione definita da

$$F(x) = P(X \leq x), \qquad -\infty < x < \infty$$

è detta *funzione di distribuzione* (o *di ripartizione*) di X.

Quindi, la funzione di distribuzione $F(x)$ di una variabile aleatoria X rappresenta la probabilità che X assuma un valore minore o uguale a x, per ogni $x \in \mathbb{R}$.

Proposizione. Una funzione di distribuzione $F(x)$ è caratterizzata dalle proprietà:

1. $F(x)$ è una funzione monotona non decrescente, ovvero $F(a) \leq F(b)$ se $a < b$.
2. $\lim\limits_{x\to\infty} F(x) = 1$.
3. $\lim\limits_{x\to-\infty} F(x) = 0$.
4. $F(x)$ è continua a destra, ossia $\lim\limits_{h\to 0^+} F(x+h) = F(x)$ per ogni $x \in \mathbb{R}$ fissato.

Queste proprietà contraddistinguono una funzione di distribuzione, nel senso che se una funzione $F(x)$ soddisfa tali proprietà allora esiste una variabile aleatoria che ammette $F(x)$ come funzione di distribuzione.

Dimostrazione. La Proprietà 1 segue dal fatto che se $a < b$, allora

$$\{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\},$$

con i 2 eventi a secondo membro incompatibili. Per la proprietà di additività finita:

$$F(b) = P(X \leq b) = P(X \leq a) + P(a < X \leq b) \geq F(a).$$

La Proprietà 2 si dimostra notando che se $\{x_n\}_{n\geq 1}$ è una successione di reali che cresce verso ∞, allora gli eventi $\{X \leq x_n\}$ formano una successione crescente di eventi il cui limite è $\{X < \infty\}$. Quindi, per la proprietà di continuità della probabilità, si ha

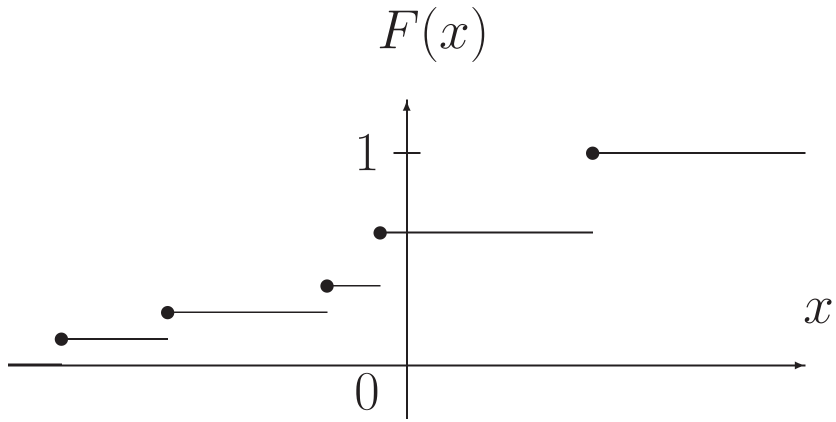
$$\lim_{x\to\infty} F(x) = \lim_{n\to\infty} P(X \leq x_n) = P(X < \infty) = 1.$$

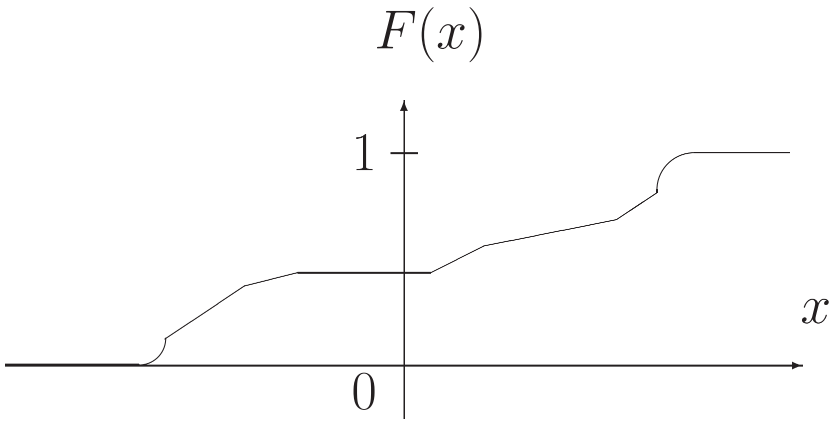
La Proprietà 3 si dimostra in modo analogo alla Proprietà 2, ed è lasciato come esercizio. Per dimostrare la Proprietà 4 notiamo che se $\{x_n\}_{n\geq 1}$ è una successione decrescente che converge a $x \in \mathbb{R}$, allora $\{X \leq x_n\}$, $n \geq 1$, costituisce una successione di eventi decrescente il cui limite coincide con $\{X \leq x\}$. Quindi, per la proprietà di continuità,

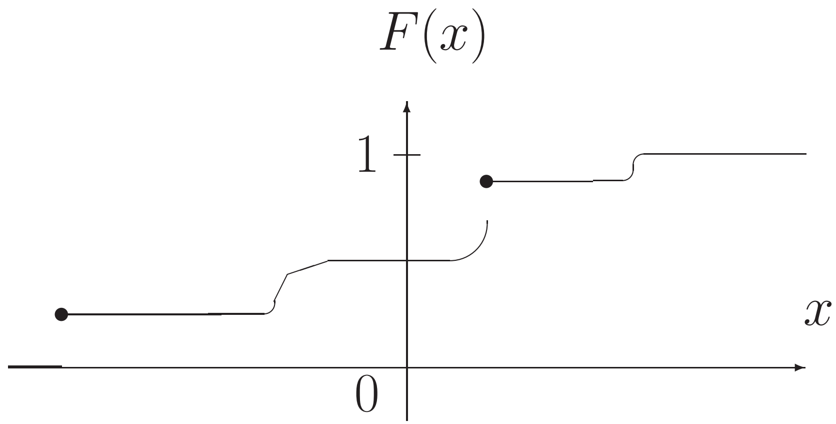
$$\lim_{h\to 0^+} F(x+h) = \lim_{n\to\infty} P(X \leq x_n) = P(X \leq x) = F(x).$$

Proprietà di $F(x) = P(X \leq x)$:

1. $F(x)$ è non decrescente.
2. $\lim_{x \to \infty} F(x) = 1$.
3. $\lim_{x \to -\infty} F(x) = 0$.
4. $F(x)$ è continua a destra.







Esercizio. Stabilire se le seguenti funzioni sono funzioni di distribuzione:

(i)

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

(ii)

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{2}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

(iii)

$$F(x) = \begin{cases} 0, & x \leq -1 \\ \dfrac{x+1}{2}, & -1 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

Proposizione. Sia $F(x)$ la funzione di distribuzione di una variabile aleatoria X. Allora,

(i) $P(a < X \leq b) = F(b) - F(a)$ per ogni $a < b$.

(ii) $P(X < b) = \lim_{n\to\infty} F(x_n)$ per ogni $b \in \mathbb{R}$ fissato e per ogni successione crescente $\{x_n\}_{n\geq 1}$ che converge a b, e quindi

$$P(X < x) = \lim_{h\to 0^-} F(x+h), \qquad \text{per ogni } x \in \mathbb{R}.$$

Dimostrazione. Nella Proposizione precedente abbiamo visto che se $a < b$, allora

$$\{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\},$$

e quindi risulta $F(b) = F(a) + P(a < X \leq b)$, da cui segue la (i).

Per dimostrare la (ii) notiamo che se $\{x_n\}_{n\geq 1}$ è una successione crescente che converge a b, allora dalla proprietà di continuità otteniamo:

$$P(X < b) = P\left(\lim_{n\to\infty} \{X \leq x_n\}\right) = \lim_{n\to\infty} P(X \leq x_n) = \lim_{n\to\infty} F(x_n).$$

Una funzione di distribuzione $F(x)$ non è necessariamente continua. Infatti, limite sinistro e limite destro di $F(x)$ in $x \in \mathbb{R}$ non necessariamente coincidono, essendo

$$F(x^-) := \lim_{h \to 0^-} F(x+h) \leq \lim_{h \to 0^+} F(x+h) = F(x).$$

Notiamo il significato probabilistico di limite sinistro e limite destro di $F(x)$:

$$F(x^-) = P(X < x), \qquad F(x) = P(X \leq x).$$

Da ciò segue che

$$P(X = x) = P(X \leq x) - P(X < x) = F(x) - F(x^-).$$

Quindi, se $F(x)$ è continua in x allora risulta

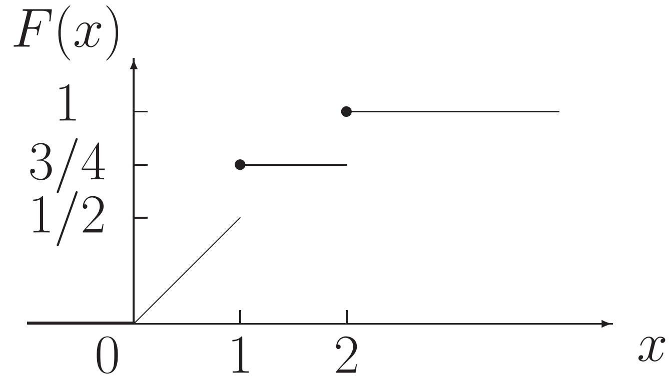
$$P(X = x) = F(x) - F(x^-) = 0.$$

Al contrario, se $F(x)$ è discontinua in x allora si ha

$$P(X = x) = F(x) - F(x^-) > 0.$$

Esempio. Sia X una variabile aleatoria avente funzione di distribuzione

$$F(x) = \begin{cases} 0, & x < 0 \\ x/2, & 0 \le x < 1 \\ 3/4, & 1 \le x < 2 \\ 1, & x \ge 2. \end{cases}$$



Calcolare (a) $P(X < 2)$, (b) $P(X = 1)$, (c) $P(X > 1/2)$, (d) $P(1 < X \le 3)$, (e) $P(1 \le X \le 2)$, (f) $P(X > 1|X > 1/2)$.

Soluzione. Si ha

(a) $P(X < 2) = F(2^-) = 3/4$;

(b) $P(X = 1) = F(1) - F(1^-) = 3/4 - 1/2 = 1/4$;

(c) $P(X > 1/2) = 1 - P(X \le 1/2) = 1 - F(1/2) = 1 - 1/4 = 3/4$;

(d) $P(1 < X \le 3) = F(3) - F(1) = 1 - 3/4 = 1/4$;

(e) $P(1 \le X \le 2) = P(X \le 2) - P(X < 1) = F(2) - F(1^-) = 1 - 1/2 = 1/2$;

(f) $P(X > 1|X > 1/2) = \dfrac{P(X > 1)}{P(X > 1/2)} = \dfrac{1 - P(X \le 1)}{3/4} = \dfrac{1 - F(1)}{3/4} = \dfrac{1/4}{3/4} = \dfrac{1}{3}.$

4.3 Variabili aleatorie discrete

Una variabile aleatoria X che può assumere al più un'infinità numerabile di valori è detta *discreta*. In tal caso definiamo la *densità discreta* (o *funzione di probabilità*) $p(k)$ di X come

$$p(k) = P(X = k).$$

La densità discreta $p(k)$ è positiva al più per un'infinità numerabile di valori di k. Quindi, se X assume i valori $x_1, x_2, \ldots$, allora

$$p(x_i) \geq 0 \qquad i = 1, 2, \ldots,$$

$$p(x) = 0 \qquad \text{altrimenti.}$$

Poiché X deve assumere almeno uno dei valori x_i, si ha che somma dei valori della funzione di probabilità $p(x)$ su tutti i valori x_i è unitaria, ossia

$$\sum_{i=1}^{\infty} p(x_i) = \sum_{i=1}^{\infty} P(X = x_i) = P\left(\bigcup_{i=1}^{\infty}\{X = x_i\}\right) = 1.$$

Può essere utile rappresentare la densità discreta in forma grafica ponendo i valori x_i in ascissa e $p(x_i)$ in ordinata.

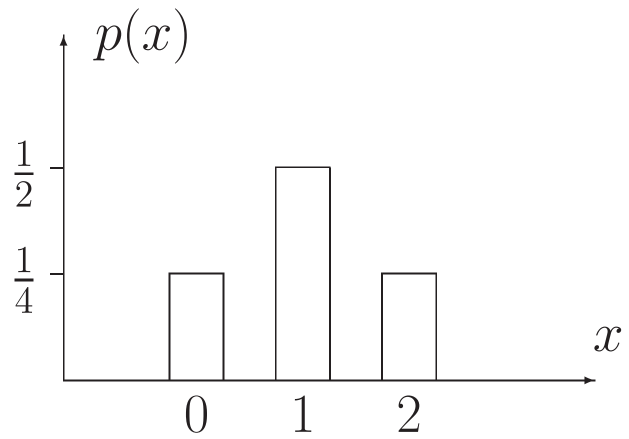
Esempio. Se X è la variabile aleatoria discreta che descrive quante volte esce testa nel lancio di due monete non truccate, allora la densità discreta di X è

$$p(k) = P(X = k) = \binom{2}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{2-k}, \qquad k = 0, 1, 2$$

ossia

$$p(0) = \frac{1}{4}, \qquad p(1) = \frac{1}{2}, \qquad p(2) = \frac{1}{4}.$$

Graficamente si ha



La densità discreta consente di calcolare la probabilità che la variabile aleatoria discreta assuma valori in un sottoinsieme qualsiasi B di $\mathbb{R}$; ad esempio nell'intervallo $[a, b]$:

$$P(X \in B) = \sum_{k\colon x_k \in B} p(x_k), \qquad P(a \leq X \leq b) = \sum_{k\colon a \leq x_k \leq b} p(x_k).$$

In particolare, se $B = (-\infty, x]$ si ha che la funzione di distribuzione di una variabile aleatoria discreta X può essere così espressa:

$$F(x) = P(X \leq x) = P(X \in (-\infty, x]) = \sum_{k\colon x_k \leq x} p(x_k), \qquad x \in \mathbb{R}.$$

Esempio. Sia X la variabile aleatoria discreta che descrive il numero di bit pari a **1** in un vettore booleano, di lunghezza n, scelto a caso.

(a) Calcolare $p(k) = P(X = k)$, per $k = 0, 1, \ldots, n$.

(b) Verificare che $\sum_{k=0}^{n} p(k) = 1$.

(c) Calcolare $P(X \geq 1)$.

Soluzione. (a) Il rapporto tra il numero di vettori booleani con k bit pari a **1** ed il numero di vettori booleani di lunghezza n fornisce

$$p(k) = P(X = k) = \frac{\binom{n}{k}}{2^n}, \qquad k = 0, 1, \ldots, n.$$

(b) Pertanto,

$$\sum_{k=0}^{n} p(k) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k} = \frac{1}{2^n} \cdot 2^n = 1.$$

(c) Si ha infine

$$P(X \geq 1) = 1 - P(X < 1) = 1 - p(0) = 1 - \frac{\binom{n}{0}}{2^n} = 1 - \frac{1}{2^n}.$$

Esempio. Sia X una variabile aleatoria discreta che assume un'infinità numerabile di valori, con $p(k) = c\,\dfrac{\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$, dove λ è una costante positiva.

(a) Ricavare c in termini di λ.

(b) Si calcoli $P(X = 0)$ e $P(X > 2)$.

Soluzione. (a) Per determinare c, imponendo che sia $\displaystyle\sum_{k=0}^{\infty} p(k) = 1$ abbiamo

$$1 = c \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = c\, e^{\lambda} \qquad \Longrightarrow \qquad c = e^{-\lambda},$$

avendo ricordato che $\boxed{\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x}$. Si ha quindi $p(k) = e^{-\lambda}\,\dfrac{\lambda^k}{k!}$, $k \geq 0$.

(b) Pertanto risulta $P(X = 0) = p(0) = e^{-\lambda}\,\lambda^0/0! = e^{-\lambda}$,

$P(X > 2) = 1 - P(0 \leq X \leq 2) = 1 - [p(0) + p(1) + p(2)] = 1 - e^{-\lambda}\left(1 + \lambda + \dfrac{\lambda^2}{2}\right).$

Esempio. Determinare la densità della variabile aleatoria discreta X, che descrive il massimo che si ottiene lanciando 2 dadi a caso.

Soluzione. Notiamo che $X = 1$ se esce la coppia $(1, 1)$, $X = 2$ se esce $(1, 2)$, $(2, 1)$ oppure $(2, 2)$, e così via. Posto $p(k) = P(X = k)$ per $k = 1, \ldots, 6$, essendo $|S| = 36$, risulta quindi

$$p(1) = \frac{1}{36}, \quad p(2) = \frac{3}{36}, \quad p(3) = \frac{5}{36}, \quad p(4) = \frac{7}{36}, \quad p(5) = \frac{9}{36}, \quad p(6) = \frac{11}{36},$$

ossia

$$p(k) = \frac{2k-1}{36}, \qquad k = 1, 2, \ldots, 6.$$

Notiamo che risulta

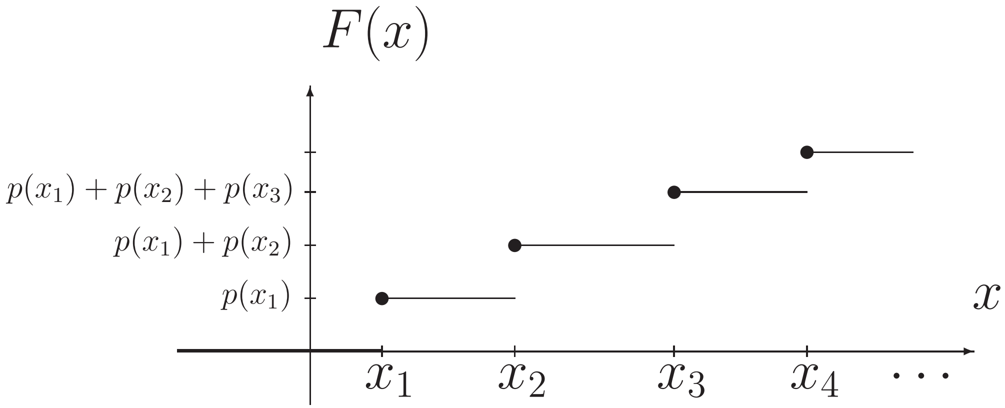
$$\sum_{k=1}^{6} p(k) = \sum_{k=1}^{6} \frac{2k-1}{36} = \frac{1}{18} \sum_{k=1}^{6} k - \frac{1}{6} = \frac{1}{18} \frac{6 \cdot 7}{2} - \frac{1}{6} = \frac{7}{6} - \frac{1}{6} = 1,$$

avendo ricordato che

$$\boxed{\sum_{k=1}^{n} k = \frac{n(n+1)}{2}}$$

Se X è una variabile aleatoria discreta che assume i valori $x_1, x_2, x_3, \ldots$, dove $x_1 < x_2 < x_3 < \ldots$, allora la sua funzione di distribuzione $F(x)$ è costante negli intervalli $[x_{i-1}, x_i)$, ed in x_i ha un salto di ampiezza pari a $p(x_i) = F(x_i) - F(x_i^-)$. Quindi $F(x)$ può essere così espressa in termini della densità discreta:

$$F(x) = \sum_{k:\, x_k \leq x} p(x_k) = \begin{cases} 0, & x < x_1 \\ p(x_1), & x_1 \leq x < x_2 \\ p(x_1) + p(x_2), & x_2 \leq x < x_3 \\ p(x_1) + p(x_2) + p(x_3), & x_3 \leq x < x_4 \\ \vdots \end{cases}$$

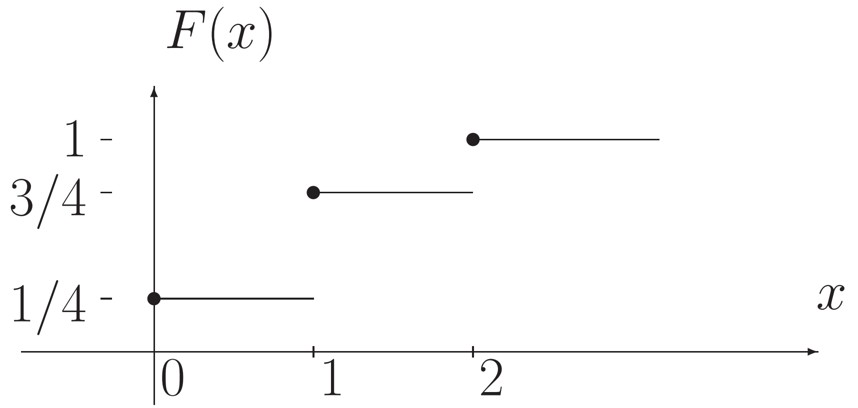


Esempio. Se X assume valori $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, con densità discreta

$$p(0) = \frac{1}{4}, \quad p(1) = \frac{1}{2}, \quad p(2) = \frac{1}{4},$$

la funzione di distribuzione di X è data da

$$F(x) = \sum_{k\colon x_k \le x} p(x_k) = \begin{cases} 0, & x < 0 \\ p(0) = \frac{1}{4}, & 0 \le x < 1 \\ p(0) + p(1) = \frac{3}{4}, & 1 \le x < 2 \\ p(0) + p(1) + p(2) = 1, & x \ge 2. \end{cases}$$



Esempio. In un gioco, in cui si lancia una moneta per 3 volte, si vincono k euro se esce testa per la prima volta al lancio k-esimo ($k = 1, 2, 3$), e si perdono $c > 0$ euro se non esce mai testa. Sia X la variabile aleatoria che descrive la vincita. Ricavarne la funzione di distribuzione $F(x)$, e determinare il valore di c affinché il gioco sia equo.

Soluzione. Indicando con X la vincita al gioco, e posto $p(k) = P(X = k)$, con $k \in \{-c, 1, 2, 3\}$, per l'indipendenza dei lanci risulta

$$p(-c) = P(\overline{T_1} \cap \overline{T_2} \cap \overline{T_3}) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}, \qquad p(1) = P(T_1) = \frac{1}{2},$$

$$p(2) = P(\overline{T_1} \cap T_2) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}, \qquad p(3) = P(\overline{T_1} \cap \overline{T_2} \cap T_3) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}.$$

Pertanto: $F(x) = 0$ per $x < -c$, $F(x) = 1/8$ per $-c \le x < 1$, $F(x) = 5/8$ per $1 \le x < 2$, $F(x) = 7/8$ per $2 \le x < 3$, $F(x) = 1$ per $x \ge 3$. La vincita media è

$$\sum_k k \cdot p(k) = -c \cdot \frac{1}{8} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} = \frac{11 - c}{8}.$$

Ne segue che il gioco è equo se $c = 11$.

Esercizio. Cosa cambia se si vincono k euro se esce testa k volte? ($k = 1, 2, 3$)

4.4 Valore atteso

Introduciamo uno dei più importanti concetti in calcolo delle probabilità.

Definizione. Se X è una variabile aleatoria discreta con densità discreta $p(x)$, il valore atteso (o valore medio, o speranza matematica) di X è definito da

$$E(X) = \sum_{x:\,p(x)>0} x \cdot p(x).$$

Il valore atteso di X è un indice di tendenza centrale dei valori della variabile aleatoria, essendo uguale alla media pesata di tutti i possibili valori che X può assumere, ognuno pesato con la probabilità che X lo assuma. Ad esempio, se $p(0) = p(1) = \frac{1}{2}$:

$$E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}.$$

Se, invece, risulta $p(0) = \frac{1}{3}$, $p(1) = \frac{2}{3}$, allora

$$E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}.$$

Esempio. Calcolare il valore atteso nel lancio di un dado equilibrato.

Soluzione. Essendo

$$p(k) = \frac{1}{6}, \qquad k = 1, 2, \ldots, 6,$$

otteniamo

$$E(X) = \sum_{k=1}^{6} k \cdot p(k) = \frac{1}{6} \sum_{k=1}^{6} k = \frac{1}{6} \cdot \frac{6 \cdot 7}{2} = \frac{7}{2} = 3{,}5$$

avendo usato la relazione

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

Esempio. Calcolare il valore atteso nel lancio di un dado equilibrato avente n facce.

Soluzione. Essendo

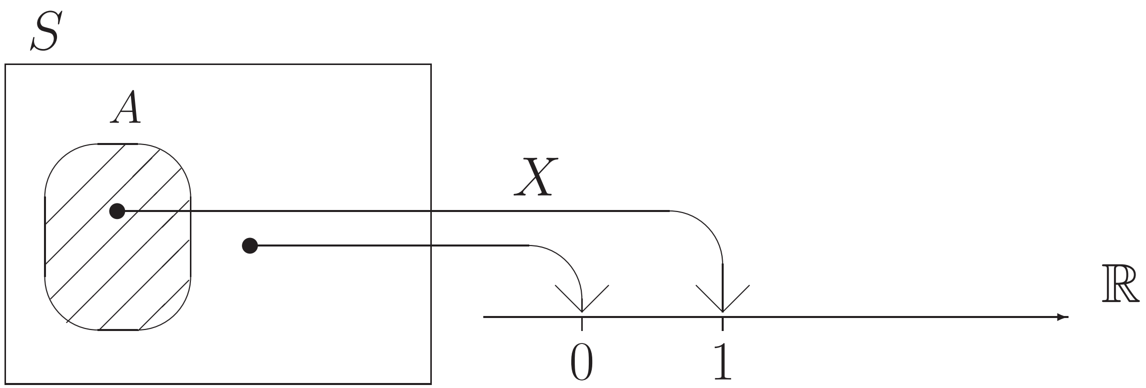
$$p(k) = \frac{1}{n}, \qquad k = 1, 2, \ldots, n,$$

si ha

$$E(X) = \sum_{k=1}^{n} k \cdot p(k) = \frac{1}{n} \sum_{k=1}^{n} k = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Esempio. Si calcoli il valore atteso $E(I_A)$, dove $I_A : S \to \mathbb{R}$ è la *funzione indicatrice* di un evento A, ossia una variabile aleatoria discreta così definita:

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



Soluzione. Poiché $p(1) = P(I_A = 1) = P(A)$ e $p(0) = P(I_A = 0) = P(\overline{A})$, abbiamo

$$E(I_A) = \sum_{k=0}^{1} k \cdot p(k) = 0 \cdot p(0) + 1 \cdot p(1) = p(1) = P(A).$$

Pertanto, il valore atteso della variabile indicatrice di un evento è uguale alla probabilità dell'evento stesso.

Esempio. Il concorrente di un gioco a quiz deve rispondere a due domande, $\mathcal{D}_1$ e $\mathcal{D}_2$, ed è libero di scegliere in che ordine rispondere. Se risponde prima alla domanda $\mathcal{D}_i$ gli sarà consentito di rispondere all'altra domanda solo se avrà risposto correttamente alla prima. Egli riceve v_i euro se risponde correttamente alla domanda $\mathcal{D}_i$, $i = 1, 2$. Se la probabilità di $E_i = \{\text{conosce la risposta di } \mathcal{D}_i\}$ è P_i, con E_1 e E_2 indipendenti per ipotesi, a quale domanda dovrà rispondere prima per massimizzare il guadagno atteso?

Soluzione. Sia X_i la vincita del concorrente se risponde prima a $\mathcal{D}_i$ $(i = 1, 2)$; si ha

$$
\begin{array}{ll}
P(X_1 = 0) = P(\overline{E}_1) = 1 - P_1 & \qquad P(X_2 = 0) = 1 - P_2 \\
P(X_1 = v_1) = P(E_1)P(\overline{E}_2) = P_1\,(1 - P_2) & \qquad P(X_2 = v_2) = P_2\,(1 - P_1) \\
P(X_1 = v_1 + v_2) = P(E_1)P(E_2) = P_1\,P_2 & \qquad P(X_2 = v_1 + v_2) = P_1\,P_2.
\end{array}
$$

Se il concorrente risponde prima a $\mathcal{D}_1$ il guadagno atteso è

$$E[X_1] = \sum_{x \in \{0, v_1, v_1 + v_2\}} x \cdot P(X_1 = x) = v_1\,P_1\,(1 - P_2) + (v_1 + v_2)\,P_1\,P_2$$

mentre se risponde prima a $\mathcal{D}_2$ si ha

$$E[X_2] = \sum_{x \in \{0, v_2, v_1 + v_2\}} x \cdot P(X_2 = x) = v_2\,P_2\,(1 - P_1) + (v_1 + v_2)\,P_1\,P_2$$

Notiamo che è più vantaggioso rispondere prima a $\mathcal{D}_1$ se risulta

$$E[X_1] \geq E[X_2]$$

ossia se

$$v_1\, P_1\, (1 - P_2) + (v_1 + v_2)\, P_1\, P_2 \geq v_2\, P_2\, (1 - P_1) + (v_1 + v_2)\, P_1\, P_2.$$

Semplificando si ottiene

$$v_1\, P_1\, (1 - P_2) \geq v_2\, P_2\, (1 - P_1),$$

pertanto è più vantaggioso rispondere prima a $\mathcal{D}_1$ se

$$\frac{v_1\, P_1}{1 - P_1} \geq \frac{v_2\, P_2}{1 - P_2}.$$

Notiamo che tale condizione si estende facilmente al caso di più domande:

$$\frac{v_1\, P_1}{1 - P_1} \geq \frac{v_2\, P_2}{1 - P_2} \geq \cdots \geq \frac{v_n\, P_n}{1 - P_n},$$

e ciò suggerisce l'applicazione di un metodo *greedy* per massimizzare il guadagno atteso.

Esempio. Una comitiva di 120 studenti viene condotta in gita in 3 autobus. Nel primo ci sono 36 studenti, nel secondo 40, e nel terzo 44. Se si sceglie un autobus a caso, sia Y il numero di studenti che viaggiano sull'autobus scelto a caso. Calcolare $E(Y)$. All'arrivo si sceglie a caso uno studente tra i 120. Se X denota il numero di studenti che hanno viaggiato sull'autobus dello studente scelto a caso, calcolare $E(X)$.

Soluzione. Si ha che Y è una variabile aleatoria discreta, con

$$P(Y = 36) = P(Y = 40) = P(Y = 44) = \frac{1}{3},$$

quindi il numero medio di studenti presenti su un autobus scelto a caso è

$$E(Y) = \sum_y y \cdot P(Y = y) = \frac{1}{3}(36 + 40 + 44) = \frac{1}{3}(120) = 40.$$

Per X si ha $P(X = 36) = \dfrac{36}{120}$, $P(X = 40) = \dfrac{40}{120}$, $P(X = 44) = \dfrac{44}{120}$ e quindi

$$E(X) = 36 \cdot \frac{36}{120} + 40 \cdot \frac{40}{120} + 44 \cdot \frac{44}{120} = \frac{4832}{120} = 40{,}2667 \ > \ E(Y).$$

Questo fenomeno si verifica perché per quanti più studenti sono presenti su un dato autobus, tanto più sarà probabile che lo studente scelto a caso provenga proprio da quell'autobus. Così si assegna peso maggiore agli autobus con più studenti.

Nota: più in generale, se vi sono k autobus con $n_1, n_2, \ldots, n_k$ studenti, si ha

$$E(X) = \sum_{i=1}^{k} n_i \cdot \frac{n_i}{\sum_{j=1}^{k} n_j} = \frac{\sum_{i=1}^{k} n_i^2}{\sum_{j=1}^{k} n_j} \quad \left(\geq \frac{\sum_{i=1}^{k} n_i}{k} \right)$$

Esempio. In seguito a trial clinici risulta che un farmaco sperimentale produce un miglioramento del 50% in pazienti gravi, un miglioramento del 5% in pazienti di media gravità e un peggioramento dell'1% in pazienti lievi. Si supponga che i pazienti affetti dalla patologia specifica siano per il 10% gravi, per il 15% medi e per il 75% lievi. Qual è il miglioramento medio prodotto dal farmaco sperimentale?

Soluzione. Il miglioramento in percentuale prodotto dal farmaco sperimentale è descritto da una variabile aleatoria X discreta che assume valori -1, 5, 50, con

$$P(X=-1)=0{,}75 \qquad P(X=5)=0{,}15 \qquad P(X=50)=0{,}10.$$

Il miglioramento medio prodotto dal farmaco è del 5%, essendo

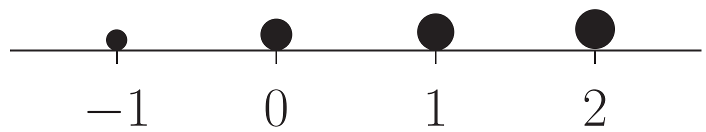
$$E(X)=\sum_{x\in\{-1,5,50\}} x\cdot P(X=x) = -1\cdot 0{,}75+5\cdot 0{,}15+50\cdot 0{,}10=5.$$

Notiamo che se fosse stabilito che un farmaco si può porre in commercio se produce un miglioramento medio non minore del 5% in pazienti affetti da patologia specifica, quel farmaco sarebbe accettato sebbene causi un peggioramento nel 75% dei pazienti.
Notiamo che in questo caso risulta $E(X)>0$ sebbene sia $P(X>0)<1/2$.

Il concetto di valore atteso è analogo al concetto fisico di *baricentro* di una distribuzione di masse. Sia X una variabile aleatoria discreta di densità discreta $p(x_i)$, $i \geq 1$. Immaginiamo una sbarra di peso trascurabile su cui sono poste delle masse

$$p(-1) = 0{,}10 \qquad p(0) = 0{,}25 \qquad p(1) = 0{,}30 \qquad p(2) = 0{,}35$$

nei punti $x_1 = -1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$.



Allora il punto m nel quale la sbarra rimane in equilibrio è il centro di gravità, rispetto al quale è nulla la somma dei momenti dei pesi delle singole porzioni di massa, ossia

$$\sum_i [x_i - m]\, p(x_i) = 0 \qquad \Longrightarrow \qquad m = \frac{\sum_i x_i\, p(x_i)}{\sum_i p(x_i)} = \sum_i x_i\, p(x_i) = E(X).$$

Nell'esempio, $E(X) = \displaystyle\sum_{x=-1}^{2} x \cdot p(x) = -1 \cdot 0{,}10 + 0 \cdot 0{,}25 + 1 \cdot 0{,}30 + 2 \cdot 0{,}35 = 0{,}9.$

Esempio. Da un'urna contenente k biglie nere e $n-k$ biglie bianche si effettuano n estrazioni a caso senza reinserimento. Determinare $E(X)$, dove X denota l'estrazione in cui si estrae una biglia nera per la prima volta.

Soluzione. Poiché le estrazioni sono senza reinserimento l'evento $\{X=j\}$ si verifica se le prime $j-1$ biglie estratte sono bianche, la j-esima è nera, e nelle rimanenti $n-j$ estrazioni vi sono $k-1$ biglie nere. Quindi risulta (schema ipergeometrico)

$$p(j) = P(X=j) = \frac{\binom{j-1}{0}\binom{1}{1}\binom{n-j}{k-1}}{\binom{n}{k}} = \frac{\binom{n-j}{k-1}}{\binom{n}{k}}, \qquad j = 1, 2, \ldots, n-k+1.$$

Notiamo che

$$\sum_j p(j) = \frac{1}{\binom{n}{k}} \sum_{j=1}^{n-k+1} \binom{n-j}{k-1} = \frac{1}{\binom{n}{k}} \sum_{r=k-1}^{n-1} \binom{r}{k-1} = \frac{1}{\binom{n}{k}} \binom{n}{k} = 1,$$

avendo posto $r=n-j$ e fatto uso dell'identità $\sum_{r=k}^{n}\binom{r}{k} = \binom{n+1}{k+1}$.

Si ha dunque

$$E(X) = \sum_j j\,p(j) = \frac{1}{\binom{n}{k}} \sum_{j=1}^{n-k+1} j \binom{n-j}{k-1} = \frac{1}{\binom{n}{k}} \sum_{j=1}^{n-k+1} \sum_{h=1}^{j} \binom{n-j}{k-1}$$

$$= \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \sum_{j=h}^{n-k+1} \binom{n-j}{k-1} = \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \sum_{r=k-1}^{n-h} \binom{r}{k-1},$$

avendo posto $r = n - j$. Facendo uso dell'identità $\sum_{r=k}^{n} \binom{r}{k} = \binom{n+1}{k+1}$ segue

$$E(X) = \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \binom{n-h+1}{k} = \frac{1}{\binom{n}{k}} \sum_{r=k}^{n} \binom{r}{k} = \frac{1}{\binom{n}{k}} \binom{n+1}{k+1},$$

avendo posto $r = n-h+1$ e usato nuovamente l'identità $\sum_{r=k}^{n} \binom{r}{k} = \binom{n+1}{k+1}$. Pertanto,

$$E(X) = \frac{n+1}{k+1}.$$

Nel caso particolare in cui $k = 1$, risulta

$$p(j) = P(X = j) = \frac{\binom{n-j}{0}}{\binom{n}{1}} = \frac{1}{n}, \qquad j = 1, 2, \ldots, n.$$

In tal caso X si dice avere distribuzione uniforme discreta, e risulta

$$E(X) = \frac{n+1}{2}.$$

Notiamo che il problema considerato nell'esempio precedente può essere anche interpretato come il problema della ricerca sequenziale, in cui è data una lista di n elementi, k dei quali sono del tipo da individuare e sono distribuiti a caso. L'algoritmo si basa sulla scansione sequenziale della lista, fermandosi in corrispondenza del primo dei k elementi da individuare. La variabile aleatoria X descrive il numero di confronti necessari per individuare tale elemento, e quindi dà una misura della complessità temporale dell'algoritmo. Pertanto la complessità nel caso medio è data da

$$E(X) = \frac{n+1}{k+1}.$$

Esempio. Da un'urna contenente k biglie nere e $n-k$ biglie bianche si effettuano ripetutamente estrazioni a caso con reinserimento. Determinare $E(Y)$, dove Y denota l'estrazione in cui si estrae una biglia nera per la prima volta.

Soluzione. Poiché le estrazioni sono con reinserimento, per l'indipendenza risulta (schema geometrico)

$$p(j) = P(Y=j) = P(\overline{A}_1 \cap \overline{A}_2 \cap \cdots \cap \overline{A}_{j-1} \cap A_j) = p(1-p)^{j-1}, \qquad j = 1, 2, \ldots,$$

dove $A_k = \{\text{si estrae una biglia nera nel caso } k\text{-esimo}\}$, e $p = P(A_k) = \dfrac{k}{n}$. Quindi

$$E(Y) = p \sum_{j=1}^{\infty} j(1-p)^{j-1} = p \sum_{j=1}^{\infty} \sum_{k=1}^{j} (1-p)^{j-1} = p \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} (1-p)^{j-1}.$$

Ponendo $n = j-k$ si ha

$$E(Y) = p \sum_{k=1}^{\infty} (1-p)^{k-1} \sum_{n=0}^{\infty} (1-p)^n = p\,\frac{1}{p}\,\frac{1}{p} = \frac{1}{p} = \frac{n}{k}.$$

Dal confronto con l'esempio precedente: $E(X) = \dfrac{n+1}{k+1} \leq \dfrac{n}{k} = E(Y)$, per $k \leq n$.

4.5 Valore atteso di una funzione di una variabile aleatoria

Sia X una variabile aleatoria discreta di cui è nota la sua densità discreta $p(x_i)$.
Si desidera calcolare il valore atteso di una qualche funzione di X, diciamo $Y = g(X)$.

Essendo $Y = g(X)$ stessa una variabile aleatoria discreta, avrà una densità discreta che possiamo determinare conoscendo quella di X. Ricavata la densità discreta di Y possiamo calcolare $E(Y) = E[g(X)]$ utilizzando la definizione di valore atteso.
Esempio. Sia X una variabile aleatoria che assuma i valori -1, 0, 1 con probabilità $P(X = -1) = 0{,}2 \quad P(X = 0) = 0{,}5 \quad P(X = 1) = 0{,}3$. Calcolare $E(X^2)$.
Soluzione. Sia $Y = X^2$; è immediato verificare che la densità discreta di Y è

$$P(Y = 1) = P(X^2 = 1) = P(X = \pm 1) = P(X = -1) + P(X = 1) = 0{,}2 + 0{,}3 = 0{,}5$$

$$P(Y = 0) = P(X^2 = 0) = P(X = 0) = 0{,}5$$

Quindi $E(X^2) = E(Y) = \sum_{y=0}^{1} y \cdot P(Y = y) = 0 \cdot 0{,}5 + 1 \cdot 0{,}5 = 0{,}5$.
Si noti che risulta $0{,}5 = E(X^2) \neq [E(X)]^2 = 0{,}01$ essendo

$$[E(X)]^2 = \left[\sum_x x \cdot P(X = x)\right]^2 = [-1 \cdot 0{,}2 + 0 \cdot 0{,}5 + 1 \cdot 0{,}3]^2 = [0{,}1]^2 = 0{,}01.$$

Se si considera che $g(X) = g(x)$ quando $X = x$, appare ragionevole che $E[g(X)]$ sia la media pesata dei valori $g(x)$, assegnando $P(X = x)$ come peso a $g(x)$.

Proposizione. Se X è una variabile aleatoria discreta che assume i valori x_i, $i \geq 1$, con probabilità $p(x_i)$, allora per ogni funzione a valori reali $g(x)$ risulta

$$E[g(X)] = \sum_i g(x_i)\, p(x_i).$$

Dimostrazione. Denotiamo con y_j, $j \geq 1$, i diversi valori di $g(x_i)$, $i \geq 1$. Allora, raggruppando tutti gli x_i che hanno stesso valore, abbiamo

$$\begin{aligned}\sum_i g(x_i)\, p(x_i) &= \sum_j \sum_{i:\, g(x_i)=y_j} g(x_i)\, p(x_i) = \sum_j y_j \sum_{i:\, g(x_i)=y_j} p(x_i) \\ &= \sum_j y_j\, P[g(X) = y_j] = E[g(X)],\end{aligned}$$

avendo fatto uso dell'identità $\boxed{P(X \in B) = \sum_{k:\, x_k \in B} p(x_k)}$ per $B = \{x_i : g(x_i) = y_j\}$.

Esempio. Sia X una variabile aleatoria che assume i valori -1, 0, 1 con probabilità $P(X=-1)=0{,}2 \quad P(X=0)=0{,}5 \quad P(X=1)=0{,}3$. Calcolare $E(X^2)$ facendo uso di $E[g(X)]=\sum_i g(x_i)\,p(x_i)$.

Soluzione. Otteniamo $E(X^2)=\sum_i x_i^2\,p(x_i)=(-1)^2\cdot 0{,}2+0^2\cdot 0{,}5+1^2\cdot 0{,}3=0{,}5$.

Esempio. Sia X una variabile aleatoria che descrive il risultato del lancio di un dado non truccato. Calcolare $E[g(X)]$, con $g(X)=\min\{X,4\}$.

Soluzione. Si ha

$$E[g(X)]=\sum_i g(x_i)\,p(x_i)=\sum_{k=1}^{6}\min\{k,4\}\,p(k)=\frac{1}{6}\,(1+2+3+4+4+4)=3.$$

In generale, è facile ricavare che se $g(x)\le h(x)$ allora $E[g(X)]\le E[h(X)]$.

Infatti, se $g(x)\le h(x)$, si ha

$$E[g(X)]=\sum_x g(x)\,p(x)\le\sum_x h(x)\,p(x)=E[h(X)].$$

Nell'esempio precedente, $g(x)=\min\{x,4\}\le x$, con $E[g(X)]=3<E(X)=3{,}5$.

Proposizione. (Proprietà di linearità) Sia X una variabile aleatoria discreta. Se a e b sono costanti reali, allora

$$E(a\,X + b) = a\,E(X) + b.$$

Dimostrazione. Ricordando che $E[g(X)] = \sum_i g(x_i)\,p(x_i)$, per $g(x) = a\,x + b$ è

$$E[g(X)] = \sum_i (a\,x_i + b)\,p(x_i) = a\sum_i x_i\,p(x_i) + b\sum_i p(x_i) = a\,E(X) + b,$$

essendo $E(X) = \sum_i x_i\,p(x_i)$ e $\sum_i p(x_i) = 1$.

Più in generale si può facilmente ricavare che la proprietà di linearità è esprimibile anche nel seguente modo:

$$E[a\,g(X) + b] = a\,E[g(X)] + b.$$

Definizione. Sia X una variabile aleatoria discreta; la quantità $E(X^n)$, $n \geq 1$, è detta *momento di ordine* n di X. Risulta

$$E(X^n) = \sum_{x:\, p(x)>0} x^n \, p(x).$$

Segue che il momento di ordine 1 di X coincide con il valore atteso di X. Inoltre, nel caso $n = 0$ risulta $E(X^n) = 1$, essendo $E(1) = \sum\limits_{x:\, p(x)>0} p(x) = 1$.

Esempio. Sia X una variabile aleatoria che assuma i valori -1, 0, 1 con probabilità $P(X = -1) = 0{,}2 \quad P(X = 0) = 0{,}5 \quad P(X = 1) = 0{,}3$. Calcolare $E(X^n)$.

Soluzione. Il momento di ordine n di X è dato da

$$E(X^n) = \sum_{x=-1}^{1} x^n \, P(X = x) = (-1)^n \, 0{,}2 + 0^n \, 0{,}5 + 1^n \, 0{,}3 = (-1)^n \, 0{,}2 + 0{,}3$$

quindi $E(X^n) = 0{,}5$ per n pari e $E(X^n) = 0{,}1$ per n dispari.

Esempio. Determinare il momento di ordine n della variabile aleatoria $Y = aX + b$, con a e b costanti reali, esprimendolo in termini dei momenti $E(X^k)$, $k = 0, 1, \ldots, n$.

Soluzione. Facendo uso del teorema del binomio della proprietà di linearità si ha

$$E[Y^n] = E[(aX+b)^n] = E\left[\sum_{k=0}^{n} \binom{n}{k} a^k X^k b^{n-k}\right] = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} E[X^k].$$

Esempio. Calcolare $E(X)$, dove X è la variabile aleatoria discreta che descrive il numero di bit pari a **1** in un vettore booleano, di lunghezza n, scelto a caso.

Soluzione. Ricordiamo che

$$p(k) = P(X = k) = \frac{1}{2^n}\binom{n}{k}, \qquad k = 0, 1, \ldots, n$$

quindi

$$E(X) = \sum_{k=0}^{n} k\, p(k) = \frac{1}{2^n}\sum_{k=0}^{n} k\binom{n}{k} = \frac{n}{2^n}\sum_{k=1}^{n}\binom{n-1}{k-1} = \frac{n}{2^n}\, 2^{n-1} = \frac{n}{2},$$

essendo

$$k\binom{n}{k} = k\,\frac{n(n-1)!}{k(k-1)!(n-k)!} = n\,\frac{(n-1)!}{(k-1)!(n-1-(k-1))!} = n\binom{n-1}{k-1}.$$

Esempio. Sia X la variabile aleatoria, detta di Bernoulli, tale che

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

Calcolare il minimo della funzione così definita:

$$\psi(a) := E[(X - a)^2], \qquad a \in \mathbb{R}.$$

Soluzione. Risulta

$$\begin{aligned} \psi(a) &= E[(X-a)^2] = \sum_{x=0}^{1} (x-a)^2\, p(x) = (-a)^2\,(1-p) + (1-a)^2\, p \\ &= a^2 - a^2 p + p - 2ap + a^2 p = a^2 - 2ap + p \end{aligned}$$

e quindi

$$\frac{d}{da}\psi(a) = 2a - 2p = 0 \qquad \Longleftrightarrow \qquad a = p.$$

Ne segue che il minimo di $\psi(a)$ è p, che coincide con $E(X)$. Infatti, si ha

$$E(X) = \sum_{x=0}^{1} x\, p(x) = 0\, p(0) + 1\, p(1) = p.$$

Esempio. Sia X una variabile aleatoria discreta. Calcolare il minimo di

$$\psi(a) := E[(X-a)^2], \qquad a \in \mathbb{R}.$$

Soluzione. Risulta

$$\psi(a) = E[(X-a)^2] = E[X^2 - 2aX + a^2] = E[X^2] - 2aE[X] + a^2,$$

avendo usato la proprietà di linearità, e quindi

$$\frac{d}{da}\psi(a) = -2E[X] + 2a = 0 \qquad \Longleftrightarrow \qquad a = E[X],$$

da cui si trae che il minimo di $\psi(a)$ è $E[X]$.

Notiamo che se si usa a per prevedere il valore che assume X, allora $X-a$ è lo scarto tra il valore osservato, X, e il valore previsto, a. Pertanto, $\psi(a) = E[(X-a)^2]$ rappresenta il valore atteso dello scarto quadratico, che quindi è minimo per $a = E(X)$.

Concludendo, se si usa il valore atteso $E(X)$ per prevedere il risultato di un esperimento descritto dalla variabile aleatoria X, esso rende minimo il valore atteso dello scarto quadratico di previsione.

4.6 Varianza

Sebbene il valore atteso fornisca una media pesata dei possibili valori di una variabile aleatoria, esso non dà alcuna informazione riguardo alla variabilità, o dispersione, di questi valori. Per esempio, le variabili aleatorie date da

$$W = 0 \ \text{ con prob. } 1, \qquad Y = \begin{cases} -1 \text{ con prob. } 1/2 \\ +1 \text{ con prob. } 1/2, \end{cases} \qquad Z = \begin{cases} -100 \text{ con prob. } 1/2 \\ +100 \text{ con prob. } 1/2 \end{cases}$$

hanno lo stesso valore atteso,

$$E(W) = \sum_w w\, p_W(w) = 0 \cdot 1 = 0, \qquad E(Y) = \sum_y y\, p_Y(y) = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0,$$

$$E(Z) = \sum_z z\, p_Z(z) = -100 \cdot \frac{1}{2} + 100 \cdot \frac{1}{2} = 0,$$

ma è presente maggiore dispersione nei valori di Y piuttosto che per W, e nei valori di Z rispetto a quelli di Y.

Poiché ci si attende che X assuma valori disposti attorno al suo valore atteso $E(X)$, appare ragionevole misurare la variabilità di X mediante la media della distanza dal valor medio che i possibili valori di X assumono, ad esempio considerando la quantità $E(|X - \mu|)$, dove $\mu = E(X)$. Per superare alcune difficoltà di tipo matematico si preferisce invece adoperare la media delle differenze al quadrato tra X ed il suo valore atteso μ, e ciò conduce alla seguente definizione.

Definizione. Sia X una variabile aleatoria discreta di valore atteso $\mu = E(X)$; la varianza di X, denotata con $\mathrm{Var}(X)$, è così definita:

$$\mathrm{Var}(X) = E[(X-\mu)^2] = \sum_{x:\, p(x)>0} (x-\mu)^2\, p(x).$$

Notiamo che $\mathrm{Var}(X) = \psi(\mu)$, dove $\psi(a) = E[(X-a)^2]$.

Formula alternativa per la varianza: $\mathrm{Var}(X) = E(X^2) - \mu^2$.
Questa formula si ottiene usando la proprietà di linearità:

$$\begin{aligned}\mathrm{Var}(X) &= E[(X-\mu)^2] = E(X^2 - 2\mu\, X + \mu^2) = E(X^2) - 2\mu\, E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2.\end{aligned}$$

La maniera più semplice di valutare la varianza di X consiste quindi nel calcolare la differenza tra il momento del secondo ordine di X e il quadrato del suo valore atteso:

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \sum_x x^2\, p(x) - \left(\sum_x x\, p(x)\right)^2.$$

Esempio. Calcolare $\mathrm{Var}(X)$, dove X è l'esito del lancio di un dado, e $\mathrm{Var}(Y)$, dove Y è l'esito del lancio di un dado truccato, con le facce numerate 2, 3, 3, 4, 4, 5.

Soluzione. Essendo $p_X(x) = \frac{1}{6}$ per $x = 1, 2, \ldots, 6$ e $\mu = E(X) = \frac{7}{2}$ otteniamo

$$\begin{aligned}\mathrm{Var}(X) &= E(X^2) - \mu^2 = \sum_{x=1}^{6} x^2 \cdot p_X(x) - \left(\frac{7}{2}\right)^2 \\ &= \frac{1+4+9+16+25+36}{6} - \frac{49}{4} = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12} \approx 2{,}9167.\end{aligned}$$

Per Y si ha $p_Y(y) = \frac{1}{6}$ per $y = 2$, 5 e $p_Y(y) = \frac{1}{3}$ per $y = 3$, 4. Pertanto

$$E(Y) = \frac{7}{2} = E(X), \qquad E(Y^2) = \sum_{y=2}^{5} y^2 \cdot p_Y(y) = \frac{79}{6}, \qquad \mathrm{Var}(Y) = \frac{79}{6} - \frac{49}{4} = \frac{11}{12}.$$

Dalla definizione

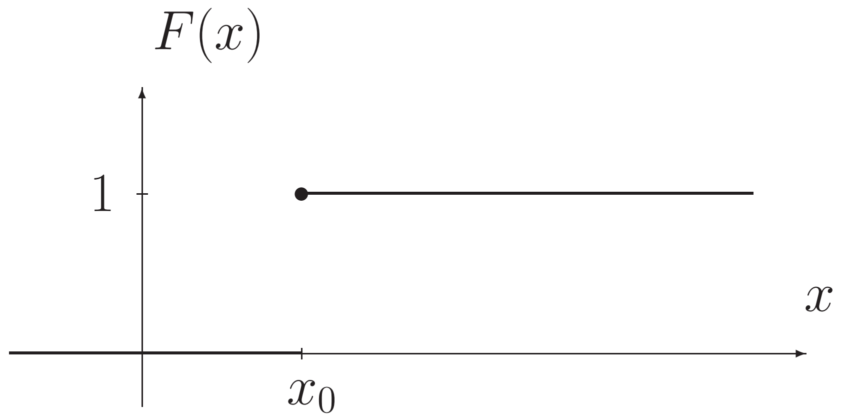
$$\mathrm{Var}(X) = E[(X-\mu)^2] = \sum_{x:\, p(x)>0} (x-\mu)^2\, p(x)$$

discende immediatamente che la varianza di X è non negativa.

Inoltre, la varianza di una variabile aleatoria X è prossima a 0 quando i valori di X sono molto concentrati in prossimità del valore atteso $E[X]$. In tal caso $E[X]$ è molto significativo per la previsione del valore che assume X nella realizzazione di un esperimento. Al contrario, la varianza di X è molto grande quando i suoi valori sono molto distanti dal valore atteso, e quindi $E[X]$ è poco significativo per la previsione del valore che assume X.

Una variabile aleatoria X si dice *degenere* se esiste un reale x_0 tale che $P(X = x_0) = 1$; quindi la funzione di distribuzione di X è

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < x_0 \\ 1, & x \geq x_0. \end{cases}$$



In tal caso si ha

$$E(X) = x_0, \qquad \mathrm{Var}(X) = 0.$$

Infatti, essendo $p(x_0) = 1$, risulta

$$E(X) = \sum_x x\, p(x) = x_0\, p(x_0) = x_0,$$

$$Var(X) = \sum_x (x - \mu)^2\, p(x) = (x_0 - x_0)^2\, p(x_0) = 0 \cdot 1 = 0.$$

Proposizione. (Proprietà della varianza) Sia X una variabile aleatoria discreta. Se a e b sono costanti reali, allora

$$\mathrm{Var}(a\,X+b)=a^2\,\mathrm{Var}(X).$$

Dimostrazione. Ricordando che, per $Y=a\,X+b$ si ha

$$E(Y)=E(a\,X+b)=a\,E(X)+b, \qquad \mathrm{Var}(Y)=E[\{Y-E(Y)\}^2],$$

otteniamo

$$\begin{aligned}\mathrm{Var}(a\,X+b) &= E[\{a\,X+b-E(a\,X+b)\}^2]=E[\{a\,X+b-a\,E(X)-b\}^2]\\ &= E[\{a\,X-a\,E(X)\}^2]=E[a^2\{X-E(X)\}^2]\\ &= a^2\,E[\{X-E(X)\}^2]=a^2\,\mathrm{Var}(X).\end{aligned}$$

Si noti che risulta:
$\mathrm{Var}(a\,X+b)\geq \mathrm{Var}(X)$ se e solo se $a^2\geq 1$, ossia $|a|\geq 1$;
$\mathrm{Var}(a\,X+b)\leq \mathrm{Var}(X)$ se e solo se $a^2\leq 1$, ossia $|a|\leq 1$.

Esempio. Sia X tale che $p_X(1) = 1/4$ e $p_X(3) = 3/4$. Verificare che $E(Y) = 4$ e $Var(Y) = 3$, se $Y = 2X - 1$.

Soluzione. Risulta $E(X) = \sum_{x=1}^{3} x \cdot p_X(x) = 1 \cdot (1/4) + 3 \cdot (3/4) = 5/2$ e inoltre $E(X^2) = \sum_{x=1}^{3} x^2 \cdot p_X(x) = 1 \cdot (1/4) + 9 \cdot (3/4) = 7$. Pertanto:

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = 7 - \left(\frac{5}{2}\right)^2 = \frac{3}{4}.$$

Per $Y = 2X - 1$ si ha $P(Y = 1) = P(X = 1) = \frac{1}{4}$ e $P(Y = 5) = P(X = 3) = \frac{3}{4}$,

$$E(Y) = 1 \cdot \frac{1}{4} + 5 \cdot \frac{3}{4} = 4, \quad E(Y^2) = 1 \cdot \frac{1}{4} + 25 \cdot \frac{3}{4} = 19, \quad \mathrm{Var}(Y) = 19 - 16 = 3.$$

In alternativa: $E(Y) = 2\,E(X) - 1 = 2\,\frac{5}{2} - 1 = 4$ e $Var(Y) = 4\,Var(X) = 4\,\frac{3}{4} = 3$.

Notiamo che $E(X)$ ha stesse unità di misura di X, mentre $\mathrm{Var}(X)$ ha stesse unità di misura di X^2. La seguente misura di variabilità ha invece stesse unità di X.

Definizione. Sia X una variabile aleatoria discreta; la radice quadrata di $\mathrm{Var}(X)$, che denotiamo con σ_X oppure $\sigma(X)$, è detta *deviazione standard* di X. Cioè

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

4.7 Le variabili aleatorie di Bernoulli e binomiali

Supponiamo di eseguire un esperimento i cui possibili esiti appartengono a due categorie, ovvero possono essere classificati come *successo* e *insuccesso*.

Poniamo $X = 1$ quando l'esito è un successo, e $X = 0$ quando l'esito è un insuccesso. La densità discreta della variabile aleatoria X è quindi

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

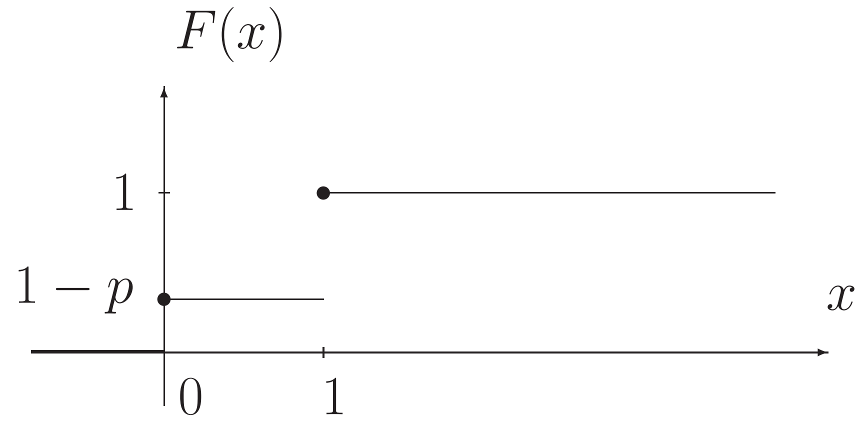
dove p rappresenta la probabilità che la prova abbia avuto successo.

La variabile aleatoria X siffatta è detta variabile aleatoria di Bernoulli; la sua densità discreta può essere così espressa:

$$p(x) = p^x\,(1-p)^{1-x} = \begin{cases} p^0\,(1-p)^{1-0} = 1 - p, & \text{per } \; x = 0, \\ p^1\,(1-p)^{1-1} = p, & \text{per } \; x = 1. \end{cases}$$

Notiamo che per una variabile aleatoria di Bernoulli X la funzione di distribuzione è

$$F(x) = \sum_{k \leq x} p(k) = \sum_{k \leq x} p^k (1-p)^{1-k} = \begin{cases} 0, & x < 0 \\ p(0) = 1 - p, & 0 \leq x < 1 \\ p(0) + p(1) = 1, & x \geq 1 \end{cases}$$



Il momento di ordine n è dato da

$$E[X^n] = \sum_{x=0}^{1} x^n p(x) = 0^n p(0) + 1^n p(1) = p(1) = p.$$

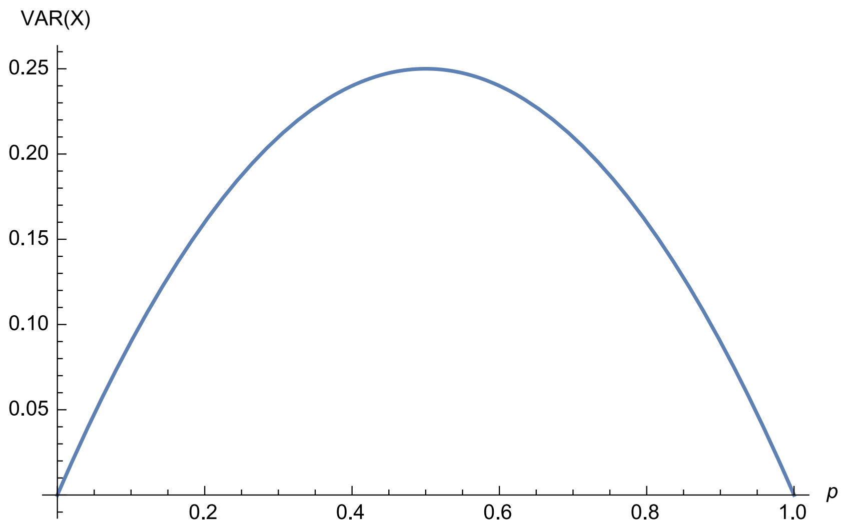
Essendo $E[X^n] = p$, si ha che valore atteso e varianza sono

$$E[X] = p,$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 = p - p^2 = p(1-p).$$

Notiamo che $\text{Var}(X) = 0$ se $p = 0$ e $p = 1$, quindi in questi casi X è una variabile degenere. Inoltre, $\text{Var}(X)$ è massima per $p = 1/2$, essendo

$$\frac{d}{dp}\text{Var}(X) = \frac{d}{dp}p(1-p) = 1 - 2p = 0 \qquad \Longleftrightarrow \qquad p = \frac{1}{2}.$$



Supponiamo ora di eseguire n prove sotto ipotesi d'indipendenza, ognuna avente come possibili risultati successo (con probabilità p) e insuccesso (con probabilità $1-p$).

Sia X il numero di successi che si ottengono nelle n prove. Allora, X è detta variabile aleatoria binomiale di parametri $(n;p)$. Quindi, una variabile aleatoria di Bernoulli è semplicemente una variabile aleatoria binomiale di parametri $(1;p)$.

La densità discreta di una variabile aleatoria binomiale di parametri $(n;p)$ è data da

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \dots, n \qquad (0 \le p \le 1).$$

Ogni sequenza di n esiti contenenti x successi e $n-x$ insuccessi si verifica con probabilità $p^x(1-p)^{n-x}$, grazie all'ipotesi di indipendenza delle prove. La densità discreta $p(x)$ segue allora dal fatto che ci sono $\binom{n}{x}$ differenti sequenze di n esiti contenenti x successi e $n-x$ insuccessi, essendoci esattamente $\binom{n}{x}$ differenti modi di scegliere le x prove in cui si verifichino i successi. Per esempio, ci sono $\binom{4}{2} = 6$ differenti modi di avere $x = 2$ successi in $n = 4$ prove: $(ssii)$, $(sisi)$, $(siis)$, $(issi)$, $(isis)$, $(iiss)$.

Si ricava facilmente che le probabilità date dalla densità discreta

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n$$

danno somma 1 per ogni $p \in [0,1]$; infatti dal teorema del binomio si ha

$$\sum_{x=0}^{n} p(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = [p + (1-p)]^n = 1^n = 1.$$

Esempio. Determinare la densità discreta della variabile aleatoria X che descrive il numero di volte che esce testa in $n = 5$ lanci indipendenti di una moneta non truccata.

Soluzione. La variabile aleatoria in questione è binomiale di parametri $(5; \frac{1}{2})$ quindi

$p(x)$	$p(0)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$
$\binom{5}{x}\left(\frac{1}{2}\right)^x\left(1-\frac{1}{2}\right)^{5-x} = \binom{5}{x}\frac{1}{2^5} = \binom{5}{x}\frac{1}{32}$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Esempio. Durante il "Progetto Manhattan" (per la creazione della bomba nucleare), il celebre fisico Enrico Fermi chiese al Gen. Groves, direttore del progetto, quale fosse la definizione di un "grande" generale. Groves replicò che ogni generale che ha vinto 5 battaglie in sequenza può sicuramente definirsi grande. Fermi allora chiese quanti generali fossero grandi. Groves rispose circa 3 su 100. Fermi congetturò che considerando che schieramenti opposti per molte battaglie sono approssimativamente di uguali forze, la probabilità di vincere una battaglia è $1/2$ e la probabilità di vincere 5 battaglie in sequenza è $(1/2)^5 = 1/32 = 0{,}03125$. Così concluse: "Quindi hai ragione Generale, circa 3 su 100. Probabilità matematica, non genio."

Invero, nell'ipotesi di indipendenza tra battaglie, la soluzione di Fermi è la probabilità che una variabile aleatoria binomiale X di parametri $(5; \frac{1}{2})$ assuma valore 5:

$$P(X = 5) = \binom{5}{5} \left(\frac{1}{2}\right)^5 \left(1 - \frac{1}{2}\right)^0 = \left(\frac{1}{2}\right)^5 = \frac{1}{32} = 0{,}03125.$$

Esempio. Nel gioco della roulette la probabilità che esca un numero rosso è $p = 18/37 = 0{,}4865$. Se un giocatore gioca 5 volte puntando sul rosso qual è la probabilità che non vinca mai? E che vinca almeno 2 volte? Cosa cambia se gioca 10 volte?

Soluzione. Sia X il numero di vincite realizzate in n tentativi (indipendenti); evidentemente ha distribuzione binomiale di parametri $(n; p)$. Quindi, se $n = 5$ si ha

$$P(X = 0) = \binom{5}{0}\left(\frac{18}{37}\right)^0\left(1 - \frac{18}{37}\right)^5 \approx 0{,}0357$$

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{5}{0}\left(\frac{18}{37}\right)^0\left(1 - \frac{18}{37}\right)^5 - \binom{5}{1}\left(\frac{18}{37}\right)^1\left(1 - \frac{18}{37}\right)^4 \approx 0{,}7952. \end{aligned}$$

Se $n = 10$ risulta

$$P(X = 0) = \binom{10}{0}\left(\frac{18}{37}\right)^0\left(1 - \frac{18}{37}\right)^{10} \approx 0{,}0013$$

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{10}{0}\left(\frac{18}{37}\right)^0\left(1 - \frac{18}{37}\right)^{10} - \binom{10}{1}\left(\frac{18}{37}\right)^1\left(1 - \frac{18}{37}\right)^9 \approx 0{,}9866. \end{aligned}$$

A titolo di esempio nella tabella seguente è riportata la densità discreta di una variabile aleatoria binomiale per $n = 5$ e per 2 scelte del parametro p:

$p(x)$	$p = 1/2$	$p = 18/37$
$p(0)$	0,0312	0,0357
$p(1)$	0,1562	0,1691
$p(1)$	0,3125	0,3205
$p(3)$	0,3125	0,3036
$p(4)$	0,1562	0,1438
$p(5)$	0,0312	0,0272

Esempio. In un gioco d'azzardo un giocatore scommette su uno dei numeri compresi tra 1 e 6. Si lanciano 3 dadi. Se il numero su cui ha scommesso appare k volte, con $k = 1, 2, 3$, allora il giocatore vince vince k euro. Se invece il numero non esce, allora il giocatore perde un euro. Il gioco è equo? Ovvero, il capitale finale atteso è zero?

Soluzione. Il numero di dadi che mostra il numero su cui si è puntato è una variabile aleatoria binomiale di parametri $(3; \frac{1}{6})$. Quindi, denotando con X la vincita si ha

$$P(X = -1) = \binom{3}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^3 = \frac{125}{216}, \qquad P(X = 1) = \binom{3}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^2 = \frac{75}{216},$$

$$P(X = 2) = \binom{3}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1 = \frac{15}{216}, \qquad P(X = 3) = \binom{3}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^0 = \frac{1}{216},$$

da cui segue che la vincita attesa è

$$E(X) = \sum_{x\colon p(x)>0} x\,p(x) = -1\cdot\frac{125}{216} + 1\cdot\frac{75}{216} + 2\cdot\frac{15}{216} + 3\cdot\frac{1}{216} = -\frac{17}{216} \approx -0{,}079.$$

Quindi, giocando ripetutamente, ci si attende di perdere 17 euro ogni 216 partite.

Proposizione. Se X è una variabile aleatoria binomiale di parametri $(n; p)$, allora

$$E(X) = n\,p, \qquad \mathrm{Var}(X) = n\,p\,(1-p).$$

Dimostrazione. Determiniamo il momento di X di ordine k:

$$E(X^k) = \sum_{x:\,p(x)>0} x^k \cdot p(x) = \sum_{i=0}^{n} i^k \cdot \binom{n}{i} p^i\,(1-p)^{n-i} = \sum_{i=1}^{n} i^k \cdot \binom{n}{i} p^i\,(1-p)^{n-i}.$$

Utilizzando l'identità

$$i\binom{n}{i} = i\,\frac{n!}{i!\,(n-i)!} = n\,\frac{(n-1)!}{(i-1)!\,(n-i)!} = n\binom{n-1}{i-1}$$

ricaviamo che

$$E(X^k) = np\sum_{i=1}^{n} i^{k-1}\binom{n-1}{i-1} p^{i-1}\,(1-p)^{n-i} = np\sum_{j=0}^{n-1} (j+1)^{k-1}\binom{n-1}{j} p^j\,(1-p)^{n-1-j}$$

avendo posto $j = i-1$. Si ha quindi

$$E(X^k) = n\,p\,E[(Y+1)^{k-1}]$$

dove Y è una variabile aleatoria binomiale di parametri $(n-1; p)$.

Ponendo $k = 1$ nella formula $E(X^k) = n\,p\,E[(Y+1)^{k-1}]$ ricaviamo

$$E(X) = n\,p$$

così il valore atteso di successi che si verificano in n prove indipendenti quando la probabilità di successo vale p, è pari a $n\,p$. Quindi risulta $E(Y+1) = (n-1)\,p + 1$. Pertanto, ponendo $k = 2$ nella formula $E(X^k) = n\,p\,E[(Y+1)^{k-1}]$ ricaviamo

$$E(X^2) = n\,p\,E(Y+1) = n\,p\,[(n-1)\,p + 1].$$

Ricordando che $E(X) = n\,p$ si ottiene infine

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = n\,p\,[(n-1)\,p + 1] - n^2\,p^2 = n\,p\,(1-p),$$

così la varianza del numero di successi che si verificano in n prove indipendenti quando la probabilità di successo vale p, è pari a $n\,p\,(1-p)$.

Per $n = 1$ la variabile aleatoria binomiale coincide con la variabile di Bernoulli, ed infatti in tal caso si ha $E(X) = p$ e $\mathrm{Var}(X) = p\,(1-p)$.

Esempio. Da un'urna contenente n biglie numerate da 1 a n si effettuano n estrazioni con reinserimento. Diciamo che si ha una concordanza all'estrazione k-esima se in tale estrazione si estrae la biglia numero k. Se X è il numero di concordanze che si verificano nelle n estrazioni, determinare la distribuzione di X, il suo valore atteso e la varianza.

Soluzione. Poiché le estrazioni si effettuano con reinserimento, possiamo riguardarle come prove indipendenti aventi probabilità di successo $p = \frac{1}{n}$. Pertanto X è una variabile aleatoria binomiale di parametri n e $p = \frac{1}{n}$, con densità discreta

$$p(k) = P(X = k) = \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k}, \qquad 0 \le k \le n,$$

che rappresenta la probabilità di avere k concordanze in n estrazioni. Quindi, si ha

$$E[X] = n\,p = n\,\frac{1}{n} = 1, \qquad Var[X] = n\,p\,(1-p) = n\,\frac{1}{n}\left(1 - \frac{1}{n}\right) = 1 - \frac{1}{n}.$$

Notiamo che per $n = 1$ si ha $Var[X] = 0$, ossia X è degenere. Inoltre, per $n = 5$:

$p(0)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$
0,3277	0,4096	0,2048	0,0512	0,0064	0,0003

Esempio. Se X è il numero di successi che si verificano in n prove indipendenti quando la probabilità di successo vale p, determinare valore atteso e varianza della frequenza relativa $F_n = X/n$ del numero di successi.

Soluzione. Poiché X è una variabile aleatoria binomiale di parametri $(n; p)$, risulta $E(X) = n\,p$ e $\mathrm{Var}(X) = n\,p\,(1-p)$. Ricordando la proprietà di linearità del valore atteso $E(aX+b) = a\,E(X)+b$, e la proprietà della varianza $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$, si ha pertanto

$$E(F_n) = E\left(\frac{X}{n}\right) = \frac{1}{n}\,E(X) = \frac{1}{n}\,n\,p = p,$$

$$\mathrm{Var}(F_n) = \mathrm{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2}\,\mathrm{Var}(X) = \frac{1}{n^2}\,n\,p\,(1-p) = \frac{p\,(1-p)}{n}.$$

Osserviamo, in particolare, che $E(F_n)$ è costante in n, mentre $\mathrm{Var}(F_n)$ è decrescente in n, e tende a 0 quando n tende a $+\infty$.

Si noti, inoltre, che F_n è una variabile aleatoria discreta che assume valori $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$. Infatti, ax $x = 0, 1, 2, \ldots, n$ corrisponde $\frac{x}{n} = 0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$.

Proposizione. Se X è una variabile aleatoria binomiale di parametri $(n; p)$, con $0 < p < 1$, allora per $k = 0, 1, \ldots, n$ la densità discreta sarà inizialmente strettamente crescente e successivamente strettamente decrescente, con massimo in corrispondenza del più grande intero $k \leq (n+1)p$.

Dimostrazione. Si ha

$$
\begin{aligned}
\frac{P(X=k)}{P(X=k-1)} &= \frac{\dfrac{n!}{k!\,(n-k)!}\, p^k\,(1-p)^{n-k}}{\dfrac{n!}{(k-1)!\,(n-k+1)!}\, p^{k-1}\,(1-p)^{n-k+1}} \\
&= \frac{(n-k+1)\,p}{k\,(1-p)}.
\end{aligned}
$$

Quindi $P(X = k) \geq P(X = k-1)$ se e solo se

$$(n-k+1)\,p \geq k\,(1-p)$$

ossia

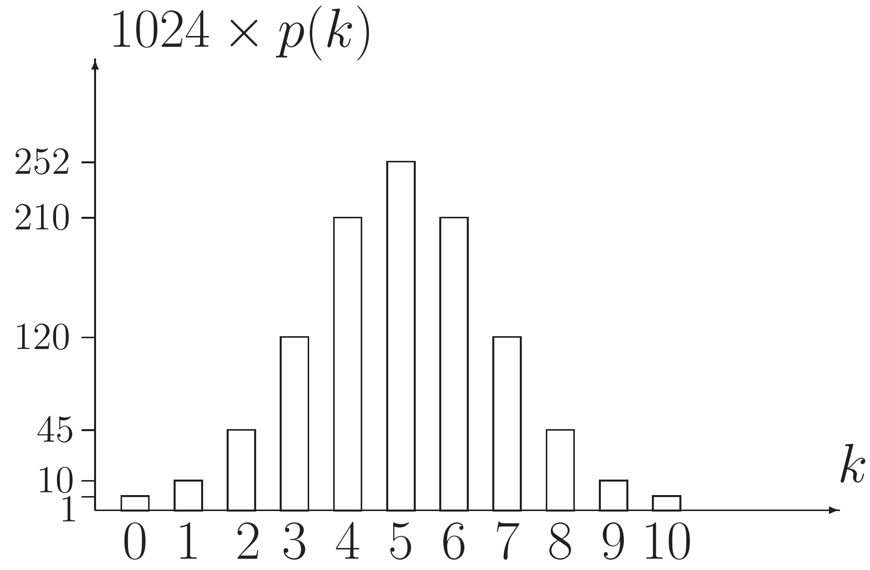
$$k \leq (n+1)p.$$

Nel caso di una variabile binomiale di parametri $(10, \frac{1}{2})$ la densità discreta

$$p(k) = \binom{10}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{10-k}$$

$$= \binom{10}{k} \left(\frac{1}{2}\right)^{10}, \quad 0 \leq k \leq 10$$

è strettamente crescente per $k \leq 5$, ed è simmetrica: $p(k) = p(10 - k) \ \ \forall k$.



Proposizione. Se X e Y sono variabili aleatorie binomiali di parametri $(n; p)$ e $(n; 1 - p)$, rispettivamente, allora

$$P(X = k) = P(Y = n - k), \qquad k = 0, 1, \dots, n.$$

Dimostrazione. Si ricava immediatamente notando che:

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

Esempio. In un sistema elaborativo l'unità centrale prova a connettersi con n unità periferiche, ogni prova avente successo con probabilità p. Determinare media e varianza del numero totale di unità connesse, inclusa la centrale. Se una risorsa viene condivisa tra le unità connesse, determinare la frazione attesa di risorsa per ogni unità.

Soluzione. Nell'ipotesi di indipendenza delle prove, il numero di periferiche connesse è descritto da una variabile aleatoria binomiale X, di parametri n e p. Quindi il numero totale di unità connesse è $X+1$, pertanto $E(X+1) = E(X)+1 = np+1$ e $Var(X+1) = Var(X) = np(1-p)$. La frazione attesa di risorsa per ogni unità è

$$E\left(\frac{1}{X+1}\right) = \sum_{k=0}^{n} \frac{1}{k+1}\binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \frac{1}{k+1}\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \frac{1}{n+1}\sum_{k=0}^{n} \frac{(n+1)!}{(k+1)!(n-k)!} p^k (1-p)^{n-k} = \frac{1}{n+1}\frac{1}{p}\sum_{j=1}^{n+1}\binom{n+1}{j} p^j (1-p)^{n+1-j}$$

$$= \frac{1-(1-p)^{n+1}}{(n+1)p} \neq \frac{1}{E(X+1)} \qquad \text{(per } j = k+1 \text{ e per la formula del binomio).}$$

4.8 La variabile aleatoria di Poisson

Una variabile aleatoria X, che assuma i valori $0, 1, 2, \ldots$, è detta variabile aleatoria di Poisson con parametro $\lambda > 0$ se

$$p(k) = P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

Notiamo che

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!} > 0 \quad \forall k;$$

inoltre si ha

$$\sum_{k=0}^{\infty} p(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = e^0 = 1,$$

essendo

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}.$$

Si può usare la variabile aleatoria di Poisson come approssimazione di una variabile aleatoria binomiale Y di parametri $(n;p)$, quando n è grande e p è piccolo in modo che il prodotto $n\,p$ tenda ad un valore positivo finito. Per $\lambda = n\,p$, ossia $p = \frac{\lambda}{n}$, si ha

$$\begin{aligned} P(Y=k) &= \binom{n}{k} p^k\,(1-p)^{n-k} = \frac{n!}{k!\,(n-k)!}\,p^k\,(1-p)^{n-k} \\ &= \frac{(n)_k}{k!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!}\cdot\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}\cdot\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}. \end{aligned}$$

Per n grande risulta

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \approx 1, \qquad \left(1-\frac{\lambda}{n}\right)^n \approx e^{-\lambda}, \qquad \left(1-\frac{\lambda}{n}\right)^k \approx 1.$$

Pertanto, quando n è grande e p è piccolo in modo che $n\,p = \lambda > 0$ si ha

$$P(Y=k) = \binom{n}{k} p^k\,(1-p)^{n-k} \approx \frac{\lambda^k}{k!}\,e^{-\lambda} = \frac{(np)^k}{k!}\,e^{-np} \qquad (k=0,1,\ldots).$$

Illustriamo il significato dell'approssimazione seguente (valida per n grande)

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, \ldots).$$

Se si eseguono n prove indipendenti, ognuna che dia successo con probabilità p, allora per n grande e p piccolo in modo che $n\,p$ sia un valore positivo finito, il numero totale di successi è ben approssimato da una variabile aleatoria di Poisson di parametro $\lambda = n\,p$.

Ad esempio, se $n = 90$ e $p = 1/18$ si ha $\lambda = n\,p = 90/18 = 5$ e quindi

$$P(Y = k) = \binom{90}{k} \left(\frac{1}{18}\right)^k \left(1 - \frac{1}{18}\right)^{90-k} \qquad \approx \frac{(5)^k}{k!} e^{-5};$$

$$P(Y = 0) = \left(1 - \frac{1}{18}\right)^{90} = 0{,}0058 \qquad \approx \frac{(5)^0}{0!} e^{-5} = e^{-5} = 0{,}0067.$$

Se $n = 900$ e $p = 1/180$ si ha ancora $\lambda = n\,p = 5$ e quindi

$$P(Y = 0) = \left(1 - \frac{1}{180}\right)^{900} = 0{,}0066 \qquad \approx e^{-5} = 0{,}0067.$$

Esempio. Supponiamo che il numero di errori tipografici di una pagina di un libro sia descritto da una variabile aleatoria X di Poisson con parametro $\lambda = \frac{1}{2}$. Calcolare la probabilità che ci sia almeno un errore in una pagina fissata.
Soluzione. Si ha $P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{(1/2)^0}{0!}\, e^{-1/2} \approx 0{,}393$.

Esempio. Supponiamo che un pezzo prodotto da un macchinario sia difettoso con probabilità pari a 0,1. Determinare la probabilità che un lotto di 10 pezzi ne contenga al più uno difettoso, e la sua approssimazione di Poisson.
Soluzione. Sia X la variabile aleatoria che descrive il numero di pezzi difettosi presenti in un lotto. Supponendo che il processo di produzione di ogni pezzo avvenga in condizioni di indipendenza, X ha distribuzione binomiale di parametri $n = 10$ e $p = 0{,}1$. Pertanto, la probabilità desiderata è

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \binom{10}{0}(0{,}1)^0(0{,}9)^{10} + \binom{10}{1}(0{,}1)^1(0{,}9)^9 = 0{,}7361$$

mentre l'approssimazione di Poisson, per $\lambda = np = 10 \cdot 0{,}1 = 1$, fornisce

$$\sum_{k=0}^{1} \frac{\lambda^k}{k!} e^{-\lambda} = \frac{1^0}{0!} e^{-1} + \frac{1^1}{1!} e^{-1} = 2\, e^{-1} = 0{,}7358.$$

Esempio. Il numero di richieste di stampa che giunge ad una stampante aziendale è in media 3,2 al minuto. Approssimare la probabilità che non giungano più di 2 richieste.

Soluzione. Assumiamo che le richieste di stampa giungano da un grande numero n di utenti, ognuno dei quali ha probabilità $p = 3{,}2/n$ di fare una richiesta al minuto. Allora, per l'approssimazione di Poisson, si ha

$$P(X \leq 2) = \sum_{k=0}^{2} P(X = k) = \sum_{k=0}^{2} \frac{\lambda^k}{k!} e^{-\lambda} = \left(1 + 3{,}2 + \frac{(3{,}2)^2}{2}\right) e^{-3{,}2} \approx 0{,}3799.$$

dove X denota una variabile aleatoria di Poisson di parametro $\lambda = 3{,}2$.

Proposizione. Se X è una variabile aleatoria di Poisson di parametro λ, allora

$$E(X) = \lambda, \qquad \mathrm{Var}(X) = \lambda.$$

Dimostrazione. Risulta

$$p(k) = e^{-\lambda}\,\frac{\lambda^k}{k!} = \frac{\lambda}{k}\,e^{-\lambda}\,\frac{\lambda^{k-1}}{(k-1)!} = \frac{\lambda}{k}\,p(k-1), \qquad k \geq 1.$$

Quindi si ha $E(X) = \displaystyle\sum_{k=0}^{\infty} k\,p(k) = \lambda \sum_{k=1}^{\infty} p(k-1)$. Ponendo $j = k-1$ e ricordando che $\displaystyle\sum_{j=0}^{\infty} p(j) = 1$ si ottiene $E(X) = \lambda \displaystyle\sum_{j=0}^{\infty} p(j) = \lambda$.

Analogamente, si ha $E(X^2) = \displaystyle\sum_{k=0}^{\infty} k^2\,p(k) = \lambda \sum_{k=1}^{\infty} k\,p(k-1)$. Per $j = k-1$ si ottiene $E(X^2) = \lambda \displaystyle\sum_{j=0}^{\infty} (j+1)\,p(j) = \lambda\,E(X+1) = \lambda\,[E(X)+1] = \lambda\,(\lambda+1)$, avendo usato la proprietà di linearità. Pertanto, $\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \lambda\,(\lambda+1) - \lambda^2 = \lambda$.

Esempio. Le linee di trasmissione $\mathcal{L}_1$ e $\mathcal{L}_2$ hanno velocità di 8 Mbit/sec e 16 Mbit/sec rispettivamente, e sono soggette ad errori con frequenza X_1 e X_2 al minuto, con X_1 e X_2 variabili di Poisson di parametro $\lambda_1 = 2$ e $\lambda_2 = 4{,}4$ rispettivamente. Sia $U(X_i) = n_i - 200X_i$ il ricavo dall'uso della linea $\mathcal{L}_i$, $i = 1, 2$, dove n_i è il numero di Mbit trasmessi in un minuto usando tale linea. Stabilire quale linea sia più conveniente.

Soluzione. Ricordando la proprietà di linearità del valor medio, e che $E[X_i] = \lambda_i$, si ottiene che il ricavo atteso è identico per le 2 linee:

$$E[U(X_i)] = E[n_i - 200X_i] = n_i - 200E[X_i] = \begin{cases} 8 \cdot 60 - 200 \cdot 2 = 80, & i = 1 \\ 16 \cdot 60 - 200 \cdot 4{,}4 = 80, & i = 2. \end{cases}$$

Analogamente, usando la proprietà della varianza e $Var[X_i] = \lambda_i$, si ha:

$$Var[U(X_i)] = Var[n_i - 200X_i] = (200)^2 Var[X_i] = \begin{cases} 8 \cdot 10^4, & i = 1 \\ 17{,}6 \cdot 10^4, & i = 2. \end{cases}$$

Essendo $Var[U(X_2)] > Var[U(X_1)]$ si trae che il ricavo dall'uso della linea $\mathcal{L}_2$ presenta maggiore variabilità. Infatti, ad esempio:

$$P[U(X_i) < 0] = P[n_i - 200X_i < 0] = P[X_i > n_i/200] = \begin{cases} P[X_1 > 2{,}4], & i = 1 \\ P[X_2 > 4{,}8], & i = 2. \end{cases}$$

Risulta:

$$P[U(X_1) < 0] = P[X_1 > 2{,}4] = P[X_1 \geq 3] = 1 - P[X_1 = 0] - P[X_1 = 1] - P[X_1 = 2]$$

ossia, ricordando che $\lambda_1 = 2$:

$$P[U(X_1) < 0] = 1 - e^{-\lambda_1}\left(1 + \lambda_1 + \frac{\lambda_1^2}{2}\right) = 1 - 5e^{-2} = 0{,}3233.$$

Analogamente, poiché $\lambda_2 = 4{,}4$ si ha

$$P[U(X_2) < 0] = P[X_2 > 4{,}8] = 1 - \sum_{k=0}^{4} P[X_2 = k] = 1 - e^{-4{,}4} \sum_{k=0}^{4} \frac{(4{,}4)^k}{k!} = 0{,}4488.$$

Inoltre, procedendo in modo simile (con $n_1 = 480$ e $n_2 = 960$) si ottiene ad esempio:

$$\begin{aligned} P[U(X_i) > 100] &= P[n_i - 200X_i > 100] = P[X_i < (n_i - 100)/200] \\ &= \begin{cases} P[X_1 < 1{,}9] = P[X_1 \leq 1] = e^{-2} \sum_{k=0}^{1} \frac{(2)^k}{k!} = 0{,}406, & i = 1 \\ P[X_2 < 4{,}3] = P[X_2 \leq 4] = e^{-4{,}4} \sum_{k=0}^{4} \frac{(4{,}4)^k}{k!} = 0{,}5522, & i = 2. \end{cases} \end{aligned}$$

Ciò conferma che il ricavo dall'uso della linea $\mathcal{L}_2$ presenta maggiore variabilità.

4.9 Ulteriori distribuzioni di probabilità discrete

La variabile aleatoria geometrica

Supponiamo di ripetere (in condizioni di indipendenza) una prova avente probabilità di successo p, con $0 < p < 1$, fintanto che non si verifica il primo successo. Denotando con X il numero di prove necessarie per ottenere il primo successo, allora

$$P(X = n) = (1-p)^{n-1}p, \qquad n = 1, 2, \dots$$

Si ha $X = n$ se le prime $n-1$ prove siano state un insuccesso e l'n-esima prova un successo, e quindi per l'indipendenza tra gli eventi $A_k = \{\text{successo alla prova } k\text{-esima}\}$,

$$P(X = n) = P(\overline{A_1} \cap \cdots \cap \overline{A_{n-1}} \cap A_n) = P(\overline{A_1}) \cdots P(\overline{A_{n-1}})\, P(A_n) = (1-p)^{n-1}p,$$

essendo $P(\overline{A_i}) = 1 - p$ e $P(A_i) = p$. Inoltre, risulta:

$$P(X = n) = (1-p)^{n-1}p > 0, \qquad \sum_{n=1}^{\infty} P(X = n) = p \sum_{n=1}^{\infty} (1-p)^{n-1} = \frac{p}{1-(1-p)} = 1,$$

avendo ricordato che $1 + c + c^2 + c^3 + \ldots = 1/(1-c)$ se $-1 < c < 1$.
La variabile aleatoria X è detta variabile aleatoria geometrica di parametro p.

Esempio. Un'urna contiene N biglie bianche e M biglie nere. Si estrae a caso una biglia alla volta, con reinserimento, fino a che non esce la prima biglia nera. Calcolare la probabilità che si debbano estrarre (a) esattamente n biglie, (b) più di k biglie.

Soluzione. Sia X il numero di biglie che si estraggono per ottenere la prima biglia nera. Allora X ha distribuzione geometrica di parametro $p = M/(M+N)$. Quindi

$$
\begin{aligned}
\text{(a)} \qquad P(X=n) &= (1-p)^{n-1}\,p; \\
\text{(b)} \qquad P(X>k) &= \sum_{n=k+1}^{\infty} P(X=n) = p \sum_{n=k+1}^{\infty} (1-p)^{n-1} \\
&= p \sum_{j=0}^{\infty} (1-p)^{k+j} = p\,(1-p)^k \sum_{j=0}^{\infty} (1-p)^j \\
&= p\,(1-p)^k \, \frac{1}{1-(1-p)} = (1-p)^k,
\end{aligned}
$$

avendo posto $j = n-k-1$, ossia $n-1 = k+j$, e avendo ricordato che

$$\sum_{j=0}^{\infty} c^j = \frac{1}{1-c} \qquad \text{per } -1 < c < 1.$$

Notiamo che la probabilità $P(X > k) = (1-p)^k$ si può calcolare anche direttamente, poiché la probabilità che siano necessarie più di k prove per ottenere il primo successo è uguale alla probabilità che le prime k prove diano come esito k insuccessi.

Quindi, per una variabile aleatoria geometrica vale $P(X > k) = (1-p)^k$, $k \geq 0$.

Proposizione. (Proprietà di assenza di memoria) Se X è una variabile aleatoria geometrica di parametro p, allora per ogni $k, n \geq 0$ risulta

$$P(X > n + k | X > k) = P(X > n).$$

Dimostrazione. Per ogni $k, n \geq 0$ si ha

$$P(X > n + k | X > k) = \frac{P(\{X > n + k\} \cap \{X > k\})}{P(X > k)} = \frac{P(X > n + k)}{P(X > k)}$$

$$= \frac{(1-p)^{n+k}}{(1-p)^k} = (1-p)^n = P(X > n).$$

Esaminiamo un'interpretazione della proprietà di assenza di memoria: in una sequenza di partite ripetute in condizioni d'indipendenza un giocatore d'azzardo punta ripetutamente su un esito avente probabilità pari a p. Se X denota il tentativo in cui si realizza la vincita per la prima volta, tale variabile aleatoria è geometrica di parametro p.

Pertanto $P(X > n)$ è la probabilità di non realizzare la vincita nei primi n tentativi. Invece, $P(X > n+k|X > k)$ è la probabilità condizionata di non realizzare la vincita nei prossimi n tentativi sapendo che nei primi k tentativi non c'è stata vincita.

La proprietà di assenza di memoria mostra quindi che tali probabilità sono identiche, e pertanto avere informazioni sulle mancate vincite non altera la probabilità di vincita in tentativi successivi. Ne segue, ad esempio, che puntare sui numeri ritardatari nel gioco nel lotto non incrementa la probabilità di vincita.

La proprietà di assenza di memoria può esprimersi anche così: per ogni $k, n \geq 0$ risulta

$$P(X = n+k|X > k) = \frac{P(X = n+k)}{P(X > k)} = \frac{(1-p)^{n+k-1}p}{(1-p)^k} = (1-p)^{n-1}p = P(X = n).$$

Esempio. Si lancia a caso una moneta. Calcolare la probabilità che esca testa per la prima volta dopo il 5° lancio sapendo che nei primi 3 lanci non esce testa.

Soluzione. Essendo $p = 1/2$, per la proprietà di assenza di memoria si ha:

$$P(X > 5|X > 3) = P(X > 2) = (1-p)^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

Proposizione. Se X è una variabile aleatoria geometrica di parametro p, allora

$$E(X) = \frac{1}{p}, \qquad \mathrm{Var}(X) = \frac{1-p}{p^2}.$$

Dimostrazione. Posto $q = 1 - p$ abbiamo che:

$$E(X) = \sum_{n=0}^{\infty} n\, q^{n-1} p = p \sum_{n=1}^{\infty} \frac{d}{dq}(q^n) = p \frac{d}{dq}\left(\sum_{n=0}^{\infty} q^n\right).$$

Utilizzando l'identità $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$ ricaviamo il valore atteso

$$E(X) = p \frac{d}{dq}\left(\frac{1}{1-q}\right) = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

Calcoliamo ora il momento del secondo ordine di X:

$$E(X^2) = \sum_{n=0}^{\infty} n^2\, q^{n-1} p = p \sum_{n=1}^{\infty} \frac{d}{dq}(nq^n) = p\frac{d}{dq}\left(\sum_{n=1}^{\infty} nq^n\right) = p\frac{d}{dq}\left(\frac{q}{p}\sum_{n=1}^{\infty} nq^{n-1}p\right).$$

Ricordando che $E(X) = \sum_{n=1}^{\infty} n\, q^{n-1} p = \frac{1}{p}$, si ha

$$E(X^2) = p\frac{d}{dq}\left[\frac{q}{(1-q)^2}\right] = p\left[\frac{(1-q)^2 + 2q(1-q)}{(1-q)^4}\right] = p\left[\frac{1}{p^2} + \frac{2(1-p)}{p^3}\right] = \frac{2}{p^2} - \frac{1}{p}.$$

Infine si ottiene la varianza:

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Se si ripete una serie di prove indipendenti, ognuna con probabilità di successo p, allora il numero atteso di prove affiché si verifichi il primo successo è

$$E(X) = \frac{1}{p} \equiv k, \qquad \text{con} \;\; \mathrm{Var}(X) = \frac{1-p}{p^2} \equiv k(k-1).$$

Ad esempio, per la prima uscita di testa nel lancio ripetuto di una moneta si ha

$$p = \frac{1}{2}, \qquad E(X) = 2, \qquad \mathrm{Var}(X) = 2.$$

Per la prima uscita di un fissato valore nel lancio ripetuto di un dado si ha

$$p = \frac{1}{6}, \qquad E(X) = 6, \qquad \mathrm{Var}(X) = 30.$$

Per la prima uscita di un singolo estratto in tentativi ripetuti nel gioco del lotto si ha

$$p = \frac{1}{18}, \qquad E(X) = 18, \qquad \mathrm{Var}(X) = 306.$$

Per la prima uscita di un singolo numero tra 90 in estrazioni indipendenti si ha

$$p = \frac{1}{90}, \qquad E(X) = 90, \qquad \mathrm{Var}(X) = 8010.$$

Esempio. Sia X una variabile aleatoria geometrica di parametro p, con $E(X) = 1/p$ intero. Calcolare $P\{X \leq E(X)\}$ e valutare tale probabilità quando $p \to 1^-$ e quando $p \to 0^+$.

Soluzione. Ponendo $E(X) = 1/p = k$ e ricordando che $P(X \leq k) = 1 - (1-p)^k$ si ha

$$P\{X \leq E(X)\} = 1 - (1-p)^{1/p},$$

con

$$\lim_{p \to 1^-} P\{X \leq E(X)\} = 1 - \lim_{p \to 1^-} (1-p)^{1/p} = 1,$$

$$\lim_{p \to 0^+} P\{X \leq E(X)\} = 1 - \lim_{p \to 0^+} (1-p)^{1/p} = 1 - \lim_{k \to \infty} \left(1 - \frac{1}{k}\right)^k = 1 - e^{-1} \approx 0{,}632.$$

La variabile aleatoria ipergeometrica

Nell'estrarre n biglie senza reinserimento da un'urna che contiene N biglie, di cui m sono bianche e $N-m$ nere, sia X il numero di biglie bianche presenti tra le n estratte. Allora X è detta variabile aleatoria *ipergeometrica*, e per opportuni valori n, m, N ha densità discreta

$$P(X=k)=\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} \qquad k=0,1,\ldots,n.$$

Ricordando che $\binom{r}{k}>0$ quando $0 \le k \le r$, risulta $\binom{m}{k}\binom{N-m}{n-k}>0$ per $\begin{cases} 0\le k\le m \\ 0\le n-k\le N-m \end{cases}$ ossia $\begin{cases} 0\le k\le m \\ n+m-N\le k\le n \end{cases}$ da cui segue che

$$\begin{aligned} &P(X=k)>0 \quad \text{se } \max(0,n+m-N)\le k\le \min(n,m), \\ &P(X=k)=0 \quad \text{altrimenti.} \end{aligned}$$

Notiamo che, in virtù della formula di Vandermonde, risulta:

$$\sum_{k=0}^{n} P(X=k) = \sum_{k=0}^{n} \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{k=0}^{n} \binom{m}{k}\binom{N-m}{n-k} = \frac{\binom{N}{n}}{\binom{N}{n}} = 1.$$

Esempio. In un sistema multiutente vi sono 8 utenti collegati, di cui 5 richiedono l'accesso ad Internet ed i rimanenti non lo richiedono. Se si scelgono a caso 4 utenti collegati, qual è la probabilità che al più 3 di essi richiedano l'accesso ad Internet?

Soluzione. Sia X il numero di utenti che richiede l'accesso ad Internet tra i 4 utenti scelti. Poiché X ha distribuzione ipergeometrica di parametri $n=4$, $m=5$, $N=8$ è:

$$p(k) = P(X=k) = \frac{\binom{5}{k}\binom{3}{4-k}}{\binom{8}{4}}, \qquad \text{con} \quad \binom{8}{4} = \frac{(8)_4}{4!} = \frac{8\cdot 7\cdot 6\cdot 5}{4\cdot 3\cdot 2\cdot 1} = 70.$$

Notiamo che: $p(0)=0$, $p(1)=\frac{1}{14}$, $p(2)=\frac{6}{14}$, $p(3)=\frac{6}{14}$, $p(4)=\frac{1}{14}$. Pertanto, la probabilità richiesta è $P(X\le 3) = 1 - P(X=4) = 1 - \frac{1}{14} = \frac{13}{14}$.

Esempio. Una lista contiene 15 nominativi di studenti, di cui 8 sono del I anno, 4 del II anno, e 3 del III anno. Se si scelgono a caso 3 studenti, quanto vale la probabilità
(a) che siano tutti del I anno?
(b) che ci sia almeno uno del III anno?

Soluzione. (a) Sia X il numero studenti del I anno estratti; allora X ha distribuzione ipergeometrica di parametri $n = 3$, $m = 8$, $N = 15$, quindi

$$P(X = 3) = \frac{\binom{8}{3}\binom{7}{0}}{\binom{15}{3}} = \frac{\frac{(8)_3}{3!} \cdot 1}{\frac{(15)_3}{3!}} = \frac{8 \cdot 7 \cdot 6}{15 \cdot 14 \cdot 13} = \frac{8}{65} = 0{,}1231.$$

(b) Sia Y il numero studenti del III anno estratti; allora Y ha distribuzione ipergeometrica di parametri $n = 3$, $m = 3$, $N = 15$, quindi

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \frac{\binom{3}{0}\binom{12}{3}}{\binom{15}{3}} = 1 - \frac{1 \cdot \frac{(12)_3}{3!}}{\frac{(15)_3}{3!}} = 1 - \frac{12 \cdot 11 \cdot 10}{15 \cdot 14 \cdot 13} = 0{,}5165.$$

Esempio. Un rivenditore acquista componenti elettriche a lotti di 10. Controlla a caso 3 componenti in ogni lotto e lo accetta solo se nessuno dei 3 pezzi controllati è difettoso. Se il 30% dei lotti ha 4 pezzi difettosi e il 70% ha 1 pezzo difettoso, qual è la percentuale dei lotti che il rivenditore rifiuterà?

Soluzione. Sia X il numero di pezzi difettosi tra i 3 controllati, e sia $B = \{$il lotto ha 4 pezzi difettosi$\}$, cosicché $\overline{B} = \{$il lotto ha 1 pezzo difettoso$\}$; si ha

$$\begin{aligned} P(X=0) &= P(X=0|B)\,P(B) + P(X=0|\overline{B})\,P(\overline{B}) \\ &= \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} \cdot \frac{3}{10} + \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} \cdot \frac{7}{10} = \frac{6\cdot5\cdot4}{10\cdot9\cdot8}\cdot\frac{3}{10} + \frac{9\cdot8\cdot7}{10\cdot9\cdot8}\cdot\frac{7}{10} \\ &= 0{,}05 + 0{,}49 = 0{,}54 \end{aligned}$$

pertanto il rivenditore rifiuterà il 46% dei lotti. Notiamo inoltre che

$$P(B|X=0) = \frac{P(X=0|B)\,P(B)}{P(X=0)} = \frac{0{,}05}{0{,}54} \approx 0{,}0926.$$

Se si scelgono a caso n biglie senza reinserimento da un insieme di N biglie, delle quali la frazione $p = m/N$ è bianca, allora il numero di biglie bianche selezionate X ha distribuzione ipergeometrica. È ragionevole supporre che se m e N sono grandi rispetto a n, allora il fatto che si effettui o meno reinserimento ad ogni estrazione possa essere trascurabile. Non tenendo conto delle biglie già estratte, ogni altra estrazione darà una biglia bianca con probabilità approssimativamente pari a p, se m e N sono grandi rispetto a n. In tal caso si può approssimare la distribuzione di X, notando che

$$P(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} = \frac{\dfrac{(m)_k}{k!} \cdot \dfrac{(N-m)_{n-k}}{(n-k)!}}{\dfrac{(N)_n}{n!}}.$$

Quindi si ha

$$P(X = k) = \frac{n!}{k!\,(n-k)!} \frac{(m)_k \cdot (N-m)_{n-k}}{(N)_n} = \binom{n}{k} \frac{(m)_k \cdot (N-m)_{n-k}}{(N)_n}.$$

Per m grande risulta $(m)_k = m(m-1)\cdots(m-k-1) \approx m^k$.

Pertanto, per m e N grandi rispetto a n e k risulta

$$P(X=k) \approx \binom{n}{k} \frac{m^k\,(N-m)^{n-k}}{N^n} = \binom{n}{k} \frac{m^k\,(N-m)^{n-k}}{N^{k+n-k}}$$

$$= \binom{n}{k} \left(\frac{m}{N}\right)^k \left(\frac{N-m}{N}\right)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

dove $p = \dfrac{m}{N}$ è la probabilità di estrarre una biglia bianca nella prima estrazione.

Concludendo, possiamo affermare che la distribuzione ipergeometrica può essere approssimata da una distribuzione binomiale:

$$P(X=k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} \approx \binom{n}{k} \left(\frac{m}{N}\right)^k \left(1-\frac{m}{N}\right)^{n-k}$$

quando m e N sono grandi rispetto a n e k.

Proposizione. Se X è una variabile aleatoria ipergeometrica di parametri n, N e m, allora per $p = \dfrac{m}{N}$ si ha

$$E(X) = n\,p, \qquad \mathrm{Var}(X) = n\,p\,(1-p)\left(1 - \frac{n-1}{N-1}\right).$$

Dimostrazione. Il momento di ordine k di X è dato da:

$$E(X^k) = \sum_{i=0}^{n} i^k\, P(X=i) = \sum_{i=1}^{n} i^k \binom{m}{i}\binom{N-m}{n-i}\Big/\binom{N}{n}.$$

Utilizzando le identità

$$i\binom{m}{i} = \frac{i\,m!}{i!\,(m-i)!} = \frac{m\,(m-1)!}{(i-1)!\,(m-i)!} = m\binom{m-1}{i-1}, \qquad n\binom{N}{n} = N\binom{N-1}{n-1}$$

si ha

$$E(X^k) = \frac{n\,m}{N}\sum_{i=1}^{n} i^{k-1}\binom{m-1}{i-1}\binom{N-m}{n-i}\Big/\binom{N-1}{n-1}.$$

Ponendo $j = i - 1$ in $E(X^k) = \dfrac{n\,m}{N} \displaystyle\sum_{i=1}^{n} i^{k-1} \binom{m-1}{i-1}\binom{N-m}{n-i}\Big/\binom{N-1}{n-1}$ si ha

$$E(X^k) = \frac{n\,m}{N} \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{m-1}{j}\binom{N-m}{n-1-j}\Big/\binom{N-1}{n-1} = \frac{n\,m}{N}\, E[(Y+1)^{k-1}]$$

con Y variabile aleatoria ipergeometrica di parametri $n-1$, $N-1$ e $m-1$. Ponendo $k = 1$ e $k = 2$ si ha rispettivamente:

$$E(X) = n\,\frac{m}{N} = n\,p, \qquad E(X^2) = \frac{n\,m}{N}\, E(Y+1) = \frac{n\,m}{N}\left[\frac{(n-1)\,(m-1)}{N-1} + 1\right].$$

Da ciò, ricordando che $p = m/N$, segue

$$\begin{aligned}
\mathrm{Var}(X) &= \frac{n\,m}{N}\left[\frac{(n-1)\,(m-1)}{N-1} + 1 - \frac{n\,m}{N}\right] = n\,p\left[\frac{(n-1)\,(Np-1)}{N-1} + 1 - n\,p\right] \\
&= n\,p\left[(n-1)\left(p - \frac{1-p}{N-1}\right) + 1 - n\,p\right] \\
&= n\,p\left[(n-1)\,p - (n-1)\,\frac{1-p}{N-1} + 1 - n\,p\right] = n\,p\,(1-p)\left(1 - \frac{n-1}{N-1}\right).
\end{aligned}$$

Poiché risulta

$$\mathrm{Var}(X) = n\,p\,(1-p)\left(1 - \frac{n-1}{N-1}\right),$$

è facile vedere che quando $N \to \infty$ la varianza di una variabile aleatoria ipergeometrica tende alla varianza di una variabile aleatoria binomiale, data da $n\,p\,(1-p)$.

Inoltre, per $n = 1$ la varianza di una variabile aleatoria ipergeometrica coincide con la varianza di una variabile aleatoria binomiale, in quanto entrambe le variabili aleatorie coincidono con una variabile aleatoria di Bernoulli.

La variabile aleatoria uniforme discreta

Nell'estrarre una biglia da un'urna contenente N biglie numerate da 1 a N, denotiamo con X il numero estratto. Allora

$$P(X = k) = \frac{1}{N}, \qquad k = 1, 2, \ldots, N.$$

La variabile aleatoria avente tale densità discreta è detta *uniforme discreta.* Notiamo che la funzione di distribuzione $F(x) = P(X \leq x)$ è data da:

$$F(x) = 0, \ \text{ per } x < 1,$$

$$F(x) = \sum_{r \leq x} P(X = r) = \sum_{r=1}^{k} \frac{1}{N} = \frac{k}{N}, \ \text{ per } k \leq x < k+1 \ \ (k = 1, 2, \ldots, N-1),$$

$$F(x) = 1, \ \text{ per } x \geq N.$$

Proposizione. Se X è una variabile aleatoria uniforme discreta di parametro N, allora

$$E(X) = \frac{N+1}{2}, \qquad \mathrm{Var}(X) = \frac{N^2 - 1}{12}.$$

Dimostrazione. Ricordando che $\displaystyle\sum_{k=1}^{N} k = \frac{N(N+1)}{2}$, si ha

$$E(X) = \sum_{k=1}^{N} k\, P(X=k) = \frac{1}{N}\sum_{k=1}^{N} k = \frac{1}{N}\cdot\frac{N(N+1)}{2} = \frac{N+1}{2}.$$

Analogamente, poiché $\displaystyle\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}$, si ha

$$E(X^2) = \sum_{k=1}^{N} k^2\, P(X=k) = \frac{1}{N}\sum_{k=1}^{N} k^2 = \frac{(N+1)(2N+1)}{6}.$$

Segue pertanto

$$\begin{aligned}\mathrm{Var}(X) &= E(X^2) - [E(X)]^2 = \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4}\\ &= \frac{N+1}{12}\,[2(2N+1) - 3(N+1)] = \frac{N+1}{12}\,(N-1)\\ &= \frac{N^2-1}{12}.\end{aligned}$$

($\star\star$) Appendice. Processo di Poisson

Nello studio di eventi che si verificano ad istanti (aleatori) di tempo supponiamo che, data una costante positiva λ, siano verificate le seguenti ipotesi:

1. La probabilità che un evento si verifichi in un dato intervallo di tempo di lunghezza h è uguale a $\lambda\, h + o(h)$, dove $o(h)$ indica una qualsiasi funzione $f(h)$ per la quale risulta $\lim_{h\to 0} f(h)/h = 0$. [Ad esempio, $f(h) = h^2$ è $o(h)$, mentre $f(h) = h$ non è $o(h)$.]

2. La probabilità che due o più eventi si verifichino in un intervallo di tempo di lunghezza h è uguale a $o(h)$.

3. Sia $E_i = \{$nell'intervallo $(t_{i-1}, t_i]$ si verificano k_i eventi$\}$, $i = 1, 2, \ldots, n$. Per ogni scelta degli istanti di tempo $0 \leq t_0 < t_1 < t_2 < \ldots < t_n$ e degli interi non negativi $k_1, k_2, \ldots, k_n$ gli eventi $E_1, E_2, \ldots, E_n$ sono indipendenti.

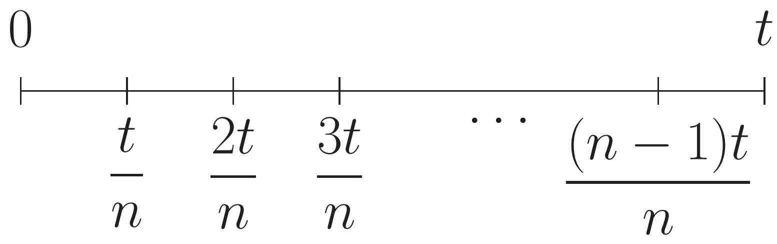
Dalle ipotesi 1 e 2 segue la seguente condizione:

4. La probabilità che non si verifichino eventi in un intervallo di tempo di lunghezza h è uguale a $1 - \lambda\, h + o(h)$.

Mostreremo che se valgono le ipotesi 1, 2 e 3, allora il numero di eventi che si verificano in ogni intervallo di lunghezza t ha distribuzione di Poisson di parametro $\lambda\, t$.

Denotiamo con $N(t)$ il numero di eventi che si verificano nell'intervallo $[0, t]$.

Suddividiamo l'intervallo $[0, t]$ in n sottointervalli disgiunti di lunghezza t/n.



Risulta

$$P\{N(t) = k\} = P(A) + P(B),$$

dove

$A = \{k$ degli n sottointervalli contengono 1 evento e gli altri $n - k$ ne contengono $0\}$

$$B = \{N(t) = k \text{ e almeno 1 sottointervallo contiene 2 o più eventi}\}$$

Mostreremo ora che $P(B) = 0$ e, al limite per $n \to \infty$, $P(A) = e^{-\lambda t}\dfrac{(\lambda t)^k}{k!}$, e così $P\{N(t) = k\} = e^{-\lambda t}\dfrac{(\lambda t)^k}{k!}$. Si ha

$$\begin{aligned}
P(B) &= P\{N(t) = k \text{ e almeno 1 sottointervallo contiene 2 o più eventi}\} \\
&\le P\{\text{almeno 1 sottointervallo contiene 2 o più eventi}\} \\
&= P\left(\bigcup_{i=1}^{n}\{\text{l'}i\text{-esimo sottointervallo contiene 2 o più eventi}\}\right) \\
&\le \sum_{i=1}^{n} P\{\text{l'}i\text{-esimo sottointervallo contiene 2 o più eventi}\} \\
&= \sum_{i=1}^{n} o\left(\frac{t}{n}\right) \\
&= n\, o\left(\frac{t}{n}\right) \\
&= t\left[\frac{o(t/n)}{t/n}\right]
\end{aligned}$$

Pertanto,
$$P(B) \leq t\left[\frac{o(t/n)}{t/n}\right].$$

$\forall t > 0$ si ha $\dfrac{t}{n} \to 0$ quando $n \to \infty$. Quindi, $\dfrac{o(t/n)}{t/n} \to 0$ quando $n \to \infty$ e perciò
$$\lim_{n\to\infty} P(B) = 0.$$

Si ha inoltre

$$\begin{aligned} P(A) &= P\{k \text{ degli } n \text{ sottointervalli contengono 1 evento} \\ &\quad \text{e gli altri } n-k \text{ ne contengono } 0\} \\ &= \binom{n}{k} p^k (1-p)^{n-k}, \end{aligned}$$

dove $p = P\{\text{in un intervallo di ampiezza } \dfrac{t}{n} \text{ si ha 1 evento}\} = \dfrac{\lambda t}{n} + o\left(\dfrac{\lambda t}{n}\right)$. Quindi,

$$P(A) = \binom{n}{k}\left[\frac{\lambda t}{n} + o\left(\frac{\lambda t}{n}\right)\right]^k \left[1 - \frac{\lambda t}{n} - o\left(\frac{\lambda t}{n}\right)\right]^{n-k}$$

Abbiamo ricavato che $P(A)$ è la probabilità di una variabile aleatoria binomiale di parametri $(n;p)$, con $p = \dfrac{\lambda t}{n} + o\left(\dfrac{\lambda t}{n}\right)$. Infatti, per $k = 0, 1, \ldots, n$:

$$P(A) = \binom{n}{k} \left[\frac{\lambda t}{n} + o\left(\frac{\lambda t}{n}\right)\right]^k \left[1 - \frac{\lambda t}{n} - o\left(\frac{\lambda t}{n}\right)\right]^{n-k}.$$

Procedendo al limite per $n \to \infty$ notiamo che $\lim\limits_{n\to\infty} p = 0$ e che

$$\lim_{n\to\infty} n\,p = \lim_{n\to\infty} n \left[\frac{\lambda t}{n} + o\left(\frac{\lambda t}{n}\right)\right] = \lambda\, t + t \lim_{n\to\infty} \frac{o(t/n)}{t/n} = \lambda\, t.$$

Quindi, un ragionamento analogo a quello fatto in precedenza per mostrare che la distribuzione di Poisson approssima quella binomiale, ci dà

$$\lim_{n\to\infty} P(A) = e^{-\lambda t}\,\frac{(\lambda t)^k}{k!}.$$

In conclusione ricaviamo, per $n \to \infty$, che

$$P\{N(t) = k\} = P(A) + P(B) = e^{-\lambda t}\,\frac{(\lambda t)^k}{k!} \qquad (k = 0, 1, \ldots).$$

In conclusione, nelle ipotesi 1, 2 e 3, il numero di eventi che si verificano in un intervallo fissato di lunghezza t si distribuisce secondo una variabile aleatoria di Poisson di media $\lambda\, t$, e diremo che gli eventi si verificano secondo un processo di Poisson di intensità λ.

Il valore λ è uguale al numero medio di eventi che si verificano per unità di tempo: $E[N(1)] = \lambda$; è una costante che si può determinare empiricamente.

Esempio. In un processo di Poisson, (a) ricavare la distribuzione di probabilità dell'istante in cui si verificherà il prossimo evento, (b) determinare la probabilità che avvengano almeno 3 eventi in 2 unità temporali, nel caso di intensità $\lambda = 2$.
Soluzione. (a) Sia X il tempo che trascorrerà fino al prossimo evento. Poiché $X \geq t$ se e solo se nessun evento si verificherà prima delle prossime t unità temporali, si ha

$$P(X > t) = P[N(t) = 0] = e^{-\lambda t}, \qquad t \geq 0,$$

e quindi $F(t) = P(X \leq t) = 1 - P(X > t) = 1 - e^{-\lambda t}$, $t \geq 0$.
(b) Si ha $P[N(2) \geq 3] = 1 - \sum_{k=0}^{2} P[N(2) = k] = 1 - \left(1 + 4 + \dfrac{4^2}{2}\right) e^{-4} = 1 - 13\, e^{-4}$.

Esercizi per casa

4.1) Sia X una variabile aleatoria discreta che assume valori $0, 1, 2, 3$ e tale che

$$p(x) = \frac{1}{2}\, p(x-1) \qquad \text{per } x = 1, 2, 3.$$

(i) Determinare la funzione di probabilità $p(x) = P(X = x)$, per $x = 0, 1, 2, 3$.
(ii) Ricavare $E(X)$.

4.2) Un venditore di automobili ha fissato due appuntamenti. La probabilità di vendere un'automobile nell'i-esimo appuntamento è $p_i = (\frac{1}{2})^i$ $(i = 1, 2)$. Ogni vendita ha la stessa probabilità di riguardare la versione base (del valore di 9000 euro) oppure la versione lusso (del valore di 12000 euro). Indicata con X la variabile aleatoria che descrive il totale dei guadagni del venditore, si determini:
(i) la densità discreta e la funzione di distribuzione di X,
(ii) il valore atteso di X.

4.3) Vi sono due dadi, di cui uno è truccato, nel senso che le facce del dado sono $\{1, 1, 2, 3, 4, 5\}$. Se si sceglie a caso uno dei due dadi e lo si lancia 5 volte, sia X il numero di volte che esce 1. Valutare
(i) $P(X \geq 2)$,
(ii) $E(X)$,
(iii) $Var(X)$.

4.4) Giorgio dispone di due monete truccate: la prima fornisce testa con probabilità $\frac{2}{3}$, la seconda con probabilità $\frac{1}{3}$. Fabrizio sceglie a caso una delle due monete e la lancia finché non esce testa.
(i) Qual è la probabilità che la moneta mostri testa per la prima volta al quarto lancio?
(ii) Qual è la probabilità che siano necessari almeno cinque lanci perché la moneta mostri testa per la prima volta?
(iii) Se Fabrizio sceglie la seconda moneta, qual è il numero medio di lanci che deve effettuare affinché la moneta mostri testa per la prima volta?

4.5) Un esperimento consiste nell'estrarre ripetutamente biglie da un'urna che contiene 4 biglie bianche e 1 nera. Sia Y l'estrazione in cui si estrae la biglia nera per la prima volta. Determinare $P(Y \leq 3)$, $E(Y)$ e $Var(Y)$ nel caso di estrazioni
(i) con reinserimento,
(ii) senza reinserimento.

4.6) Ogni mattina alla stazione ferroviaria arriva esattamente un treno tra le 8:00 e le 9:00, ed un altro treno arriva tra le 9:00 e le 10:00. I tempi di arrivo e le probabilità sono qui riportati:

treno A	8:10	8:30	8:50
treno B	9:10	9:30	9:50
probabilità	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

Se un viaggiatore arriva alle 8:20 qual è il suo tempo medio di attesa del treno?

4.7) Sia X una variabile aleatoria con funzione di distribuzione

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < -2, \\ \frac{1}{3}, & -2 \leq x < 0, \\ \frac{1}{2}, & 0 \leq x < 2, \\ \frac{3}{4}, & 2 \leq x < 3, \\ 1, & x \geq 3. \end{cases}$$

Calcolare
(i) la densità discreta di X,
(ii) $P(X > -1)$, $P(1 < X \leq 3)$, $P(X < 2)$ e $P(X < 2|X > -1)$.

4.8) La variabile aleatoria discreta X può assumere solo i valori -1, 0, 1. Posto $P(X = -1) = a$, $P(X = 0) = b$ e $P(X = 1) = c$, determinare i valori a, b, c in modo che sia $E(X) = \frac{1}{2}$ e $\text{Var}(X) = \frac{3}{10}$.

4.9) In un canale di trasmissione binario il segnale trasmesso X e il segnale ricevuto Y sono entrambi variabili aleatorie che possono assumere solamente i valori "**1**" e "**0**". Si sa che:
– Il segnale trasmesso è "**1**" nel 70% dei casi e "**0**" nei rimanenti.
– A causa di disturbi il segnale trasmesso è ricevuto correttamente nel 95% dei casi se è stato trasmesso "**1**" e nel 98% dei casi se è stato trasmesso "**0**".
Calcolare:
(i) la probabilità $P(Y = 1)$ di ricevere "**1**";
(ii) la probabilità $P(X = 1|Y = 1)$ che sia stato trasmesso "**1**" se viene ricevuto "**1**".

4.10) Una scatola contiene 24 fusibili, di cui 4 difettosi. Antonio sceglie a caso 10 fusibili e Biagio prende i rimanenti.
(i) Con quale probabilità i fusibili difettosi capiteranno tutti ad Antonio (= evento A)?
(ii) Con quale probabilità i fusibili difettosi capiteranno tutti alla stessa persona (= evento B)?
(iii) Con quale probabilità i fusibili difettosi capiteranno tutti ad Antonio sapendo che capiteranno tutti alla stessa persona?
(iv) Qual è il numero atteso di fusibili difettosi che capiteranno ad Antonio?

CAPITOLO 5 – Variabili aleatorie continue

5.1 Introduzione

Le variabili aleatorie discrete assumono un numero finito o un'infinità numerabile di valori.

Esistono comunque variabili aleatorie il cui insieme dei valori è non numerabile, come ad esempio l'ora di arrivo di un treno o il tempo di vita di un dispositivo elettronico.

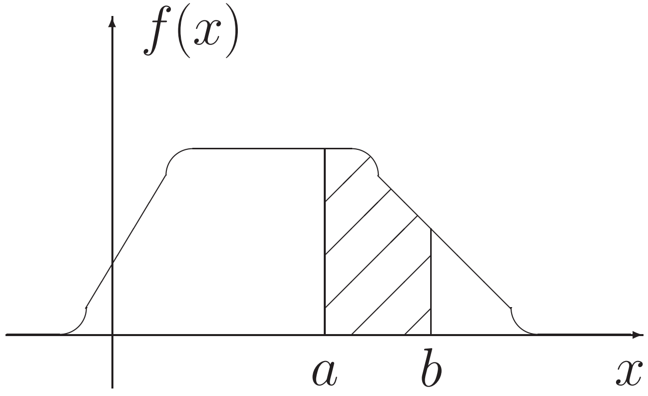
Definizione. Una variabile aleatoria X è detta continua (o, anche, assolutamente continua) se esiste una funzione $f\colon \mathbb{R} \to [0,+\infty)$ tale che per ogni sottoinsieme B di numeri reali risulta

$$P(X \in B) = \int_B f(x)\, dx.$$

La funzione f è chiamata *funzione di densità* della variabile aleatoria X, ed è di fatto caratterizzata dalle seguenti proprietà:

$$f(x) \geq 0 \quad \text{per } x \in \mathbb{R}, \qquad \int_{-\infty}^{\infty} f(x)\, dx = 1.$$

Per $B = (-\infty, x]$ si ha $F(x) = P(X \leq x) = P(X \in (-\infty, x]) = \int_{-\infty}^{x} f(t)\, dt$.



$$P(a \leq X \leq b) = \text{ area della regione tratteggiata } = \int_a^b f(x)\,dx.$$

Tale relazione ricalca la seguente, già vista in passato, che sussiste se X è una variabile aleatoria discreta:

$$P(a \leq X \leq b) = \sum_{k\colon a \leq x_k \leq b} p(x_k).$$

Abbiamo visto che se X è una variabile aleatoria continua, si ha

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

Quindi, se $b = a$ si ricava

$$P(X = a) = \int_a^a f(x)\,dx = 0.$$

Pertanto, a differenza del caso discreto, la probabilità che una variabile aleatoria continua X assuma un singolo valore è uguale a zero, cosicché risulta

$$P(X < a) = P(X \le a) = F(a) = \int_{-\infty}^a f(x)\,dx$$

e quindi

$$F(b) - F(a) = P(a < X \le b) = \int_a^b f(x)\,dx.$$

Ne segue che la funzione di distribuzione $F(x)$ di una variabile aleatoria X continua è una funzione continua per ogni x reale.

Esempio. Sia X una variabile aleatoria continua di densità

$$f(x) = \begin{cases} c(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{altrimenti.} \end{cases}$$

(a) Calcolare c.

(b) Determinare $P(X > 1)$.

Soluzione. (a) Deve essere $f(x) \geq 0$ per ogni $x \in \mathbb{R}$, ossia $c2x(2-x) \geq 0$ per $0 < x < 2$, da cui segue $c \geq 0$. Inoltre, dalla condizione $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$ si ha

$$1 = c\int_0^2 (4x - 2x^2)\,dx = c\left[2x^2 - \frac{2x^3}{3}\right]_{x=0}^{x=2} = c\left[8 - \frac{16}{3}\right] = c\,\frac{8}{3} \quad\Longrightarrow\quad c = \frac{3}{8},$$

che è una soluzione accettabile in quanto soddisfa la condizione $c \geq 0$.

(b) Risulta quindi

$$P(X > 1) = \int_1^{\infty} f(x)\,dx = \frac{3}{8}\int_1^2 (4x - 2x^2)\,dx = \frac{3}{8}\left[2x^2 - \frac{2x^3}{3}\right]_{x=1}^{x=2} = 1 - \frac{1}{2} = \frac{1}{2}.$$

Esempio. Il tempo (in ore) che un computer funzioni prima di bloccarsi è una variabile aleatoria continua X di densità

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

(a) Calcolare la probabilità che il computer funzioni tra 50 e 150 ore senza bloccarsi.

(b) Calcolare la probabilità che il computer si blocchi prima di 100 ore.

Soluzione. (a) Per determinare λ, imponendo che risulti $\int_{-\infty}^{\infty} f(x)\, dx = 1$ otteniamo

$$1 = \lambda \int_0^{\infty} e^{-x/100}\, dx = \lambda \Big[-100\, e^{-x/100}\Big]_{x=0}^{x=\infty} = \lambda\, 100 \quad \Longrightarrow \quad \lambda = \frac{1}{100}.$$

Tale soluzione è accettabile, in quanto risulta $f(x) \geq 0$ per $\lambda \geq 0$. Pertanto si ha

$$P(50 < X < 150) = \int_{50}^{150} \frac{1}{100}\, e^{-x/100}\, dx = \Big[-e^{-x/100}\Big]_{50}^{150} = e^{-1/2} - e^{-3/2} \approx 0{,}383.$$

(b) Analogamente risulta

$$P(X < 100) = \int_0^{100} \frac{1}{100}\, e^{-x/100}\, dx = \Big[-e^{-x/100}\Big]_0^{100} = 1 - e^{-1} \approx 0{,}632.$$

Esempio. Il tempo (in ore) di vita di certe pile per la radio è una variabile aleatoria continua X di densità

$$f(x) = \begin{cases} 0 & x \leq 100 \\ 100/x^2 & x > 100. \end{cases}$$

Determinare la probabilità che esattamente 2 pile della radio su 5 debbano essere sostituite entro le 150 ore di attività, supponendo che gli eventi $E_i = \{$l'i-esima pila va rimpiazzata entro 150 ore d'uso$\}$, $1 \leq i \leq 5$, siano indipendenti.

Soluzione. Risulta $P(E_i) = P(X \leq 150)$, e quindi

$$P(E_i) = \int_{-\infty}^{150} f(x)\,dx = 100 \int_{100}^{150} \frac{1}{x^2}\,dx = 100 \left[-\frac{1}{x}\right]_{x=100}^{x=150} = 100 \left[\frac{1}{100} - \frac{1}{150}\right] = \frac{1}{3}.$$

Pertanto, per l'indipendenza degli eventi E_i, la probabilità richiesta (di tipo binomiale) corrisponde alla probabilità che si realizzino 2 eventi di tipo E_i e 3 eventi di tipo $\overline{E_i}$, quindi si ha:

$$\binom{5}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3 = 10 \cdot \frac{1}{9} \cdot \frac{8}{27} = \frac{80}{243} \approx 0{,}329.$$

La relazione tra la funzione di distribuzione F di una variabile aleatoria continua e la densità f è data da

$$F(x) = P\{X \in (-\infty, x]\} = \int_{-\infty}^{x} f(t)\,dt, \qquad \forall\, x \in \mathbb{R}.$$

Quindi, derivando ambo i membri si ha

$$\frac{d}{dx}F(x) = f(x) \qquad \forall\, x \in \mathbb{R} \ \text{ in cui } F \text{ è derivabile,}$$

cosicché la densità si ottiene derivando la funzione di distribuzione.

Un'interpretazione intuitiva della densità segue da:

$$P\left\{x - \frac{\varepsilon}{2} \leq X \leq x + \frac{\varepsilon}{2}\right\} = \int_{x-\varepsilon/2}^{x+\varepsilon/2} f(t)\,dt \approx \varepsilon\, f(x),$$

dove l'approssimazione vale quando ε è prossimo a 0 e se la densità f è continua in x. Pertanto, la probabilità che X assuma valori in un intorno di x avente ampiezza ε è approssimativamente uguale a $\varepsilon\, f(x)$.

Esempio. Sia X una variabile aleatoria continua con densità

$$f(x) = \begin{cases} c(1-x^2) & -1 \leq x \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Determinare la costante c, la funzione di distribuzione di X, e $P(X \leq 0|X \leq 0{,}5)$.

Soluzione. Imponendo che sia $f(x) \geq 0$ per $x \in \mathbb{R}$ e $\int_{-\infty}^{\infty} f(t)\,dt = 1$ si ha

$$c(1-x^2) \geq 0 \quad \text{per } -1 \leq x \leq 1 \quad \Rightarrow \quad c \geq 0,$$

$$1 = \int_{-\infty}^{\infty} f(t)\,dt = c\int_{-1}^{1}(1-t^2)\,dt = c\left[t - \frac{t^3}{3}\right]_{-1}^{1} = c\left[1 - \frac{1}{3} - \left(-1 + \frac{1}{3}\right)\right] = c\,\frac{4}{3}$$

pertanto si ha $c = \dfrac{3}{4}$ (accettabile in quanto $c \geq 0$). Ricaviamo ora $F(x) = P(X \leq x)$.
Per $x < -1$ si ha:

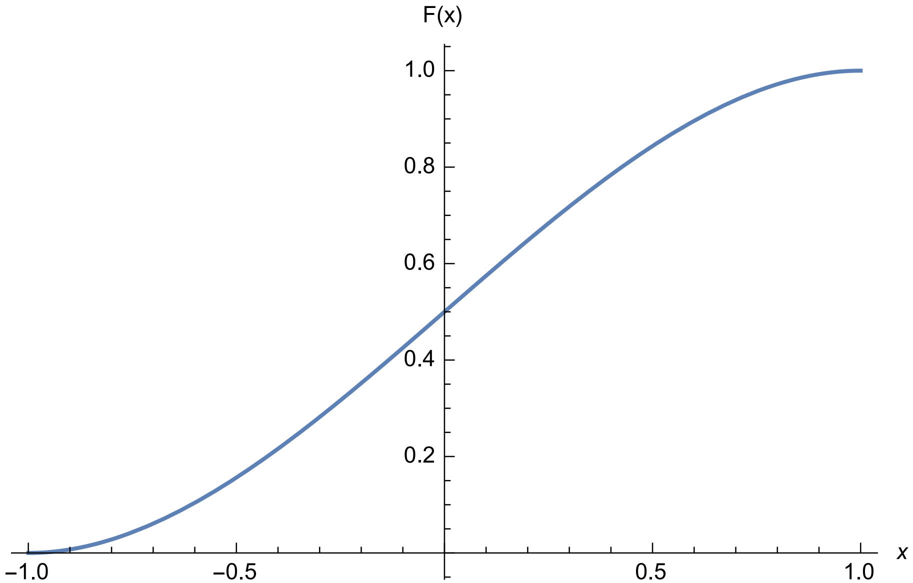
$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0\,dt = 0.$$

Per $-1 \leq x < 1$ si ha:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{3}{4}\int_{-1}^{x}(1-t^2)dt = \frac{3}{4}\left[t - \frac{t^3}{3}\right]_{-1}^{x} = \frac{3}{4}\left[x - \frac{x^3}{3} + \frac{2}{3}\right].$$

Per $x \geq 1$ si ha:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{3}{4}\int_{-1}^{1}(1-t^2)dt = \frac{3}{4}\left[t - \frac{t^3}{3}\right]_{-1}^{1} = \frac{3}{4}\left[1 - \frac{1}{3} + \frac{2}{3}\right] = 1.$$



Infine si ha

$$P(X \leq 0 | X \leq 0{,}5) = \frac{P(X \leq 0, X \leq 0{,}5)}{P(X \leq 0{,}5)} = \frac{P(X \leq 0)}{P(X \leq 0{,}5)} = \frac{F(0)}{F(0{,}5)} = 0{,}5925$$

essendo $F(0) = \frac{1}{2}$ e $F(0{,}5) = \frac{3}{4}\left[\frac{1}{2} - \frac{1}{3}\cdot\frac{1}{8} + \frac{2}{3}\right] = 0{,}84375.$

Esempio. Sia X una variabile aleatoria continua con funzione di distribuzione F_X e densità f_X. Determinare la densità di $Y = aX + b$, con $a \neq 0$.

Soluzione. Se $a > 0$, la funzione di distribuzione di Y è così esprimibile:

$$F_Y(x) = P(Y \leq x) = P(aX + b \leq x) = P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right),$$

mentre per $a < 0$ risulta:

$$F_Y(x) = P\left(X \geq \frac{x-b}{a}\right) = 1 - P\left(X < \frac{x-b}{a}\right) = 1 - F_X\left(\frac{x-b}{a}\right).$$

Derivando rispetto ad x, ricordando che $f(x) = \frac{d}{dx}F(x)$ se $F(x)$ è derivabile in x, per $a \neq 0$ si ottiene infine:

$$f_Y(x) = \frac{d}{dx}F_Y(x) = \begin{cases} \dfrac{1}{a}\, f_X\left(\dfrac{x-b}{a}\right) & \text{se } a > 0 \\ -\dfrac{1}{a}\, f_X\left(\dfrac{x-b}{a}\right) & \text{se } a < 0 \end{cases} = \frac{1}{|a|}\, f_X\left(\frac{x-b}{a}\right).$$

Esempio. Sia X una variabile aleatoria continua con densità

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$

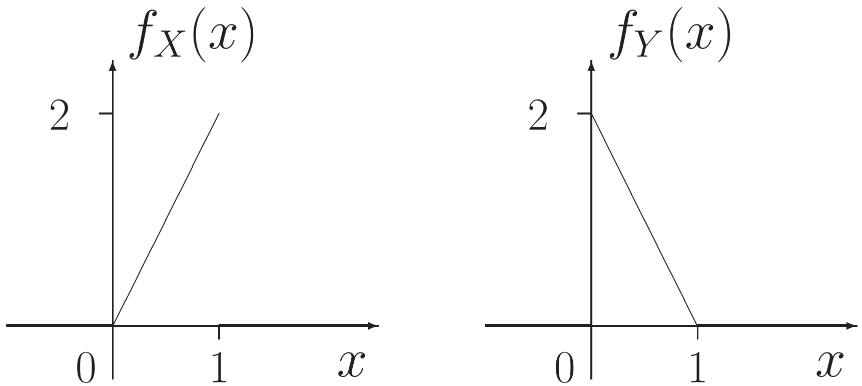
Determinare la densità di $Y = 1 - X$.

Soluzione. Ricordando che la densità di $aX + b$ è data da

$$f_{aX+b}(x) = \frac{1}{|a|} f_X\left(\frac{x-b}{a}\right),$$

per $a = -1$ e $b = 1$ segue che la densità di $Y = aX + b = 1 - X$ è:

$$f_Y(x) = f_X(1-x) = \begin{cases} 2(1-x) & 0 \leq 1-x \leq 1 \\ 0 & \text{altrimenti} \end{cases} = \begin{cases} 2(1-x) & 0 \leq x \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$



Alternativamente, è possibile ricavare la densità di Y notando che la funzione di distribuzione di $Y = 1 - X$ è data da

$$F_Y(x) = P(Y \leq x) = P(1 - X \leq x) = P(X \geq 1 - x) = 1 - P(X < 1 - x).$$

Poiché X è una variabile aleatoria continua, risulta

$$P(X < 1 - x) = P(X \leq 1 - x) = F_X(1 - x).$$

Pertanto la densità di probabilità di Y si può ricavare come segue:

$$f_Y(x) = \frac{d}{dx} F_Y(x) = \frac{d}{dx}\left[1 - F_X(1 - x)\right] = -f_X(1 - x)\frac{d}{dx}(1 - x) = f_X(1 - x).$$

Usando tale indentità e ricordando che la densità di probabilità di X è

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{altrimenti} \end{cases}$$

si ottiene infine la densità di Y:

$$f_Y(x) = \begin{cases} 2(1 - x) & 0 \leq x \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$

5.2 Valore atteso e varianza di una variabile aleatoria continua

Come visto in precedenza, il valore atteso di una variabile aleatoria discreta X è

$$E[X] = \sum_x x\, P(X = x).$$

Analogamente, se X è una variabile aleatoria continua con densità $f(x)$, poiché

$$f(x)\, dx \approx P(x \leq X \leq x + dx) \qquad \text{per } dx \text{ piccolo},$$

si definisce il valore atteso di X come

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

Esempio. Determinare $E[X]$ sapendo che la densità di X è

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Soluzione. Si ha

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_0^1 x\,(2x)\, dx = 2\int_0^1 x^2\, dx = 2\left[\frac{x^3}{3}\right]_0^1 = \frac{2}{3}.$$

Esempio. Determinare $E\left[e^X\right]$ sapendo che la densità di X è

$$f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Soluzione. Per determinare la funzione di distribuzione di $Y = e^X$, notiamo che per $0 < X < 1$ risulta $1 < e^X < e$. Quindi, per $1 \le y \le e$ si ha

$$F_Y(y) = P(Y \le y) = P\left(e^X \le y\right) = P(X \le \log y) = \int_0^{\log y} f_X(x)\,dx = \log y.$$

Derivando $F_Y(y)$ si ottiene la densità di probabilità di Y:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \log y = \frac{1}{y}, \qquad 1 \le y \le e,$$

da cui segue

$$E\left[e^X\right] = E[Y] = \int_{-\infty}^{\infty} y\, f_Y(y)\,dy = \int_1^e y\,\frac{1}{y}\,dy = \int_1^e dy = e - 1 = 1{,}71828.$$

Se X è una variabile aleatoria discreta, sappiamo che

$$E[g(X)] = \sum_i g(x_i)\, p(x_i).$$

È possibile dimostrare il seguente analogo risultato:

Proposizione. Se X è una variabile aleatoria continua con densità $f(x)$, allora per ogni funzione g a valori reali,

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

Ad esempio, applicando tale Proposizione alla variabile aleatoria continua di densità

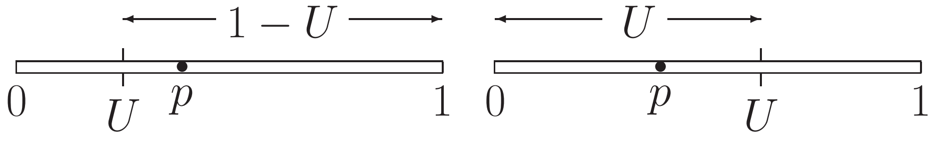
$$f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{altrimenti} \end{cases}$$

per $Y = e^X$ si ottiene

$$E[Y] = E\left[e^X\right] = \int_{-\infty}^{\infty} e^x\, f_X(x)\, dx = \int_0^1 e^x\, dx = [e^x]_0^1 = e - 1 = 1{,}71828.$$

Esempio. Un bastoncino di lunghezza 1 è spezzato in un punto U scelto a caso su di esso, e quindi distribuito uniformemente su $(0, 1)$. Determinare il valore atteso della lunghezza del pezzo di bastoncino che contiene il punto p, $0 \le p \le 1$.

Soluzione. Sia $L(U)$ la lunghezza del pezzo di bastoncino contenente p. Si ha



$$L(U) = \begin{cases} 1 - U & \text{se } 0 < U < p \\ U & \text{se } p < U < 1. \end{cases}$$

Quindi, essendo $f_U(x) = 1$ per $0 \le x \le 1$, e $f_U(x) = 0$ altrimenti, segue che

$$\begin{aligned} E[L(U)] &= \int_{-\infty}^{\infty} L(x)\, f_U(x)\, dx = \int_0^p (1 - x)\, dx + \int_p^1 x\, dx \\ &= \left[x - \frac{x^2}{2}\right]_0^p + \left[\frac{x^2}{2}\right]_p^1 = p - \frac{p^2}{2} + \frac{1}{2} - \frac{p^2}{2} = \frac{1}{2} + p(1 - p). \end{aligned}$$

Il valore atteso della lunghezza del pezzo di bastoncino contente p è massimo se p è il punto centrale, ossia se $p = 1/2$. Notiamo inoltre che $E[U] \le E[L(U)]$, essendo

$$E[U] = \int_{-\infty}^{\infty} x\, f_U(x)\, dx = \int_0^1 x\, dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2} \quad \le \quad \frac{1}{2} + p(1 - p) = E[L(U)].$$

Corollario. Come nel caso discreto, se X è una variabile aleatoria continua, con a e b costanti, sussiste la proprietà di linearità:

$$E[aX + b] = aE[X] + b.$$

Se X è una variabile aleatoria continua di valore atteso $\mu = E[X]$, in analogia col caso discreto la varianza di X è definita da

$$\mathrm{Var}(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx.$$

La formula alternativa,

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \left(\int_{-\infty}^{\infty} x\, f(x)\, dx \right)^2,$$

si dimostra imitando quanto fatto nel caso discreto. Analogamente, se a e b sono delle costanti, allora si ha

$$\mathrm{Var}(aX + b) = a^2 Var(X).$$

Esempio. Sia X una variabile aleatoria continua avente densità

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Determinare la funzione di distribuzione e la varianza di X.

Soluzione. La funzione di distribuzione di X è data da

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0, & x < 0 \\ \int_0^x 2t\,dt = [t^2]_0^x = x^2, & 0 \le x < 1 \\ \int_0^1 2t\,dt = [t^2]_0^1 = 1, & x \ge 1. \end{cases}$$

Ricaviamo la varianza di X ricordando che $\mu = E[X] = 2/3$; pertanto:

$$\begin{aligned} \mathrm{Var}(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx = \int_0^1 \left(x - \frac{2}{3}\right)^2 2x\,dx \\ &= 2\int_0^1 \left(x^3 - \frac{4}{3}x^2 + \frac{4}{9}x\right) dx = 2\left[\frac{x^4}{4} - \frac{4}{9}x^3 + \frac{2}{9}x^2\right]_0^1 = \frac{1}{18}. \end{aligned}$$

Si perviene allo stesso risultato anche notando che

$$\mathrm{Var}(X) = E(X^2) - \mu^2 = \int_0^1 2t^3\,dt - \left(\frac{2}{3}\right)^2 = \left[\frac{t^4}{2}\right]_0^1 - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.$$

Esempio. Sia X una variabile aleatoria continua avente funzione di distribuzione

$$F(x) = \begin{cases} 0 & x < 0 \\ \sqrt{x} & 0 \leq x < 1 \\ 1 & x \geq 1. \end{cases}$$

Determinare la densità, il valore atteso e la varianza di X.

Soluzione. Derivando $F(x)$ si ottiene la densità di X:

$$f(x) = \frac{d}{dx}F(x) = \begin{cases} \dfrac{d}{dx}x^{1/2} = \dfrac{1}{2}\,x^{-1/2} = \dfrac{1}{2\sqrt{x}}, & 0 < x < 1 \\ 0, & \text{altrimenti.} \end{cases}$$

Da questa si ricavano i momenti di X:

$$E\left[X^n\right] = \int_{-\infty}^{\infty} x^n\, f(x)\, dx = \int_0^1 \frac{x^{n-1/2}}{2}\, dx = \frac{1}{2}\left[\frac{x^{n+1/2}}{n+1/2}\right]_0^1 = \frac{1}{2n+1},$$

da cui seguono valore atteso e la varianza di X:

$$E[X] = \frac{1}{3}, \qquad \mathrm{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.$$

5.3 La variabile aleatoria uniforme

Una variabile aleatoria è detta *uniforme* (o *uniformemente* distribuita) sull'intervallo $(0, 1)$ se la sua densità è data da

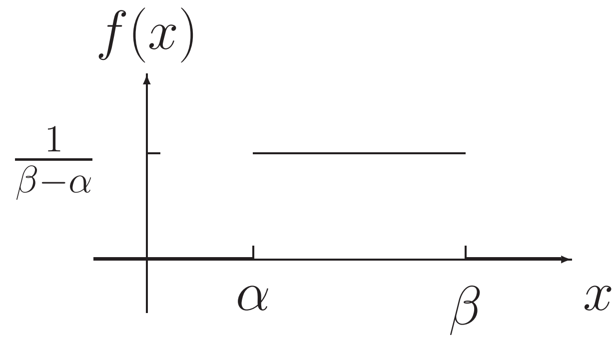
$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{altrimenti.} \end{cases}$$

Risulta $f(x) \geq 0$ e $\int_{-\infty}^{\infty} f(x)\,dx = 1$; inoltre per ogni $0 < a < b < 1$ si ha

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

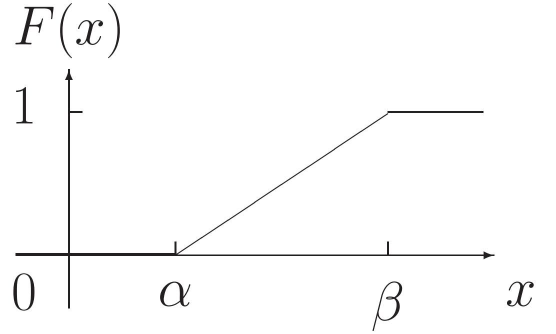
Più in generale, X è una variabile aleatoria uniforme sull'intervallo (α, β) se ha densità

$$f_X(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \alpha < x < \beta \\ 0, & \text{altrimenti.} \end{cases}$$



La funzione di distribuzione di una variabile aleatoria uniforme sull'intervallo (α, β) è

$$F(x) = \begin{cases} 0 & x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha} & \alpha \le x < \beta \\ 1 & x \ge \beta \end{cases}$$



Infatti, per $\alpha < x < \beta$ si ha $f_X(x) = \dfrac{1}{\beta - \alpha}$ e quindi

$$F(x) = \int_{-\infty}^{x} f(u)\, du = \int_{\alpha}^{x} \frac{1}{\beta-\alpha}\, du = \frac{1}{\beta-\alpha}\int_{\alpha}^{x} du = \frac{x-\alpha}{\beta-\alpha}.$$

Esempio. Sia X una variabile aleatoria continua avente distribuzione uniforme su (α, β). Determinare momenti, valore atteso e varianza di X.

Soluzione. Per $n = 1, 2, \ldots$ il momento n-esimo di X è dato da

$$E\left[X^n\right] = \int_{-\infty}^{\infty} x^n\, f(x)\, dx = \int_{\alpha}^{\beta} \frac{x^n}{\beta-\alpha}\, dx = \frac{1}{(\beta-\alpha)} \left[\frac{x^{n+1}}{n+1}\right]_{\alpha}^{\beta} = \frac{\beta^{n+1} - \alpha^{n+1}}{(n+1)(\beta-\alpha)}.$$

In particolare, per $n = 1$ e per $n = 2$ risulta

$$E[X] = \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} = \frac{(\beta - \alpha)(\beta + \alpha)}{2(\beta - \alpha)} = \frac{\alpha + \beta}{2},$$

$$E\left[X^2\right] = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{(\beta - \alpha)(\beta^2 + \alpha\beta + \alpha^2)}{3(\beta - \alpha)} = \frac{\beta^2 + \alpha\beta + \alpha^2}{3}.$$

Possiamo così ricavare facilmente la varianza di X:

$$\mathrm{Var}(X) = E\left[X^2\right] - (E[X])^2 = \frac{\beta^2 + \alpha\beta + \alpha^2}{3} - \frac{(\alpha + \beta)^2}{4}$$

$$= \frac{4(\beta^2 + \alpha\beta + \alpha^2) - 3(\beta^2 + 2\alpha\beta + \alpha^2)}{12} = \frac{\beta^2 - 2\alpha\beta + \alpha^2}{12} = \frac{(\beta - \alpha)^2}{12}.$$

Segue che:

- se α e β sono distanti, allora la varianza di X è grande e quindi il valore medio di X è poco significativo;

- se α e β sono prossimi, allora la varianza di X è piccola e quindi il valore medio di X è molto significativo.

Esempio. Sia X una variabile aleatoria avente distribuzione uniforme su $(-10, 10)$. Calcolare $P(X > 3)$, $P(X < 6)$ e $P(X > 3 \,|\, X < 6)$.

Soluzione. Dall'espressione della funzione di distribuzione, essendo $\alpha = -10$ e $\beta = 10$, si ha

$$F(x) = \frac{x-\alpha}{\beta-\alpha} = \frac{x+10}{20}, \qquad -10 \leq x \leq 10.$$

Quindi si ha

$$P(X > 3) = 1 - P(X \leq 3) = 1 - F(3) = 1 - \frac{13}{20} = \frac{7}{20},$$

$$P(X < 6) = F(6) = \frac{16}{20},$$

$$P(X > 3 \,|\, X < 6) = \frac{P(X > 3, X < 6)}{P(X < 6)} = \frac{P(3 < X < 6)}{P(X < 6)} = \frac{F(6) - F(3)}{F(6)}$$

$$= \frac{(16/20) - (13/20)}{16/20} = \frac{3/20}{16/20} = \frac{3}{16}.$$

Esempio. Gli autobus passano ad una fermata ad intervalli di 15 minuti a partire dalle ore 7, cioè alle 7, 7:15, 7:30, 7:45, ecc. Se un passeggero arriva alla fermata in un istante uniformemente distribuito tra le 7 e le 7:30, determinare la probabilità che egli aspetti l'autobus (a) meno di 5 minuti (b) più di 10 minuti.

Soluzione. Sia X il minuto tra le 7 e le 7:30 in cui arriva il passeggero. Il passeggero aspetterà meno di 5 minuti se (e solo se) egli arriva tra le 7:10 e le 7:15 o tra le 7:25 e le 7:30. Pertanto, poiché X è uniforme sull'intervallo $(0, 30)$, si ha

$$f(x) = \begin{cases} \dfrac{1}{30} & 0 < x < 30 \\ 0 & \text{altrimenti} \end{cases}$$

e quindi la probabilità cercata in (a) è

$$P(10 < X < 15) + P(25 < X < 30) = \int_{10}^{15} \frac{1}{30}\, dx + \int_{25}^{30} \frac{1}{30}\, dx = \frac{5+5}{30} = \frac{1}{3}.$$

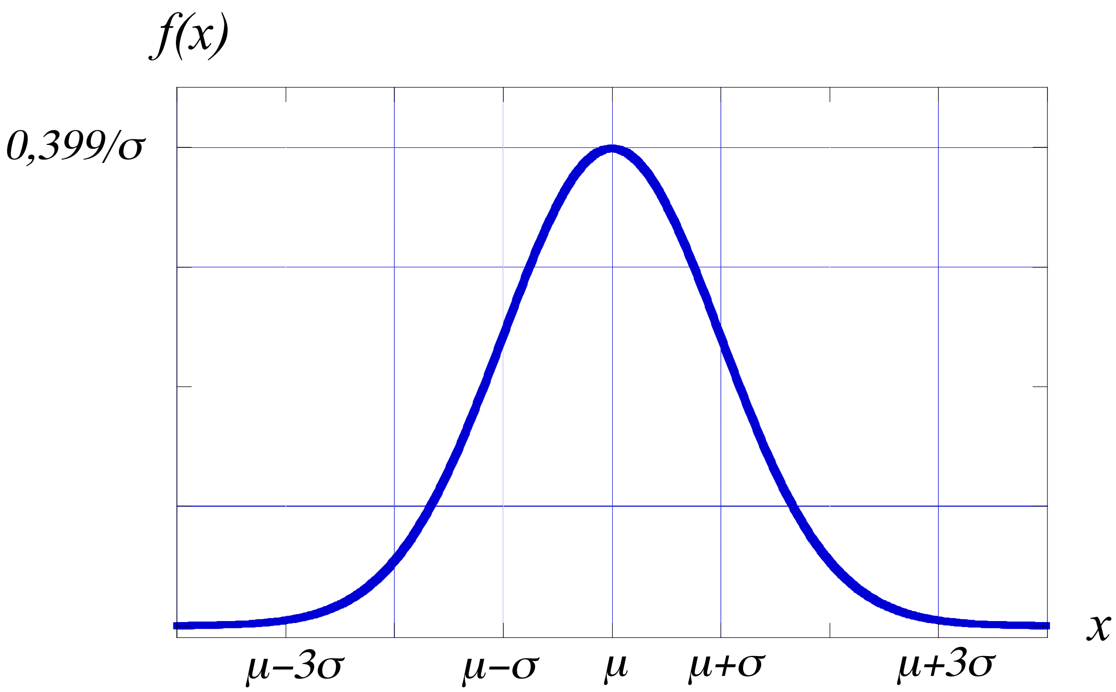
Analogamente, la probabilità cercata in (b) vale

$$P(0 < X < 5) + P(15 < X < 20) = \int_{0}^{5} \frac{1}{30}\, dx + \int_{15}^{20} \frac{1}{30}\, dx = \frac{5+5}{30} = \frac{1}{3}.$$

5.4 Variabili aleatorie normali

Una variabile aleatoria continua X è detta normale (o gaussiana) di parametri $\mu \in \mathbb{R}$ e σ^2, con $\sigma > 0$, se la sua densità di probabilità è data da

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/(2\sigma^2)} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \qquad -\infty < x < \infty.$$



La densità è una funzione con grafico a campana; è simmetrica rispetto a $x = \mu$, ossia risulta $f(\mu - t) = f(\mu + t)$ per ogni $t \geq 0$; possiede massimo in μ e due punti di flesso: $\mu - \sigma$ e $\mu + \sigma$.

Esempio. Mostrare che se X è una variabile aleatoria normale di parametri μ e σ^2, allora valore atteso e varianza sono: $E[X] = \mu$ e $Var(X) = \sigma^2$.
Soluzione. Si ha

$$E[X] = \int_{-\infty}^{\infty} x \, \frac{1}{\sqrt{2\pi}\,\sigma} \, e^{-(x-\mu)^2/2\sigma^2} \, dx.$$

Ponendo $z = (x - \mu)/\sigma$, si ottiene

$$E[X] = \sigma \, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \, e^{-z^2/2} \, dz + \mu \, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \, dz.$$

Il risultato $E[X] = \mu$ segue direttamente da

$$\int_{-\infty}^{\infty} z \, e^{-z^2/2} \, dz = -e^{-z^2/2} \Big|_{z=-\infty}^{z=\infty} = 0, \qquad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \, dz = 1,$$

dove l'ultima uguaglianza si può ricavare da opportune considerazioni.

Pertanto si ha

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \, \frac{1}{\sqrt{2\pi}\,\sigma} \, e^{-(x-\mu)^2/2\sigma^2} \, dx.$$

Ponendo ancora $z = (x-\mu)/\sigma$, si ottiene

$$\mathrm{Var}(X) = \sigma^2 \, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 \, e^{-z^2/2} \, dz = \sigma^2 \, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-z) \, \frac{d}{dz} e^{-z^2/2} \, dz.$$

Quindi (integrando per parti) si ottiene

$$\mathrm{Var}(X) = \sigma^2 \, \frac{1}{\sqrt{2\pi}} (-z) \, e^{-z^2/2} \Big|_{z=-\infty}^{z=\infty} + \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \, dz = \sigma^2.$$

Proposizione. Se X è normale di parametri μ e σ^2, e se $a, b \in \mathbb{R}$, con $a \neq 0$, allora $Y = aX + b$ è una variabile normale di parametri $\mu_Y = a\mu + b$ e $\sigma_Y^2 = a^2\sigma^2$.

Dimostrazione. Ricordando che

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\},$$

la densità di $Y = aX + b$ è data da

$$\begin{aligned} f_Y(x) &= \frac{1}{|a|} f_X\left(\frac{x-b}{a}\right) \\ &= \frac{1}{|a|} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{x-b}{a}-\mu\right)^2\right\} \\ &= \frac{1}{\sqrt{2\pi a^2\sigma^2}} \exp\left\{-\frac{[x-(a\mu+b)]^2}{2a^2\sigma^2}\right\} = \frac{1}{\sqrt{2\pi\sigma_Y^2}} \exp\left\{-\frac{[x-\mu_Y]^2}{2\sigma_Y^2}\right\}, \end{aligned}$$

che è la densità di una variabile aleatoria normale di parametri

$$\mu_Y = a\mu + b \quad \text{e} \quad \sigma_Y^2 = a^2\sigma^2.$$

Ricordiamo che $E[Y] = a\,E[X] + b = a\mu + b$ e $\mathrm{Var}(Y) = a^2\,\mathrm{Var}(X) = a^2\sigma^2$.

Una variabile aleatoria di valore medio 0 e varianza 1 è detta *standard*.

Esempio. Se X è una variabile aleatoria di valore medio μ e varianza σ^2 (con $\sigma > 0$), determinare i valori di a e b tali che $Z = aX + b$ sia standard.

Soluzione. Affinché sia $E(Z) = 0$ e $\mathrm{Var}(Z) = 1$ deve risultare

$$E(Z) = E(aX + b) = a\,E(X) + b = a\mu + b = 0,$$

$$\mathrm{Var}(Z) = \mathrm{Var}(aX + b) = a^2\mathrm{Var}(X) = a^2\sigma^2 = 1.$$

Dall'ultima equazione si ottiene $a^2 = 1/\sigma^2$, da cui si ricavano due soluzioni:

$$a = \frac{1}{\sigma} \quad \Rightarrow \quad \frac{1}{\sigma} \cdot \mu + b = 0 \quad \Rightarrow \quad b = -\frac{\mu}{\sigma} \quad \Rightarrow \quad Z = \frac{X - \mu}{\sigma},$$

$$a = -\frac{1}{\sigma} \quad \Rightarrow \quad -\frac{1}{\sigma} \cdot \mu + b = 0 \quad \Rightarrow \quad b = \frac{\mu}{\sigma} \quad \Rightarrow \quad Z = \frac{\mu - X}{\sigma}.$$

Quindi $Z = aX + b$ è standard per due diverse scelte dei valori di a e b:

$$(a,\, b) = \left(\frac{1}{\sigma},\, -\frac{\mu}{\sigma}\right), \qquad (a,\, b) = \left(-\frac{1}{\sigma},\, \frac{\mu}{\sigma}\right).$$

Dai risultati precedenti segue che se X è una variabile normale di parametri μ e σ^2, allora

$$Z = \frac{X - \mu}{\sigma} \qquad \text{e} \qquad Z' = \frac{\mu - X}{\sigma}$$

sono variabili normali di parametri 0 e 1, ossia sono *variabili normali standard*.

Quindi la densità di probabilità normale

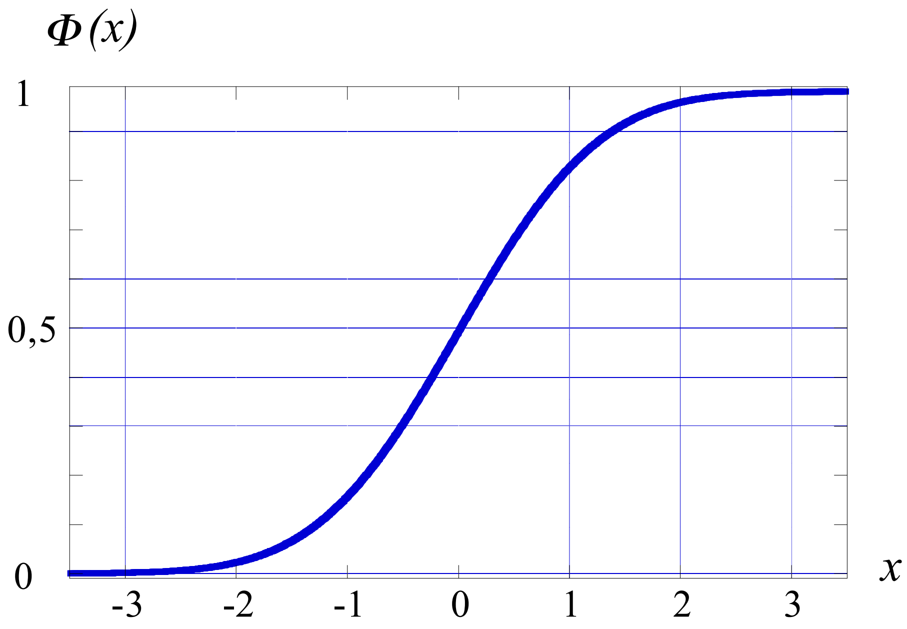
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/(2\sigma^2)}$$

nel caso standard, essendo $\mu = 0$ e $\sigma^2 = 1$, diventa

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}.$$

La funzione di distribuzione di una variabile normale standard Z si indica con $\Phi(x)$:

$$\Phi(x) = P(Z \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2}\, dz, \qquad -\infty < x < \infty.$$



Nella Tabella in Appendice sono dati i valori di $\Phi(x)$ per $x = 0$; 0,01; 0,02; . . .; 3,49. Per valori negativi di x, tale funzione si può ottenere dalla seguente formula:

$$\Phi(-x) = 1 - \Phi(x), \qquad -\infty < x < \infty.$$

Dalla simmetria della densità normale standard $f(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$ segue la relazione di simmetria per $\Phi(x) = P(Z \le x)$:

$$\Phi(-x) = 1 - \Phi(x), \qquad -\infty < x < \infty,$$

la quale afferma che se Z è una variabile normale standard allora

$$P(Z \le -x) = P(Z > x), \qquad -\infty < x < \infty.$$

Inoltre, per $a < b$ si ha

$$P(a \le Z \le b) = P(Z \le b) - P(Z < a) = \Phi(b) - \Phi(a).$$

Ricordiamo che se X è una variabile normale di parametri μ e σ^2, allora $Z = (X - \mu)/\sigma$ è una variabile normale standard, e quindi la funzione di distribuzione di X è

$$F_X(x) = P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P\left(Z \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Ne segue:

$$P(a \le X \le b) = P(X \le b) - P(X < a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Qui di seguito è riportata una routine in Linguaggio C per il calcolo di $\Phi(x)$:

```
//  File: gaussian_distribution.c
//  Routine(s): Gaussian_Distribution

#include <math.h>              // required for erf() and M_SQRT1_2

//  double Gaussian_Distribution( double x )
//
//  Description:
//  This function returns the probability that a random variable with
//  a standard Normal (Gaussian) distribution has a value less than "x".
//
//  Arguments:
//  double x   Argument of Pr[X < x] where X ~ N(0,1).
//
//  Return Values:
//  The probability of observing a value less than (or equal) to x assuming
//  a normal (Gaussian) distribution with mean 0 and variance 1.
//
//  M_SQRT1_2 = 1/(2)=0.707107     (The inverse of the square root of 2)
//
//  Example:
//  double x, pr;
//
//  pr = Gaussian_Distribution(x);

double Gaussian_Distribution( double x )
{
   return  0.5 * ( 1.0 + erf( M_SQRT1_2 * x ) );
}
```

Esempio. Sia X una variabile aleatoria normale di parametri $\mu = 3$ e $\sigma^2 = 9$. Determinare $P(2 < X < 5)$, $P(X > 0)$, $P(|X-3| > 6)$.

Soluzione. Ponendo $Z = (X-\mu)/\sigma$ e ricordando che $\Phi(-x) = 1 - \Phi(x)$ si ha:

$$
\begin{aligned}
P(2 < X < 5) &= P\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right) = P\left(-\frac{1}{3} < Z < \frac{2}{3}\right) \\
&= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right) \approx \Phi\big(0{,}67\big) - \big[1 - \Phi\big(0{,}33\big)\big] \\
&= 0{,}7486 - (1 - 0{,}6293) = 0{,}3779
\end{aligned}
$$

$$
P(X > 0) = P\left(\frac{X-3}{3} > \frac{0-3}{3}\right) = P\left(Z > -1\right) = 1 - \Phi\big(-1\big) = \Phi\big(1\big) = 0{,}8413
$$

$$
\begin{aligned}
P(|X-3| > 6) &= P\big(X - 3 > 6\big) + P\big(-X + 3 > 6\big) \\
&= P\big(X > 9\big) + P\big(X < -3\big) \\
&= P\left(\frac{X-3}{3} > \frac{9-3}{3}\right) + P\left(\frac{X-3}{3} < \frac{-3-3}{3}\right) \\
&= P\left(Z > 2\right) + P\left(Z < -2\right) = 1 - \Phi(2) + \Phi(-2) \\
&= 2\big[1 - \Phi\big(2\big)\big] = 2(1 - 0{,}9772) = 0{,}0456
\end{aligned}
$$

Esempio. In un test d'esame in cui il punteggio ottenuto è una variabile aleatoria normale X di parametri μ e σ^2, viene attribuita la lettera **A** a chi ha un punteggio superiore a $\mu+\sigma$, **B** a chi ha un punteggio tra μ e $\mu+\sigma$, **C** a chi ha un punteggio tra $\mu-\sigma$ e μ, **D** a chi ha punteggio tra $\mu-2\sigma$ e $\mu-\sigma$, ed **E** a chi ottiene un punteggio inferiore a $\mu-2\sigma$. Determinare le percentuali dei giudizi **A**, **B**, **C**, **D**, **E**.

Soluzione. Dato che $Z=(X-\mu)/\sigma$ ha distribuzione normale standard, si ha

$$P(X>\mu+\sigma)=P\Big(\frac{X-\mu}{\sigma}>1\Big)=1-\Phi\big(1\big)=1-0{,}8413=0{,}1587$$

$$P(\mu<X<\mu+\sigma)=P\Big(0<\frac{X-\mu}{\sigma}<1\Big)=\Phi\big(1\big)-\Phi\big(0\big)=0{,}8413-0{,}5=0{,}3413$$

$$P(\mu-\sigma<X<\mu)=P\Big(-1<\frac{X-\mu}{\sigma}<0\Big)=\Phi\big(0\big)-\Phi\big(-1\big)=0{,}3413$$

$$P(\mu-2\sigma<X<\mu-\sigma)=P\Big(-2<\frac{X-\mu}{\sigma}<-1\Big)=\Phi\big(2\big)-\Phi\big(1\big)=0{,}1359$$

$$P(X<\mu-2\sigma)=P\Big(\frac{X-\mu}{\sigma}<-2\Big)=\Phi\big(-2\big)=1-\Phi\big(2\big)=1-0{,}9772=0{,}0228$$

da cui segue che il 16% otterrà **A**, il 34% **B**, il 34% **C**, il 14% **D** ed il 2% **E**.

Esempio. Un bit deve essere trasmesso attraverso un canale soggetto a rumore. Per ridurre la possibilità di errore si invia il valore 2 quando il messaggio è **1** ed il valore -2 quando il messaggio è **0**. Se si spedisce x, allora il valore ricevuto è $R = x + Z$, per $x = \pm 2$, dove Z è una variabile aleatoria normale standard che descrive il rumore del canale. Si determini la probabilità di errore se il valore ricevuto è così decodificato:

$$\text{se } R \geq 0{,}5 \ \text{ si decodifica } \mathbf{1},$$
$$\text{se } R < 0{,}5 \ \text{ si decodifica } \mathbf{0}.$$

Soluzione. Si hanno 2 tipi di errore: uno è che il messaggio **1** sia erroneamente interpretato come **0**, l'altro è che **0** sia erroneamente interpretato come **1**. Il primo tipo di errore si realizza se il messaggio è **1** e $R = 2 + Z < 0{,}5$, il secondo si realizza se il messaggio è **0** e $R = -2 + Z \geq 0{,}5$. Posto $p_{\mathbf{k}} = P(\text{errore} \,|\, \text{il messaggio è}\, \mathbf{k})$, si ha

$$p_{\mathbf{1}} = P\big(2 + Z < 0{,}5\big) = P\big(Z < -1{,}5\big) = \Phi\big(-1{,}5\big) = 1 - \Phi\big(1{,}5\big) = 0{,}0668$$
$$p_{\mathbf{0}} = P\big(-2 + Z \geq 0{,}5\big) = P\big(Z \geq 2{,}5\big) = 1 - \Phi\big(2{,}5\big) = 0{,}0062.$$

Quesito. Cosa cambia se nel criterio di decodifica si usa la soglia 0 invece di 0,5?

5.5 Variabili aleatorie esponenziali

Una variabile aleatoria continua X è detta esponenziale (o distribuita esponenzialmente) di parametro $\lambda > 0$ se la densità di X è data da

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x} & x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

La funzione di distribuzione della variabile aleatoria esponenziale è data da

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \lambda \mathrm{e}^{-\lambda t} dt = \left[-\mathrm{e}^{-\lambda t}\right]_0^x = 1 - \mathrm{e}^{-\lambda x}, \qquad x \geq 0,$$

mentre è $F(x) = 0$ per $x < 0$.

Si noti che $\displaystyle\int_{-\infty}^{\infty} f(t)dt = \int_{0}^{\infty} \lambda \mathrm{e}^{-\lambda t} dt = \lim_{x\to\infty} F(x) = \lim_{x\to\infty} \left(1 - \mathrm{e}^{-\lambda x}\right) = 1.$

Inoltre, per $x \geq 0$ risulta

$$P(X > x) = 1 - F(x) = 1 - (1 - \mathrm{e}^{-\lambda x}) = \mathrm{e}^{-\lambda x}.$$

Esempio. Se X è una variabile aleatoria esponenziale di parametro λ, allora valore atteso e varianza sono: $E[X] = 1/\lambda$ e $Var(X) = 1/\lambda^2$.

Soluzione. Ricordando l'espressione della densità esponenziale, si ha

$$E[X] = \int_0^\infty x\, f(x)\, dx = \int_0^\infty x\lambda \mathrm{e}^{-\lambda x} dx$$

e quindi, integrando per parti, si ottiene

$$E[X] = -x\mathrm{e}^{-\lambda x}\Big|_0^\infty + \int_0^\infty \mathrm{e}^{-\lambda x} dx = \frac{1}{\lambda}\int_0^\infty f(x)dx = \frac{1}{\lambda}.$$

Per calcolare la varianza di X dobbiamo prima calcolare $E[X^2]$. Si ha

$$E[X^2] = \int_0^\infty x^2\lambda \mathrm{e}^{-\lambda x}\mathrm{d}x = -x^2\mathrm{e}^{-\lambda x}\Big|_0^{+\infty} + \int_0^\infty 2x\mathrm{e}^{-\lambda x}\mathrm{d}x = \frac{2}{\lambda^2}.$$

Quindi

$$Var(X) = E[X^2] - (E[X])^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

Esercizio. Se X è una variabile aleatoria esponenziale di parametro λ, mostrare che $Z = \lambda X - 1$ è una variabile aleatoria standard.

Si dice *mediana* di una variabile aleatoria qualsiasi X ogni valore m tale che

$$P(X \leq m) \geq \frac{1}{2} \quad \text{e} \quad P(X \geq m) \geq \frac{1}{2},$$

ossia tale che $F_X(m) \geq \frac{1}{2}$ e $1 - F_X(m^-) \geq \frac{1}{2}$, ovvero

$$F_X(m^-) \leq \frac{1}{2} \leq F_X(m) \qquad \left(F_X(m) = \frac{1}{2} \quad \text{se } X \text{ è continua} \right).$$

Esempio. Sia X una variabile aleatoria esponenziale di parametro λ. Valutare $P\{X > E(X)\}$ e determinare il valore m tale che $P\{X > m\} = 1/2$.

Soluzione. Ricordando che $P(X > x) = \mathrm{e}^{-\lambda x}$ per $x \geq 0$, da $E(X) = \lambda^{-1}$ segue

$$P\{X > E(X)\} = P\{X > \lambda^{-1}\} = \mathrm{e}^{-\lambda \cdot \lambda^{-1}} = \mathrm{e}^{-1} = 0{,}3679.$$

Risulta $P\{X > m\} = \dfrac{1}{2}$ per $\mathrm{e}^{-\lambda m} = \dfrac{1}{2}$ ovvero per $-\lambda m = \log \dfrac{1}{2} = -\log 2$ da cui segue

$$m = \frac{1}{\lambda} \log 2 = \frac{1}{\lambda} 0{,}6931 \qquad \text{(mediana di } X\text{)}.$$

Esercizio. Mostrare che se X è una variabile aleatoria uniforme oppure una variabile aleatoria normale, allora la mediana coincide con $E(X)$.

Una variabile aleatoria X ha la proprietà di assenza di memoria se

$$P(X > s + t \,|\, X > t) = P(X > s) \qquad \text{per ogni } s, t \geq 0.$$

Se X descrive il tempo di vita di un dispositivo, la proprietà di assenza di memoria afferma che la probabilità che esso funzioni per almeno $s + t$ ore, sapendo che ha funzionato per t ore, è uguale alla probabilità iniziale che esso funzioni per almeno s ore. In altri termini, se il dispositivo ha funzionato per t ore, la distribuzione del tempo di vita residuo è uguale a quella del tempo di vita originale.

Essendo $P(X > s + t \,|\, X > t) = \dfrac{P(X > s + t, X > t)}{P(X > t)} = \dfrac{P(X > s + t)}{P(X > t)}$, la condizione di assenza di memoria può essere scritta equivalentemente come

$$P(X > s + t) = P(X > t)\, P(X > s) \qquad \text{per ogni } s, t \geq 0.$$

La variabile aleatoria esponenziale soddisfa la proprietà di assenza di memoria, essendo

$$P(X > s + t) = e^{-\lambda(s+t)} = e^{-\lambda t}\, e^{-\lambda s} = P(X > t)\, P(X > s), \qquad s, t \geq 0,$$

ed è l'unica variabile aleatoria continua a godere di tale proprietà. Notiamo che $e^{-\lambda}$ per la variabile esponenziale ha lo stesso ruolo di $1 - p$ per la variabile geometrica.

Esempio. Sia X la durata aleatoria di una telefonata avente distribuzione esponenziale con durata media di 10 minuti. Calcolare la probabilità che la durata sia

(a) maggiore di 5 minuti;

(b) tra 10 e 20 minuti;

(c) tra 15 e 20 minuti;

(d) tra 15 e 20 minuti sapendo che è maggiore di 10 minuti.

Soluzione. Poiché X ha distribuzione esponenziale con $E(X) = \frac{1}{\lambda} = 10$, e quindi con parametro $\lambda = \frac{1}{10}$, la sua funzione di distribuzione è $F(x) = 1 - e^{-x/10}$, $x \geq 0$, pertanto

$$P(X > 5) = 1 - F(5) = 1 - (1 - e^{-\frac{1}{10}5}) = e^{-0,5} \simeq 0{,}6065;$$

$$P(10 < X < 20) = F(20) - F(10) = 1 - e^{-2} - (1 - e^{-1}) = e^{-1} - e^{-2} \simeq 0{,}2325;$$

$$P(15 < X < 20) = P(X > 15) - P(X > 20) = e^{-1,5} - e^{-2} =\simeq 0{,}0878;$$

infine, dalla proprietà di assenza di memoria si ha

$$\begin{aligned} P(15 < X < 20 \,|\, X > 10) &= P(X > 15 \,|\, X > 10) - P(X > 20 \,|\, X > 10) \\ &= P(X > 5) - P(X > 10) = e^{-0,5} - e^{-1} \simeq 0{,}2387. \end{aligned}$$

Esempio. Consideriamo un ufficio postale con due sportelli, e supponiamo che quando Carlo entra nell'ufficio veda Alice e Bruno agli sportelli. Supponiamo che il Carlo acceda al primo sportello che si libera. Se il tempo che un impiegato dedica ad un cliente è distribuito esponenzialmente con parametro λ, qual è la probabilità che, sui tre clienti, Carlo sia l'ultimo a lasciare l'ufficio?

Soluzione. Prendiamo in esame il tempo necessario affinché Carlo trovi uno sportello libero. A questo punto, Alice o Bruno hanno appena lasciato lo sportello e uno dei due può trovarsi ancora allo sportello. Tuttavia, per la proprietà di assenza di memoria della variabile esponenziale, il tempo residuo di permanenza allo sportello di quest'altra persona (Alice o Bruno) è distribuito esponenzialmente con parametro λ. Per simmetria la probabilità che questa persona finisca prima di Carlo è $1/2$, in quanto anche il tempo di permanenza allo sportello di Carlo è distribuito esponenzialmente con parametro λ.

Esempio. Si supponga che il numero di chilometri che un'auto possa percorrere prima che la batteria ceda sia una variabile aleatoria esponenziale di valore atteso 10000. Se si desidera fare un viaggio di 5000 chilometri, qual è la probabilità di effettuarlo senza dover cambiare la batteria? Cosa si può dire se la distribuzione non è esponenziale?

Soluzione. Dalla proprietà di assenza di memoria della variabile esponenziale segue che il tempo di vita rimanente (in migliaia di chilometri) della batteria è una variabile esponenziale di parametro $\lambda = 1/10$. La probabilità cercata vale quindi

$$P(\text{tempo di vita rimanente} > 5) = 1 - F(5) = \mathrm{e}^{-5\lambda} = \mathrm{e}^{-1/2} \simeq 0{,}6065.$$

Se la funzione di distribuzione $F(x)$ non fosse esponenziale, la probabilità sarebbe

$$P(\text{tempo di vita} > 5 + t \,|\, \text{tempo di vita} > t) = \frac{1 - F(t+5)}{1 - F(t)},$$

dove t è il numero di chilometri già percorsi prima dell'inizio del viaggio. Quindi, se la distribuzione non è esponenziale, occorre una informazione aggiuntiva (il valore di t) per poter calcolare la probabilità richiesta.

($\star\star$) 5.6 Distribuzione di una funzione di variabile aleatoria

Supponiamo di conoscere la distribuzione di X e di voler determinare la distribuzione di $g(X)$. A tal fine è necessario esprimere l'evento $\{g(X) \leq y\}$ in termini dell'appartenenza di X ad un insieme.

Esempio. Sia X uniformemente distribuita su $(0, 1)$. Determinare la funzione di distribuzione della variabile aleatoria $Y = -(\log X)/\lambda$, con $\lambda > 0$.
Soluzione. Ricordando che $F_X(x) = x$ per $0 \leq x \leq 1$, per $y > 0$ si ha

$$F_Y(y) = P(Y \leq y) = P(-\frac{1}{\lambda} \log X \leq y) = P(X \geq e^{-\lambda y}) = 1 - F_X(e^{-\lambda y}) = 1 - e^{-\lambda y}.$$

Pertanto, Y ha distribuzione esponenziale di parametro λ.

Esempio. Sia X uniformemente distribuita su $(0,1)$. Determinare la funzione di distribuzione e la densità di probabilità della variabile aleatoria $Y = X^\alpha$, con $\alpha > 0$.

Soluzione. Poiché $F_X(x) = x$ per $0 \le x \le 1$, per $0 \le y \le 1$ si ha

$$F_Y(y) = P(Y \le y) = P(X^\alpha \le y) = P(X \le y^{1/\alpha}) = F_X(y^{1/\alpha}) = y^{1/\alpha}.$$

Pertanto, la densità di probabilità di Y è data da

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{\alpha}\, y^{1/\alpha - 1}, \qquad 0 < y < 1.$$

Ad esempio, per $\alpha = 2$ si ha $f_Y(y) = \dfrac{1}{2\sqrt{y}}$, per $\alpha = 1/2$ si ha $f_Y(y) = 2y$, $0 < y < 1$.

Esempio. Sia X una variabile aleatoria continua con densità $f_X(x)$. Determinare la funzione di distribuzione e la densità di probabilità di $Y = X^2$ e $U = |X|$.

Soluzione. La funzione di distribuzione di $Y = X^2$, per $y \geq 0$, è data da

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Derivando rispetto a y si ottiene la densità

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{2\sqrt{y}} [f_X(\sqrt{y}) + f_X(-\sqrt{y})], \qquad y > 0.$$

Determiniamo ora la funzione di distribuzione di $U = |X|$. Per $u \geq 0$, si ha

$$F_U(u) = P(U \leq u) = P(|X| \leq u) = P(-u \leq X \leq u) = F_X(u) - F_X(-u).$$

La densità è pertanto

$$f_U(u) = \frac{d}{du} F_U(u) = f_X(u) + f_X(-u), \qquad u > 0.$$

Teorema. Sia X una variabile aleatoria continua con densità f_X, e $g(x)$ una funzione strettamente monotona (crescente o decrescente) e derivabile con continuità. Allora la variabile aleatoria Y definita da $Y = g(X)$ è continua e la sua densità è data da

$$f_Y(y) = \begin{cases} f_X[g^{-1}(y)] \left| \dfrac{d}{dy} g^{-1}(y) \right| & \text{se } y = g(x) \text{ per qualche } x \\ 0 & \text{se } y \neq g(x) \text{ per ogni } x, \end{cases}$$

dove $g^{-1}(y)$ è l'unica x tale che $y = g(x)$.

Dimostrazione. Supponiamo $g(x)$ crescente. Se y appartiene a tale intervallo è $y = g(x)$ per qualche x. Allora, se $Y = g(X)$, si ha

$$F_Y(y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)).$$

In accordo con l'enunciato del teorema, derivando si ottiene

$$f_Y(y) = f_X[g^{-1}(y)] \frac{d}{dy} g^{-1}(y),$$

dato che $g^{-1}(y)$ è non decrescente e la sua derivata è quindi non negativa.

Se y non appartiene all'insieme dei valori di g, cioè $y \neq g(x)$ per ogni x, allora $F_Y(y)$ è uguale a 0 o a 1 ed in ogni caso $f_Y(y) = 0$.

Esempio. Sia X una variabile aleatoria continua non negativa con densità f e sia $Y = X^n$. Determinare la densità $f_Y(y)$.

Soluzione. Se $g(x) = x^n$, allora

$$g^{-1}(y) = y^{1/n}, \qquad \frac{d}{dy}\, g^{-1}(y) = \frac{1}{n}\, y^{1/n-1}.$$

In virtù del teorema si ha pertanto: $f_Y(y) = \dfrac{1}{n}\, y^{1/n-1} f(y^{1/n})$.

Esempio. Sia X una variabile aleatoria uniforme in $(0, 1)$ e sia $\lambda > 0$. Mostrare che la densità di $Y = g(X) = -(\log X)/\lambda$ è esponenziale di parametro λ.

Soluzione. Risulta $g^{-1}(y) = \mathrm{e}^{-\lambda y}$, e quindi per $y > 0$ si ha:

$$f_Y(y) = f_X[g^{-1}(y)] \left|\frac{d}{dy} g^{-1}(y)\right| = 1 \cdot \left|\frac{d}{dy} \mathrm{e}^{-\lambda y}\right| = \lambda\, \mathrm{e}^{-\lambda y}.$$

Esercizi per casa

5.1) Sia $F(x) = 0$ per $x < 1$, $F(x) = c\,(x-1)^2$ per $1 \leq x < 2$, $F(x) = 1$ per $x \geq 2$.
(i) Ricavare i valori di c per cui $F(x)$ è la funzione di distribuzione di una variabile aleatoria X.
(ii) Stabilire per quale valore di c la variabile aleatoria X è continua, ed in tal caso ricavare $E(X)$.

5.2) Sia X una variabile aleatoria continua con densità di probabilità

$$f_X(x) = \begin{cases} cx^2 + |x| & -1/2 < x < 1/2, \\ 0 & \text{altrimenti.} \end{cases}$$

(i) Determinare i valori ammissibili per c.
(ii) Ricavare la funzione di distribuzione di X.
(iii) Calcolare $E(X)$ e $Var(X)$.

5.3) Il tempo di funzionamento (in ore) di una batteria di un computer è descritto da una variabile aleatoria esponenziale. Sapendo che la durata media della batteria è di 3 ore, calcolare:
(i) la probabilità che la batteria duri più di 4 ore,
(ii) la probabilità che la batteria duri in totale meno di 4 ore sapendo che il computer è già funzionante da un'ora.

5.4) Sia X una variabile aleatoria con distribuzione normale tale che $E(X+2) = 1$ e $Var(2X+2) = 4$.
(i) Calcolare $P(X \geq 0)$ e $P(-1{,}55 \leq X \leq 1{,}55 \,|\, X \geq 0)$.
(ii) Determinare il valore di c tale che $P(X > c) = 0{,}30$.

5.5) Sia X una variabile aleatoria uniformemente distribuita nell'intervallo (a, b).
(i) Determinare la funzione di distribuzione di X sapendo che $E(X) = -1/2$ e $Var(X) = 3/4$.
(ii) Calcolare $P(-1 < X < 0 \,|\, X < 1/2)$.

5.6) Sia X una variabile aleatoria che descrive la durata dell'esecuzione di un algoritmo, e sia

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < 0, \\ 1 - \dfrac{1}{x+1} & x \geq 0. \end{cases}$$

(i) Determinare la densità di probabilità di X.

(ii) Calcolare $P(X \leq 3)$, $P(X > 1)$ e $P(X \leq 3 \,|\, X > 1)$.

(iii) Determinare la mediana di X.

Calcolo delle Probabilità e Statistica Matematica – A.A. 2009/10

CAPITOLO 6 – Leggi congiunte di variabili aleatorie

6.1 Funzioni di distribuzione congiunte

6.2 Variabili aleatorie indipendenti

6.3 Somme di variabili aleatorie indipendenti

6.4 Distribuzioni condizionate: il caso discreto

6.1 Funzioni di distribuzione congiunte

Definizione. Date due variabili aleatorie X e Y, la funzione di distribuzione congiunta della variabile aleatoria doppia, o vettore aleatorio bidimensionale, (X, Y) è

$$F(x, y) = P\left(X \leq x, Y \leq y\right), \qquad x, y \in \mathbb{R},$$

dove si intende $P\left(X \leq x, Y \leq y\right) = P\left(\{X \leq x\} \cap \{Y \leq y\}\right)$. La funzione di distribuzione di X può essere ottenuta da quella congiunta di (X, Y) come segue:

$$\begin{aligned} \lim_{y\to+\infty} F(x, y) &= \lim_{y\to+\infty} P\left(X \leq x, Y \leq y\right) = P\left(\lim_{y\to+\infty} \{X \leq x, Y \leq y\}\right) \\ &= P\left(X \leq x, Y < +\infty\right) = P\left(X \leq x\right) = F_X(x). \end{aligned}$$

In maniera analoga la funzione di distribuzione di Y è data da

$$\lim_{x\to+\infty} F(x, y) = F_Y(y).$$

Le funzioni $F_X(x)$ e $F_Y(y)$ sono dette funzioni di distribuzione *marginali* di X e di Y.

Notiamo che sussistono i seguenti limiti:

$$\lim_{x\to+\infty}\lim_{y\to+\infty} F(x,y) = P\left(X < +\infty, Y < +\infty\right) = 1,$$

$$\lim_{x\to-\infty} F(x,y) = P\left(X < -\infty, Y \le y\right) = 0, \qquad \lim_{y\to-\infty} F(x,y) = 0.$$

Probabilità riferite a (X,Y) si possono esprimere in termini di $F(x,y)$.

Ad esempio, facendo uso della formula di De Morgan e della formula di inclusione/esclusione, la probabilità che X sia maggiore di a e Y sia maggiore di b è data da

$$\begin{aligned} P\left(X > a, Y > b\right) &= 1 - P\left(\{X \le a\} \cup \{Y \le b\}\right) \\ &= 1 - \left[P\left(X \le a\right) + P\left(Y \le b\right) - P\left(X \le a, Y \le b\right)\right] \\ &= 1 - F_X(a) - F_Y(b) + F(a,b). \end{aligned}$$

Inoltre, per $a_1 < a_2$ e $b_1 < b_2$ risulta

$$\begin{aligned} &P\left(a_1 < X \le a_2, b_1 < Y \le b_2\right) \\ &\qquad = P\left(X \le a_2, b_1 < Y \le b_2\right) - P\left(X \le a_1, b_1 < Y \le b_2\right) \\ &\qquad = F(a_2,b_2) - F(a_2,b_1) - F(a_1,b_2) + F(a_1,b_1). \end{aligned}$$

Procedendo al limite per $a_2 \to \infty$ e $b_2 \to \infty$ si ottiene la relazione precedente.

Se X e Y sono variabili aleatorie discrete, la densità discreta congiunta di (X, Y) è

$$p(x, y) = P\left(X = x, Y = y\right).$$

Notiamo che

$$p(x, y) \geq 0 \quad \forall x, y, \qquad \sum_x \sum_y p(x, y) = 1.$$

Per ogni sottoinsieme $\mathcal{D}$ di $\mathbb{R}^2$ si ha

$$P\{(X, Y) \in \mathcal{D}\} = \sum_{(x,y)\in\mathcal{D}} p(x, y).$$

Quindi la funzione di distribuzione congiunta di (X, Y) è così esprimibile:

$$F(x, y) = P\left(X \leq x, Y \leq y\right) = \sum_{u\,:\,u\leq x} \sum_{v\,:\,v\leq y} p(u, v).$$

Le densità discrete di X e di Y (dette marginali) si ottengono da $p(x, y)$ come segue:

$$p_X(x) = P\left(X = x\right) = \sum_{y\,:\,p(x,y)>0} p(x, y),$$

$$p_Y(y) = P\left(Y = y\right) = \sum_{x\,:\,p(x,y)>0} p(x, y).$$

Esempio. Mostrare che $p(x,y)$, definita da

$$p(0,0) = \frac{1}{4}, \qquad p(1,0) = p(0,1) = \frac{1}{8}, \qquad p(1,1) = \frac{1}{2},$$

è una densità discreta congiunta di una variabile aleatoria doppia discreta (X, Y). Ricavare le densità discrete di X e di Y, e calcolare $P(X = Y)$ e $P(X \neq Y)$.

Soluzione. Risulta $p(x,y) > 0$, per ogni $x, y \in \{0,1\}$, ed inoltre

$$\sum_{x=0}^{1}\sum_{y=0}^{1} p(x,y) = \frac{1}{4} + 2 \cdot \frac{1}{8} + \frac{1}{2} = 1.$$

Le densità discrete (marginali) di X e di Y sono:

$$p_X(0) = p(0,0) + p(0,1) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}, \qquad p_X(1) = p(1,0) + p(1,1) = \frac{1}{8} + \frac{1}{2} = \frac{5}{8},$$

$$p_Y(0) = p(0,0) + p(1,0) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}, \qquad p_Y(1) = p(0,1) + p(1,1) = \frac{1}{8} + \frac{1}{2} = \frac{5}{8}.$$

Infine, si ha:

$$P(X = Y) = \sum_{x=y} p(x,y) = p(0,0) + p(1,1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}, \qquad P(X \neq Y) = \frac{1}{4}.$$

Se X e Y sono variabili aleatorie discrete, che assumono valori in $\{x_1, x_2, \ldots, x_n\}$ e $\{y_1, y_2, \ldots, y_m\}$, rispettivamente, è possibile riportare i valori della densità discreta congiunta $p(x, y) = P(X = x, Y = y)$ in una tabella:

$x \setminus y$	y_1	y_2	$\cdots$	y_m	$p_X(x)$
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	$\cdots$	$p(x_1, y_m)$	$p_X(x_1)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	$\cdots$	$p(x_2, y_m)$	$p_X(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_n	$p(x_n, y_1)$	$p(x_n, y_2)$	$\cdots$	$p(x_n, y_m)$	$p_X(x_n)$
$p_Y(y)$	$p_Y(y_1)$	$p_Y(y_2)$	$\cdots$	$p_Y(y_m)$	1

Ai margini della tabella sono riportate le densità discrete di X e Y, ricavate da:

$$p_X(x_i) = P(X = x_i) = \sum_{j=1}^{m} p(x_i, y_j), \qquad i = 1, 2, \ldots, n$$

$$p_Y(y_j) = P(Y = y_j) = \sum_{i=1}^{n} p(x_i, y_j), \qquad j = 1, 2, \ldots, m.$$

Esempio. Un esperimento consiste nello scegliere a caso una sequenza booleana di lunghezza 5, avente 2 bit pari a **1**. Sia X la posizione in cui si trova il secondo bit pari a **1** e sia Y la lunghezza della più lunga sottosequenza composta da bit pari a **0**. Ricavare la densità congiunta di (X, Y) e le densità discrete di X e Y.

Soluzione. Lo spazio campionario è costituito da $|S| = \binom{5}{2} = 10$ sequenze equiprobabili. Quindi si ha:

	X	Y		X	Y
0 0 0 1 1	5	3	**0 1 1 0 0**	3	2
0 0 1 0 1	5	2	**1 0 0 0 1**	5	3
0 0 1 1 0	4	2	**1 0 0 1 0**	4	2
0 1 0 0 1	5	2	**1 0 1 0 0**	3	2
0 1 0 1 0	4	1	**1 1 0 0 0**	2	3

$x \setminus y$	1	2	3	$p_X(x)$
2	0	0	1/10	1/10
3	0	2/10	0	2/10
4	1/10	2/10	0	3/10
5	0	2/10	2/10	4/10
$p_Y(y)$	1/10	6/10	3/10	1

Le densità discrete di X e Y, dette *densità marginali*, si ottengono calcolando le somme sulle righe e sulle colonne, rispettivamente, e appaiono ai margini della tabella:

$$p_X(x) = \sum_y p(x, y), \qquad p_Y(y) = \sum_x p(x, y).$$

Esempio. Ricavare la densità discrete delle variabili aleatorie di Bernoulli X e Y, aventi densità congiunta

$$p(x,y) = p^{x+1-y}(1-p)^{1-x+y}, \quad x,y = 0,1 \quad (0 < p < 1).$$

Soluzione. Le densità discrete di X e di Y si ottengono da $p(x,y)$ al seguente modo:

$$p_X(x) = \sum_{y=0}^{1} p(x,y) = \sum_{y=0}^{1} p^{x+1-y}(1-p)^{1-x+y} = p^x(1-p)^{1-x}, \qquad x = 0,1;$$

$$p_Y(y) = \sum_{x=0}^{1} p(x,y) = \sum_{x=0}^{1} p^{x+1-y}(1-p)^{1-x+y} = (1-p)^y p^{1-y}, \qquad y = 0,1.$$

Pertanto X è di Bernoulli di parametro p, e Y è di Bernoulli di parametro $1-p$.

$x \setminus y$	0	1	$p_X(x)$
0	$p(1-p)$	$(1-p)^2$	$1-p$
1	p^2	$p(1-p)$	p
$p_Y(y)$	p	$1-p$	1

Inoltre, per ogni $x,y = 0,1$ si ha

$$p_X(x)\, p_Y(y) = p^x(1-p)^{1-x}\,(1-p)^y p^{1-y} = p^{x+1-y}(1-p)^{1-x+y} = p(x,y).$$

Esempio. Vengono scelte a caso 3 biglie da un'urna contenente 3 biglie rosse, 4 bianche e 5 blu. Se denotiamo con X e Y, rispettivamente, il numero di biglie rosse e bianche scelte, la densità congiunta di X e Y, $p(x,y) = P(X=x, Y=y)$, è data da

$$p(0,0) = \binom{5}{3} \Big/ \binom{12}{3} = \frac{10}{220}, \qquad p(0,1) = \binom{4}{1}\binom{5}{2} \Big/ \binom{12}{3} = \frac{40}{220},$$

$$p(0,2) = \binom{4}{2}\binom{5}{1} \Big/ \binom{12}{3} = \frac{30}{220}, \qquad p(0,3) = \binom{4}{3} \Big/ \binom{12}{3} = \frac{4}{220},$$

$$p(1,0) = \binom{3}{1}\binom{5}{2} \Big/ \binom{12}{3} = \frac{30}{220}, \qquad p(1,1) = \binom{3}{1}\binom{4}{1}\binom{5}{1} \Big/ \binom{12}{3} = \frac{60}{220},$$

$$p(1,2) = \binom{3}{1}\binom{4}{2} \Big/ \binom{12}{3} = \frac{18}{220}, \qquad p(2,0) = \binom{3}{2}\binom{5}{1} \Big/ \binom{12}{3} = \frac{15}{220},$$

$$p(2,1) = \binom{3}{2}\binom{4}{1} \Big/ \binom{12}{3} = \frac{12}{220}, \qquad p(3,0) = \binom{3}{3} \Big/ \binom{12}{3} = \frac{1}{220},$$

essendo $|S| = \binom{12}{3} = \frac{(12)_3}{3!} = \frac{12\cdot 11\cdot 10}{3\cdot 2} = 220.$

Le probabilità $p(x,y)$ possono essere facilmente tabulate, tenendo conto che

$$p(x,y)=P(X=x,Y=y)=\frac{\binom{3}{x}\binom{4}{y}\binom{5}{3-x-y}}{\binom{12}{3}} \qquad (x,y=0,1,2,3;\ x+y\le 3)$$

$x \setminus y$	0	1	2	3	$P(X=x)$
0	10/220	40/220	30/220	4/220	84/220
1	30/220	60/220	18/220	0	108/220
2	15/220	12/220	0	0	27/220
3	1/220	0	0	0	1/220
$P(Y=y)$	56/220	112/220	48/220	4/220	1

Si ha:

$$P(X\ge 2)=\frac{28}{220}\approx 0{,}1273 \qquad P(Y\le 1)=\frac{168}{220}\approx 0{,}7636$$

$$P(X\ge 2, Y\le 1)=\frac{28}{220}\approx 0{,}1273 \qquad P(X=Y)=\sum_{x=y} p(x,y)=\frac{70}{220}\approx 0{,}3182.$$

La probabilità di estrarre z biglie blu, ossia $3-X-Y=z$, è

$$P(3-X-Y=z)=\frac{\binom{5}{z}\binom{7}{3-z}}{\binom{12}{3}} \qquad (z=0,1,2,3).$$

Esempio. Supponiamo che il 15% delle famiglie in una certa comunità non abbia figli, che il 20% ne abbia 1, il 35% ne abbia 2 ed il 30% ne abbia 3. Supponiamo inoltre che in ogni famiglia ogni figlio sia con uguale probabilità maschio o femmina in maniera indipendente. Se si sceglie a caso una famiglia di questa comunità, allora il numero di maschi X e il numero di femmine Y hanno la seguente densità discreta congiunta

$x \setminus y$	0	1	2	3	$p_X(x)$
0	0,15	0,10	0,0875	0,0375	0,3750
1	0,10	0,175	0,1125	0	0,3875
2	0,0875	0,1125	0	0	0,2000
3	0,0375	0	0	0	0,0375
$p_Y(y)$	0,3750	0,3875	0,2000	0,0375	1

Risulta:

$$
\begin{aligned}
p(0,0) &= P(\text{senza figli}) = 0{,}15 \\
p(0,1) &= P(\text{1 figlio femmina}) = P(\text{1 figlio})\, P(\text{1 femmina} \,|\, \text{1 figlio}) = 0{,}2\,\frac{1}{2} = 0{,}1 \\
p(0,2) &= P(\text{2 figli femmina}) = P(\text{2 figli})\, P(\text{2 femmine} \,|\, \text{2 figli}) = 0{,}35\,\left(\frac{1}{2}\right)^2 = 0{,}0875.
\end{aligned}
$$

Diciamo che due variabili aleatorie X e Y sono *congiuntamente (assolutamente) continue* se esiste una funzione $f(x,y)$ integrabile, tale che per ogni sottoinsieme $\mathcal{C}$ dello spazio delle coppie di numeri reali risulti

$$P\{(X,Y) \in \mathcal{C}\} = \iint_{(x,y)\in\mathcal{C}} f(x,y)\,dx\,dy.$$

La funzione $f(x,y)$ è detta densità di probabilità congiunta di X e Y. Se $\mathcal{A}$ e $\mathcal{B}$ sono una qualsiasi coppia di sottoinsiemi di $\mathbb{R}$, allora per $\mathcal{C} = \{(x,y) : x \in \mathcal{A}, y \in \mathcal{B}\}$ otteniamo che

$$P\{X \in \mathcal{A}, Y \in \mathcal{B}\} = \int_{\mathcal{B}}\int_{\mathcal{A}} f(x,y)\,dx\,dy.$$

Poiché

$$F(a,b) = P\{X \in (-\infty,a], Y \in (-\infty,b]\,\} = \int_{-\infty}^{b}\int_{-\infty}^{a} f(x,y)\,dx\,dy,$$

differenziando segue che

$$f(a,b) = \frac{\partial^2}{\partial a\,\partial b}F(a,b) = \frac{\partial}{\partial a}\left[\frac{\partial}{\partial b}F(a,b)\right].$$

Esempio. La distribuzione multinomiale. La distribuzione multinomiale si ottiene quando si ripete n volte un esperimento in condizioni di indipendenza, quando ogni esperimento può avere come risultato uno qualsiasi tra r possibili esiti, con probabilità $p_1, \ldots, p_r$, rispettivamente, tali che $p_1 + \ldots + p_r = 1$. Se denotiamo con X_i il numero degli n esperimenti che hanno dato come risultato l'esito i, allora si ha

$$P\left(X_1 = n_1, X_2 = n_2, \ldots, X_r = n_r\right) = \frac{n!}{n_1! n_2! \cdots n_r!} \, p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}, \quad \text{con} \quad \sum_{i=1}^{r} n_i = n.$$

Tale formula è verificata notando che ogni successione degli esiti degli n esperimenti che portano al fatto che l'esito i si verifichi esattamente n_i volte, per $i = 1, 2, \ldots, r$, avrà probabilità pari a $p_1^{n_1} \, p_2^{n_2} \cdots p_r^{n_r}$, grazie all'ipotesi di indipendenza. La formula finale segue ricordando che $n!/(n_1! n_2! \ldots n_r!)$ è il numero di permutazioni di n oggetti di cui n_1 sono uguali tra loro, n_2 sono uguali tra loro, ... , n_r sono uguali tra loro.

Diciamo che $(X_1, \ldots, X_r)$ ha distribuzione multinomiale con parametri $(n; p_1, \ldots, p_r)$.

Osserviamo che per $r = 2$ la distribuzione multinomiale si riduce a quella binomiale.

Esempio. Nel lanciare 9 volte un dado equilibrato, indicando con X_k, $1 \leq k \leq 6$, il numero di volte che il risultato è k, si ha

$$\begin{aligned}
&P(X_1 = 3, X_2 = 2, X_3 = 2, X_4 = 1, X_5 = 1, X_6 = 0)\\
&\quad = \frac{9!}{3!2!2!1!1!0!}\left(\frac{1}{6}\right)^3\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^0 = \frac{9!}{3!2!2!}\left(\frac{1}{6}\right)^9 \approx 0{,}0015
\end{aligned}$$

$(X_1, \ldots, X_6)$ ha distribuzione multinomiale con parametri $(9; \frac{1}{6}, \ldots, \frac{1}{6})$. Inoltre

$$P(X_1+X_2+X_3 = 7, X_4 = 1, X_5+X_6 = 1) = \frac{9!}{7!1!1!}\left(\frac{1}{2}\right)^7\left(\frac{1}{6}\right)^1\left(\frac{1}{3}\right)^1 = \frac{1}{2^5} = 0{,}3125.$$

Esempio. Siano p, q, $1-p-q$ le probabilità che una squadra di calcio vinca, perda, pareggi una partita. Supponendo di giocare n partite indipendenti, con probabilità costanti, posto $X =$"numero di vincite" e $Y =$"numero di sconfitte", risulta

$$P(X = x, Y = y) = \frac{n!}{x!y!(n-x-y)!}\, p^x q^y (1-p-q)^{n-x-y}$$

(probabilità di x vincite, y sconfitte e $n-x-y$ pareggi).

6.2 Variabili aleatorie indipendenti

Le variabili aleatorie X ed Y si dicono *indipendenti* se, per ogni coppia $\mathcal{A}$ e $\mathcal{B}$ di sottoinsiemi di $\mathbb{R}$, si ha

$$P\left(X \in \mathcal{A}, Y \in \mathcal{B}\right) = P\left(X \in \mathcal{A}\right) P\left(Y \in \mathcal{B}\right).$$

In altre parole, X e Y sono indipendenti se gli eventi $\{X \in \mathcal{A}\}$ e $\{Y \in \mathcal{B}\}$ sono indipendenti per ogni $\mathcal{A}$ e $\mathcal{B}$.

La condizione di indipendenza si può equivalentemente esprimere richiedendo che per ogni coppia di numeri reali x, y risulti

$$P\left(X \leq x, Y \leq y\right) = P\left(X \leq x\right)\, P\left(Y \leq y\right),$$

In termini della distribuzione congiunta $F(x, y)$ di (X, Y), si ha che X e Y sono indipendenti se e solo se

$$F(x, y) = F_X(x)\, F_Y(y), \qquad \forall\, x, y \in \mathbb{R}.$$

Proposizione. Le variabili aleatorie discrete X ed Y sono indipendenti se e solo se la loro densità discreta congiunta può essere espressa come

$$p(x,y) = p_X(x)\,p_Y(y), \qquad \text{per ogni } x,y.$$

Dimostrazione. Se vale $P(X \in \mathcal{A}, Y \in \mathcal{B}) = P(X \in \mathcal{A})\,P(Y \in \mathcal{B})$, allora la tesi segue ponendo $\mathcal{A} = \{x\}$ e $\mathcal{B} = \{y\}$. Viceversa, se vale $p(x,y) = p_X(x)\,p_Y(y)$ per ogni x, y, allora per ogni coppia di sottoinsiemi $\mathcal{A}$ e $\mathcal{B}$ di $\mathbb{R}$ risulta

$$\begin{aligned} P(X \in \mathcal{A}, Y \in \mathcal{B}) &= \sum_{x\in\mathcal{A}}\sum_{y\in\mathcal{B}} p(x,y) = \sum_{x\in\mathcal{A}}\sum_{y\in\mathcal{B}} p_X(x)\,p_Y(y) \\ &= \sum_{x\in\mathcal{A}} p_X(x) \sum_{y\in\mathcal{B}} p_Y(y) = P(X \in \mathcal{A})\; P(Y \in \mathcal{B})\,. \end{aligned}$$

Nel caso di variabili congiuntamente continue la condizione d'indipendenza è equivalente a richiedere che la densità congiunta possa essere fattorizzata come

$$f(x,y) = f_X(x)\,f_Y(y), \qquad \text{per ogni } x,y.$$

Esempio. Nell'esperimento che consiste nel lancio di una moneta ripetuto 3 volte, sia X il numero di volte che esce testa e sia Y il numero di variazioni, ossia quanti lanci danno risultati diversi dal lancio precedente. Stabilire se X e Y sono indipendenti.

Soluzione. Dall'analisi delle corrispondenze tra i risultati dell'esperimento (ognuno di probabilità 1/8) e i valori di X e Y, ricaviamo la densità congiunta di (X, Y):

ω	$X(\omega)$	$Y(\omega)$
ccc	0	0
cct	1	1
ctc	1	2
ctt	2	1
tcc	1	1
tct	2	2
ttc	2	1
ttt	3	0

$x \setminus y$	0	1	2	$p_X(x)$
0	1/8	0	0	1/8
1	0	1/4	1/8	3/8
2	0	1/4	1/8	3/8
3	1/8	0	0	1/8
$p_Y(y)$	1/4	1/2	1/4	1

Risulta ad esempio $p(0,0) = \frac{1}{8} \neq p_X(0)\, p_Y(0) = \frac{1}{8} \cdot \frac{1}{4}$; ne segue che X e Y non sono indipendenti.

Osserviamo che X ha distribuzione binomiale di parametri $(3, \frac{1}{2})$, e Y ha distribuzione binomiale di parametri $(2, \frac{1}{2})$.

Esempio. Nell'esperimento che consiste nel lancio di una moneta truccata (esce testa con probabilità 1/3) ripetuto 3 volte, sia X il numero di volte che esce testa e sia Y il numero di variazioni. Stabilire se X e Y sono indipendenti.

Soluzione. Dall'analisi delle corrispondenze tra i risultati dell'esperimento (ognuno di probabilità indicata in tabella) e i valori di X e Y, ricaviamo la densità congiunta:

ω	$X(\omega)$	$Y(\omega)$	$P(\{\omega\})$
ccc	0	0	8/27
cct	1	1	4/27
ctc	1	2	4/27
ctt	2	1	2/27
tcc	1	1	4/27
tct	2	2	2/27
ttc	2	1	2/27
ttt	3	0	1/27

$x \setminus y$	0	1	2	$p_X(x)$
0	8/27	0	0	8/27
1	0	8/27	4/27	12/27
2	0	4/27	2/27	6/27
3	1/27	0	0	1/27
$p_Y(y)$	1/3	4/9	2/9	1

Risulta $p(0,0) = \frac{8}{27} \neq p_X(0)\, p_Y(0) = \frac{8}{27} \cdot \frac{1}{3}$; ne segue che X e Y non sono indipendenti.

Osserviamo che X ha distribuzione binomiale di parametri $(3, \frac{1}{3})$, mentre Y non ha distribuzione binomiale.

Esempio. Per la densità di probabilità congiunta in tabella, determinare i valori ammissibili di p, e le scelte di p per cui le variabili aleatorie sono indipendenti.

$x \setminus y$	0	1	2	$p_X(x)$
-1	$1/8$	p	$1/8$	$1/4+p$
1	$1/8$	$1/2-p$	$1/8$	$3/4-p$
$p_Y(y)$	$1/4$	$1/2$	$1/4$	1

Soluzione. Imponendo che sia $p(x,y) \geq 0$ e $\sum_x \sum_y p(x,y) = 1$ si ricava che $p \geq 0$ e $1/2 - p \geq 0$, quindi i valori ammissibili sono: $0 \leq p \leq 1/2$.

Per esaminare l'indipendenza di X e Y stabiliamo se $p(-1,0) = p_X(-1)\, p_Y(0)$. Si ha

$$\frac{1}{8} = p(-1,0) = p_X(-1)\, p_Y(0) = \frac{1}{4}\left(\frac{1}{4}+p\right) \quad \Leftrightarrow \quad \frac{1}{2} = \frac{1}{4}+p \quad \Leftrightarrow \quad p = \frac{1}{4}.$$

La soluzione $p = 1/4$ soddisfa la condizione $0 \leq p \leq 1/2$, e quindi ammissibile. Infine, è facile vedere che per $p = 1/4$ risulta $p(x,y) = p_X(x)\, p_Y(y)$ per ogni x, y (ossia X e Y sono indipendenti).

Esercizio. Nelle seguenti tabelle sono riportate densità di probabilità congiunte. Stabilire se le coppie di variabili aleatorie corrispondenti sono indipendenti o meno.

$x \setminus y$	0	1	2	$p_X(x)$
0	2/10	1/10	1/10	
1	1/10	1/10	1/10	
3	2/10	0	1/10	
$p_Y(y)$				

$x \setminus y$	-1	0	1	$p_X(x)$
0	0,20	0,08	0,12	
1	0,15	0,06	0,09	
2	0,15	0,06	0,09	
$p_Y(y)$				

Esempio. Supponiamo che vengano eseguite $n+m$ prove indipendenti, ognuna delle quali abbia probabilità pari a p di risultare in un successo. Se X è il numero di successi nelle prime n prove e Y il numero di successi nelle m prove successive, allora X e Y sono indipendenti, in quanto conoscere il numero dei successi nelle prime n prove non modifica la distribuzione del numero di successi nelle ulteriori m prove, in virtù dell'ipotesi di indipendenza delle prove.

Infatti, per $x = 0, 1, \ldots, n$ e $y = 0, 1, \ldots, m$ si ha

$$P(X = x, Y = y) = P(X = x)\, P(Y = y) = \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{y} p^y (1-p)^{m-y}$$

$$= \binom{n}{x}\binom{m}{y} p^{x+y} (1-p)^{n+m-(x+y)}.$$

Osserviamo che se $Z = X + Y$ è il numero totale di successi nelle $n+m$ prove, allora X e Z non sono indipendenti, e chiaramente anche Y e Z non sono indipendenti.

Esempio. Supponiamo che il numero di richieste che un centro di servizio informatici riceve in un dato giorno sia una variabile di Poisson di parametro λ. Se ogni richiesta è ad alta priorità con probabilità p e a bassa priorità con probabilità $1-p$, si provi che il numero di richieste di servizi ad alta e a bassa priorità sono variabili aleatorie di Poisson indipendenti di parametri λp e $\lambda(1-p)$, rispettivamente.

Soluzione. Denotiamo con X ed Y il numero di richieste di servizi ad alta e a bassa priorità. Condizionando rispetto a $X+Y$, per $i, j \geq 0$ si ottiene

$$\begin{aligned} P(X=i, Y=j) &= P(X=i, Y=j \,|\, X+Y=i+j)\, P(X+Y=i+j) \\ &\quad + P(X=i, Y=j \,|\, X+Y \neq i+j)\, P(X+Y \neq i+j). \end{aligned}$$

Ovviamente si ha $P(X=i, Y=j \,|\, X+Y \neq i+j) = 0$, e quindi

$$P(X=i, Y=j) = P(X=i, Y=j \,|\, X+Y=i+j)\, P(X+Y=i+j).$$

Essendo $X+Y$ il numero totale di richieste di servizi, risulta

$$P(X+Y=i+j) = \mathrm{e}^{-\lambda} \frac{\lambda^{i+j}}{(i+j)!}.$$

Essendo giunte in totale $i+j$ richieste, di cui ognuna è ad alta priorità con probabilità p, la probabilità che i di esse siano ad alta priorità (e quindi j siano a bassa priorità) è la densità discreta di una variabile binomiale di parametri $i+j$ e p valutata in i, ossia

$$P(X=i, Y=j \,|\, X+Y=i+j) = \binom{i+j}{i} p^i (1-p)^j.$$

In conclusione si perviene all'indipendenza di X e Y notando che risulta

$$\begin{aligned} P(X=i, Y=j) &= \binom{i+j}{i} p^i (1-p)^j \, \mathrm{e}^{-\lambda} \frac{\lambda^{i+j}}{(i+j)!} = \mathrm{e}^{-\lambda} \frac{(\lambda p)^i \, [\lambda(1-p)]^j}{i! \, j!} \\ &= \frac{\mathrm{e}^{-\lambda p} (\lambda p)^i}{i!} \, \frac{\mathrm{e}^{-\lambda(1-p)} [\lambda(1-p)]^j}{j!} \qquad (i=0,1,\ldots;\ j=0,1,\ldots) \end{aligned}$$

e

$$P(X=i) = \frac{\mathrm{e}^{-\lambda p}(\lambda p)^i}{i!} \sum_{j=0}^{\infty} \frac{\mathrm{e}^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!} = \frac{\mathrm{e}^{-\lambda p}(\lambda p)^i}{i!} \quad (i=0,1,\ldots)$$

$$P(Y=j) = \frac{\mathrm{e}^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!} \sum_{i=0}^{\infty} \frac{\mathrm{e}^{-\lambda p}(\lambda p)^i}{i!} = \frac{\mathrm{e}^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!} \quad (j=0,1,\ldots).$$

6.3 Somme di variabili aleatorie indipendenti

Molto spesso si è interessati a calcolare la distribuzione di $X+Y$ a partire dalle distribuzioni delle due variabili aleatorie X e Y, sotto l'ipotesi aggiuntiva che le due variabili siano indipendenti. Nel caso in cui le variabili X e Y siano assolutamente continue la variabile aleatoria $X+Y$ è a sua volta continua con densità

$$f_{X+Y}(a) = \int_{-\infty}^{+\infty} f_X(a-y) f_Y(y)\, dy.$$

In particolare si ha il seguente risultato.

Proposizione. Se X_i, $i = 1, 2, \ldots, n$, sono variabili aleatorie indipendenti con distribuzione normale di parametri, rispettivamente, μ_i e σ_i^2, allora $Y = \sum_{i=1}^{n} X_i$ è una variabile aleatoria normale di parametri $\sum_{i=1}^{n} \mu_i$ e $\sum_{i=1}^{n} \sigma_i^2$. Quindi in questo caso risulta

$$E(Y) = \sum_{i=1}^{n} \mu_i = \sum_{i=1}^{n} E(X_i), \qquad \mathrm{Var}(Y) = \sum_{i=1}^{n} \sigma_i^2 = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Esempio. Somme di variabili aleatorie indipendenti di Poisson. Se X e Y sono due variabili aleatorie di Poisson indipendenti di parametri λ_1 e λ_2, rispettivamente, si calcoli la densità discreta di $X+Y$.

Soluzione. Dalle ipotesi fatte segue che per $n = 0, 1, \ldots$

$$
\begin{aligned}
P(X+Y=n) &= \sum_{k=0}^{n} P(X=k, Y=n-k) = \sum_{k=0}^{n} P(X=k)\, P(Y=n-k) \\
&= \sum_{k=0}^{n} \mathrm{e}^{-\lambda_1} \frac{\lambda_1^k}{k!}\, \mathrm{e}^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} = \mathrm{e}^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{\lambda_1^k\, \lambda_2^{n-k}}{k!\,(n-k)!} \\
&= \frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} \lambda_1^k\, \lambda_2^{n-k} = \frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda_1^k\, \lambda_2^{n-k} \\
&= \frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{n!} (\lambda_1+\lambda_2)^n,
\end{aligned}
$$

avendo fatto uso della formula del binomio.

Si ricava che $X+Y$ ha distribuzione di Poisson di parametro $\lambda_1+\lambda_2$.

Esempio. Somme di variabili aleatorie binomiali indipendenti. Se X e Y sono due variabili aleatorie binomiali indipendenti di parametri, rispettivamente, (n, p) e (m, p), mostrare che $X + Y$ è binomiale di parametri $(n + m, p)$.

Soluzione. Dalle ipotesi fatte, per $k = 0, 1, \dots, n + m$ segue

$$
\begin{aligned}
P(X+Y=k) &= \sum_{i=0}^{k} P\left(X=i, Y=k-i\right) = \sum_{i=0}^{k} P\left(X=i\right) P\left(Y=k-i\right) \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} \\
&= p^k (1-p)^{n+m-k} \sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} = \binom{n+m}{k} p^k (1-p)^{n+m-k},
\end{aligned}
$$

avendo utilizzato l'identità combinatoria (formula di Vandermonde)

$$
\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i}.
$$

6.4 Distribuzioni condizionate: il caso discreto

Date due variabili aleatorie discrete X e Y, si definisce la densità discreta condizionata di X dato $Y = y$, come

$$p_{X|Y}(x|y) = P\left(X = x \,|\, Y = y\right) = \frac{P\left(X = x\,, Y = y\right)}{P\left(Y = y\right)} = \frac{p(x, y)}{p_Y(y)},$$

per tutti i valori di y per i quali $p_Y(y) > 0$. In maniera analoga, per ogni y tale che $p_Y(y) > 0$, si definisce la funzione di distribuzione condizionata di X dato $Y = y$:

$$F_{X|Y}(x|y) = P\left(X \leq x \,|\, Y = y\right) = \sum_{a \leq x} p_{X|Y}(a|y).$$

Se X è indipendente da Y, le funzioni condizionate $p_{X|Y}(x|y)$ e $F_{X|Y}(x|y)$ coincidono con le versioni non condizionate. Infatti se X e Y sono indipendenti, per ogni x, y è

$$\begin{aligned} p_{X|Y}(x|y) &= P\left(X = x \,|\, Y = y\right) = \frac{P\left(X = x, Y = y\right)}{P\left(Y = y\right)} \\ &= \frac{P\left(X = x\right)\, P\left(Y = y\right)}{P\left(Y = y\right)} = p_X(x). \end{aligned}$$

Viceversa, se $p_{X|Y}(x|y) = p_X(x)$ per ogni x, y, allora X e Y sono indipendenti.

Esempio. Nell'esperimento che consiste nell'effettuare 5 estrazioni a caso da un'urna contenente 2 biglie nere e 3 biglie bianche, denotiamo con X l'estrazione in cui fuoriesce la prima biglia nera, e con Y il numero di biglie nere fuoriuscite nelle prime 3 estrazioni. Calcolare le densità discrete condizionate di X, dato $Y = y$, per $y = 0, 1, 2$.

Soluzione. Dall'esame delle $\binom{5}{2} = 10$ sequenze di estrazioni possibili segue la densità discreta congiunta di (X, Y), da cui ricava la densità discreta di Y.

ω	$X(\omega)$	$Y(\omega)$
○ ○ ○ ● ●	4	0
○ ○ ● ○ ●	3	1
○ ○ ● ● ○	3	1
○ ● ○ ○ ●	2	1
○ ● ○ ● ○	2	1
○ ● ● ○ ○	2	2
● ○ ○ ○ ●	1	1
● ○ ○ ● ○	1	1
● ○ ● ○ ○	1	2
● ● ○ ○ ○	1	2

$x \setminus y$	0	1	2	$p_X(x)$
1	0	$2/10$	$2/10$	$4/10$
2	0	$2/10$	$1/10$	$3/10$
3	0	$2/10$	0	$2/10$
4	$1/10$	0	0	$1/10$
$p_Y(y)$	$1/10$	$6/10$	$3/10$	1

Notiamo che $p(1, 0) = 0 \neq p_X(1)\, p_Y(0) = 0{,}04$ quindi X e Y non sono indipendenti

Ricaviamo le densità discrete condizionate di X, dato $Y = y$, tenendo presente che

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}, \qquad y = 0, 1, 2.$$

Per $y = 0$:

$$p_{X|Y}(x|0) = \frac{p(x,0)}{p_Y(0)} = \begin{cases} 0, & x = 1,2,3 \\ \dfrac{1/10}{1/10} = 1, & x = 4. \end{cases}$$

Per $y = 1$:

$$p_{X|Y}(x|1) = \frac{p(x,1)}{p_Y(1)} = \begin{cases} \dfrac{2/10}{6/10} = \dfrac{1}{3}, & x = 1,2,3 \\ 0, & x = 4. \end{cases}$$

Per $y = 2$:

$$p_{X|Y}(x|2) = \frac{p(x,2)}{p_Y(2)} = \begin{cases} \dfrac{2/10}{3/10} = \dfrac{2}{3}, & x = 1 \\ \dfrac{1/10}{3/10} = \dfrac{1}{3}, & x = 2 \\ 0, & x = 3,4. \end{cases}$$

Notiamo che le usuali condizioni valide per densità discrete sono soddisfatte dalle densità condizionate $p_{X|Y}(x|y) = \dfrac{p(x,y)}{p_Y(y)}$ e $p_{Y|X}(y|x) = \dfrac{p(x,y)}{p_X(x)}$, essendo

$$p_{X|Y}(x|y) \geq 0, \qquad \sum_x p_{X|Y}(x|y) = \sum_x \frac{p(x,y)}{p_Y(y)} = \frac{1}{p_Y(y)} \sum_x p(x,y) = \frac{p_Y(y)}{p_Y(y)} = 1,$$

$$p_{Y|X}(y|x) \geq 0, \qquad \sum_y p_{Y|X}(y|x) = \sum_y \frac{p(x,y)}{p_X(x)} = \frac{1}{p_X(x)} \sum_y p(x,y) = \frac{p_X(x)}{p_X(x)} = 1.$$

Date due variabili aleatorie discrete X e Y, il valor medio condizionato di X dato $Y = y$, e quello di Y dato $X = x$, sono dati rispettivamente da:

$$E[X|Y = y] = \sum_x x\, p_{X|Y}(x|y), \qquad E[Y|X = x] = \sum_y y\, p_{Y|X}(y|x).$$

Analogamente, le varianze condizionate sono

$$\mathrm{Var}(X|Y = y) = \sum_x (x - E[X|Y = y])^2\, p_{X|Y}(x|y),$$

$$\mathrm{Var}(Y|X = x) = \sum_y (y - E[Y|X = x])^2\, p_{Y|X}(y|x).$$

Anche nel caso condizionato la varianza può esprimersi come differenza tra momento del secondo ordine e quadrato del valore atteso:

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

dove

$$E[X^2|Y=y] = \sum_x x^2\, p_{X|Y}(x|y).$$

Analogamente, si ha

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

con

$$E[Y^2|X=x] = \sum_y y^2\, p_{Y|X}(y|x).$$

Esempio. Nell'esperimento che consiste nel lancio di una moneta ripetuto 3 volte, sia X il numero di volte che esce testa e sia Y il numero di variazioni. Ricavare $E[X|Y=y]$ e $\mathrm{Var}[X|Y=y]$ per $y=0,1,2$.

Soluzione. La densità congiunta e le densità marginali sono riportate in tabella:

$x \setminus y$	0	1	2	$p_X(x)$
0	1/8	0	0	1/8
1	0	1/4	1/8	3/8
2	0	1/4	1/8	3/8
3	1/8	0	0	1/8
$p_Y(y)$	1/4	1/2	1/4	1

Ricordiamo la densità condizionata

$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}.$$

Pertanto: $p_{X|Y}(x|0) = \frac{1}{2}, \quad x = 0, 3.$

Per $y=0$ i momenti condizionati di ordine 1 e 2 di $[X|Y=y]$ sono quindi:

$$E[X|Y=0] = \sum_{x=0}^{3} x\, p_{X|Y}(x|0) = 0\,\frac{1}{2} + 3\,\frac{1}{2} = \frac{3}{2},$$

$$E[X^2|Y=0] = \sum_{x=0}^{3} x^2\, p_{X|Y}(x|0) = 0\,\frac{1}{2} + 9\,\frac{1}{2} = \frac{9}{2}.$$

Segue che la varianza condizionata di X dato $Y = 0$ è

$$\mathrm{Var}(X|Y=0) = E[X^2|Y=0] - (E[X|Y=0])^2 = \frac{9}{2} - \left(\frac{3}{2}\right)^2 = \frac{9}{4}.$$

Analogamente, per $y = 1$ si ha $p_{X|Y}(x|1) = \dfrac{1}{2}, \quad x = 1, 2$ e quindi:

$$E[X|Y=1] = \sum_{x=0}^{3} x\, p_{X|Y}(x|1) = 1\,\frac{1}{2} + 2\,\frac{1}{2} = \frac{3}{2},$$

$$E[X^2|Y=1] = \sum_{x=0}^{3} x^2\, p_{X|Y}(x|1) = 1\,\frac{1}{2} + 4\,\frac{1}{2} = \frac{5}{2},$$

pertanto la varianza condizionata di X dato $Y = 1$ è

$$\mathrm{Var}(X|Y=1) = E[X^2|Y=1] - (E[X|Y=1])^2 = \frac{5}{2} - \left(\frac{3}{2}\right)^2 = \frac{1}{4}.$$

Per $y = 2$ i risultati coincidono con quelli del caso $y = 1$, essendo

$$p_{X|Y}(x|1) = p_{X|Y}(x|2) \quad \text{per ogni } x = 0, 1, 2, 3.$$

Esempio. Supponiamo che X e Y siano due variabili aleatorie di Poisson indipendenti di parametri λ_1 e λ_2, rispettivamente. Si dimostri che la densità discreta condizionata di X dato $X+Y=n$ è binomiale di parametri n e $p=\lambda_1/(\lambda_1+\lambda_2)$.

Soluzione. Dalla proprietà di indipendenza, per $k=0,1,\ldots,n$ si ha

$$\begin{aligned}
P(X=k\,|\,X+Y=n) &= \frac{P(X=k,X+Y=n)}{P(X+Y=n)} = \frac{P(X=k,Y=n-k)}{P(X+Y=n)} \\
&= \frac{P(X=k)\,P(Y=n-k)}{P(X+Y=n)}.
\end{aligned}$$

Ricordando che $X+Y$ ha distribuzione di Poisson di parametro $\lambda_1+\lambda_2$, e inoltre che $p=\lambda_1/(\lambda_1+\lambda_2)$, la precedente formula diventa, per $k=0,1,\ldots,n$

$$\begin{aligned}
P(X=k\,|\,X+Y=n) &= \frac{\mathrm{e}^{-\lambda_1}\,\lambda_1^k}{k!}\,\frac{\mathrm{e}^{-\lambda_2}\,\lambda_2^{n-k}}{(n-k)!}\,\frac{n!}{\mathrm{e}^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n} \\
&= \frac{n!}{k!(n-k)!}\frac{\lambda_1^k\,\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n} = \binom{n}{k}\frac{\lambda_1^k\,\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^{k+n-k}} \\
&= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k} = \binom{n}{k}p^k(1-p)^{n-k}.
\end{aligned}$$

Esempio. Da un'urna contenente k biglie nere e $n-k$ biglie bianche si effettuano n estrazioni a caso con reinserimento. Denotiamo con Y il numero di biglie nere estratte, e con X l'estrazione in cui si estrae una biglia nera per la prima volta. Determinare $p_Y(r)$, $p_{X|Y}(j|r)$ e $p_X(j)$.

Soluzione. Poiché le estrazioni sono con reinserimento, Y ha distribuzione binomiale di parametri n e $p = k/n$; quindi risulta

$$p_Y(r) = P(Y = r) = \binom{n}{r} p^r \, (1-p)^{n-r}, \qquad r = 0, 1, \ldots, n \qquad \left(p = \frac{k}{n}\right).$$

Notiamo che, se $r \geq 1$, $p_{X|Y}(j|r)$ è la probabilità che una sequenza casuale costituita da r biglie nere e $n-r$ biglie bianche abbia la prima biglia nera al j-esimo posto, e pertanto si ha

$$p_{X|Y}(j|r) = P(X = j|Y = r) = \binom{n-j}{r-1} \Big/ \binom{n}{r}, \qquad j = 1, 2, \ldots, n-r+1.$$

Quindi, dovendo essere $r \geq 1$ e $j \leq n - r + 1$, risulta

$$
\begin{aligned}
p_X(j) &= P(X = j) = \sum_{r=0}^{n} P(X = j|Y = r)\, P(Y = r) \\
&= \sum_{r=1}^{n-j+1} \frac{\binom{n-j}{r-1}}{\binom{n}{r}} \binom{n}{r} p^r\, (1-p)^{n-r} = \sum_{r=1}^{n-j+1} \binom{n-j}{r-1} p^r\, (1-p)^{n-r}.
\end{aligned}
$$

Ponendo $s = r - 1$ si ottiene, per $j = 1, 2, \ldots, n - r + 1$

$$
\begin{aligned}
p_X(j) &= \sum_{s=0}^{n-j} \binom{n-j}{s} p^{s+1}\, (1-p)^{n-s-1} \\
&= \sum_{s=0}^{n-j} \binom{n-j}{s} p^{s}\, (1-p)^{n-j-s} \cdot p\, (1-p)^{j-1} \\
&= p\, (1-p)^{j-1} \qquad \text{(per la formula del binomio)}
\end{aligned}
$$

che è una probabilità di tipo geometrico.

Esercizi per casa

6.1) Sono date due urne, la prima contenente 2 biglie nere e 3 biglie bianche, la seconda contenente 1 biglia nera e 4 biglie bianche. Si estraggono a caso (senza reinserimento) 2 biglie da ciascuna delle due urne. Sia X il numero di biglie bianche estratte dalla prima urna, e Y il numero di biglie bianche estratte dalla seconda urna. Determinare
(i) $P(X = x)$ per $x = 0, 1, 2$ e $P(Y = y)$ per $y = 0, 1, 2$;
(ii) $P(Z = z)$ per $z = 0, 1, 2, 3, 4$ (avendo posto $Z = X + Y$);
(iii) $E(X)$, $E(Y)$, $E(Z)$, e verificare che $E(Z) = E(X) + E(Y)$.

6.2) Da un'urna contenente 2 biglie nere e 3 biglie bianche si effettuano 3 estrazioni a caso senza reinserimento. Denotiamo con Y il numero di biglie bianche estratte, e con X l'estrazione in cui si estrae una biglia bianca per la prima volta. Determinare
(i) $p_Y(r)$ per $r = 1, 2, 3$;
(ii) $p_{X|Y}(j|1)$ per $j = 1, 2, 3$;
(iii) $p_X(j)$ per $j = 1, 2, 3$.

6.3) Sia X una variabile aleatoria di Bernoulli di parametro $p = 3/4$. Se $X = 0$ allora Y è uniforme in $(0, 1)$, mentre se $X = 1$ allora Y è uniforme in $(0, 2)$. Calcolare
(i) $P(Y \leq 1/2)$,
(ii) $P(X = 0 \,|\, Y \leq 1/2)$ e $P(X = 1 \,|\, Y \leq 1/2)$.

Calcolo delle Probabilità e Statistica Matematica – A.A. 2009/10

CAPITOLO 7 – Proprietà del valore atteso

7.1 Introduzione

Ricordiamo che il valore atteso di una variabile aleatoria discreta X con densità $p(x)$ è

$$E[X] = \sum_x x\, p(x)$$

mentre se X è (assolutamente) continua con densità $f(x)$ si ha

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

Proposizione. Se X assume valori compresi tra a e b, allora il valore atteso è compreso tra a e b. In altre parole, se $P(a \le X \le b) = 1$ allora $a \le E[X] \le b$.

Dimostrazione. Poiché $p(x) = 0$ se x non appartiene ad $[a, b]$, nel caso discreto si ha

$$E[X] = \sum_{x:p(x)>0} x\, p(x) = \sum_{a\le x\le b} x\, p(x) \ge \sum_{a\le x\le b} a\, p(x) = a \sum_{a\le x\le b} p(x) = a.$$

Analogamente si ricava $E[X] \le b$. La dimostrazione nel caso continuo è analoga.

7.2 Valore atteso di somme di variabili aleatorie

Siano X e Y due variabili aleatorie, e $g\colon \mathbb{R}^2 \to \mathbb{R}$ una funzione di due variabili.

Proposizione. Se X e Y hanno densità discreta congiunta $p(x, y)$, allora

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) p(x, y).$$

Illustriamo un'importante applicazione di quanto su esposto. Se $E[X]$ e $E[Y]$ sono entrambe finite e poniamo $g(x, y) = x + y$, si ha

$$\begin{aligned} E[X+Y] &= \sum_x \sum_y (x+y)\, p(x, y) = \sum_x x \sum_y p(x, y) + \sum_y y \sum_x p(x, y) \\ &= \sum_x x\, p_X(x) + \sum_y y\, p_Y(y) = E[X] + E[Y]. \end{aligned}$$

Ragionando per induzione si dimostra che se $E[X_i]$ è finito per $i = 1, \ldots, n$, allora sussiste la proprietà di linearità per le variabili aleatorie $X_1, \ldots, X_n$:

$$E[a_1 X_1 + \cdots + a_n X_n + b] = a_1 E[X_1] + \cdots + a_n E[X_n] + b.$$

Esempio. La media campionaria. Siano $X_1, \ldots, X_n$ variabili aleatorie indipendenti e identicamente distribuite con distribuzione F e valore atteso μ. Tale sequenza costituisce un *campione casuale* della distribuzione F. La variabile aleatoria

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

è detta media campionaria di $X_1, \ldots, X_n$. Calcolare $E[\overline{X}]$.

Soluzione. Dato che $E[X_i] = \mu$, per la proprietà di linearità si ha

$$E[\overline{X}] = E\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu = \mu.$$

Pertanto il valore atteso della media campionaria è μ, ossia la media stessa della distribuzione. La media campionaria è spesso utilizzata per dare una stima del valore atteso μ della distribuzione, se questo è sconosciuto.

Esempio. Siano X e Y due variabili aleatorie tali che $X \geq Y$. In altre parole, per ogni esito dell'esperimento, il valore assunto da X è maggiore o uguale del valore assunto da Y. Dato che $X - Y \geq 0$, si ha che $E[X - Y] \geq 0$, ossia $E[X] \geq E[Y]$.

Esempio. La disuguaglianza di Boole. Siano $A_1, \ldots, A_n$ degli eventi, e definiamo le loro variabili indicatrici

$$X_i = \begin{cases} 1 & \text{se } A_i \text{ si realizza,} \\ 0 & \text{altrimenti,} \end{cases} \qquad\qquad (i = 1, \ldots, n).$$

Notiamo che X_i è una variabile di Bernoulli, con

$$P(X_i = 1) = P(A_i), \qquad P(X_i = 0) = P(\overline{A_i}), \qquad E[X_i] = P(X_i = 1) = P(A_i).$$

Poniamo $X = \displaystyle\sum_{i=1}^{n} X_i$, così X è il numero di eventi A_i che si realizzano. Infine sia

$$Y = \begin{cases} 1 & \text{se } X \geq 1, \\ 0 & \text{altrimenti,} \end{cases}$$

sicché Y è uguale ad 1 se si realizza almeno uno degli eventi A_i e vale 0 altrimenti. Notiamo che Y è una variabile di Bernoulli tale che $Y = \max\{X_1, \ldots, X_n\}$.

Risulta, per la proprietà di linearità,

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} P(X_i = 1) = \sum_{i=1}^{n} P(A_i)$$

e

$$E[Y] = P(Y = 1) = P\{\text{si realizza almeno uno degli eventi } A_i\} = P\left(\bigcup_{i=1}^{n} A_i\right).$$

Essendo chiaramente

$$Y \leq X,$$

dall'esempio precedente si deduce che

$$E[Y] \leq E[X].$$

Si ottiene quindi la seguente relazione, nota come *disuguaglianza di Boole*:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

Esempio. Valore atteso di una variabile aleatoria binomiale. Sia X una variabile aleatoria binomiale di parametri n e p. Ricordando che questa variabile rappresenta il numero di successi in n prove indipendenti, dove ogni prova ha successo con probabilità p, si ha che

$$X = X_1 + X_2 + \cdots + X_n,$$

dove

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esima prova è un successo,} \\ 0 & \text{se l'}i\text{-esima prova è un insuccesso.} \end{cases}$$

La variabile aleatoria di Bernoulli X_i ha valore atteso $E[X_i] = 0 \cdot (1-p) + 1 \cdot p = p$. Di conseguenza, facendo uso della proprietà di linearità, risulta

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = E[X_1] + E[X_2] + \cdots + E[X_n] = n\,p.$$

Esempio. Valore atteso di una variabile aleatoria ipergeometrica. Si scelgano a caso n biglie da un'urna contenente N biglie delle quali m sono bianche. Determinare il numero atteso di biglie bianche selezionate.

Soluzione. Il numero X di biglie bianche selezionate si può rappresentare come

$$X = X_1 + X_2 + \cdots + X_m,$$

dove

$$X_i = \begin{cases} 1 & \text{se è stata scelta l'}i\text{-esima biglia bianca,} \\ 0 & \text{altrimenti.} \end{cases}$$

Essendo

$$E[X_i] = P(X_i = 1) = P\{\text{è stata scelta l'}i\text{-esima biglia bianca }\} = \frac{\binom{1}{1}\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N},$$

si ottiene infine, per la proprietà di linearità,

$$E[X] = E[X_1] + \cdots + E[X_m] = \frac{m\,n}{N}.$$

Ricordiamo che $\displaystyle\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N(N-1)!}{n(n-1)!(N-n)!} = \frac{N}{n}\binom{N-1}{n-1}$.

Tale risultato può essere ottenuto anche utilizzando una rappresentazione alternativa:

$$X = Y_1 + \cdots + Y_n,$$

dove le variabili di Bernoulli

$$Y_i = \begin{cases} 1 & \text{se l'}i\text{-esima biglia selezionata è bianca,} \\ 0 & \text{altrimenti,} \end{cases} \qquad (i = 1, 2, \ldots, n).$$

non sono indipendenti, ma sono identicamente distribuite. Notiamo infatti che

$$P(Y_i = 1 \,|\, X = k) = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\dfrac{(n-1)!}{(k-1)!\,(n-k)!}}{\dfrac{n!}{k!\,(n-k)!}} = \frac{k}{n},$$

dove a denominatore c'è il numero $\binom{n}{k}$ di sequenze booleane di lunghezza n con k elementi pari a 1 e con $n-k$ elementi pari a 0, mentre a numeratore è presente il numero $\binom{n-1}{k-1}$ di sequenze booleane di lunghezza n aventi 1 nel posto i-esimo e $k-1$ elementi pari a 1 nei rimanenti $n-1$ posti.

Notiamo che la probabilità

$$P(Y_i = 1 \,|\, X = k) = \frac{k}{n}$$

si può ottenere anche in modo diretto, tenendo presente che se si devono collocare k cifre pari a 1 e $n - k$ cifre pari a 0 in una sequenza di lunghezza n, ci sono k casi favorevoli su un totale di n affinché nel posto i-esimo della sequenza sia collocata una cifra pari a 1. Ne segue che la distribuzione di Y_i non dipende da i, essendo

$$P(Y_i = 1) = \sum_{k=0}^{n} P(Y_i = 1 \,|\, X = k)\, P(X = k) = \sum_{k=0}^{n} \frac{k}{n}\, P(X = k) = \frac{1}{n}\, E(X).$$

Si ricava pertanto

$$E(X) = n\, P(Y_i = 1) = n\, P(Y_1 = 1) = n\, \frac{m}{N}.$$

Esempio. Valore atteso del numero di accoppiamenti. N persone lanciano il loro cappello nel centro di una stanza. I cappelli vengono mescolati, e ogni persona ne prende uno a caso. Determinare il valore atteso del numero di persone che selezionano il proprio cappello.

Soluzione. Indicando con X il numero di accoppiamenti, possiamo calcolare $E[X]$ scrivendo $X = X_1 + X_2 + \cdots + X_N$, dove

$$X_i = \begin{cases} 1 & \text{se la persona } i \text{ prende il suo cappello,} \\ 0 & \text{altrimenti.} \end{cases}$$

Dato che, per ogni i, l'i-esima persona può scegliere ugualmente uno degli N cappelli, si ha

$$E[X_i] = P(X_i = 1) = \frac{1}{N},$$

da cui

$$E[X] = E[X_1] + \cdots + E[X_N] = \frac{1}{N} \cdot N = 1.$$

Quindi, in media, esattamente una persona seleziona il proprio cappello.

Esempio. Al passaggio di uno stormo di n anatre, vi sono n cacciatori che sparano all'istante, e ognuno mira ad un bersaglio a caso, indipendentemente dagli altri. Se ogni cacciatore colpisce il suo bersaglio con probabilità p, calcolare il numero atteso di anatre che non sono colpite.

Soluzione. Sia $X_i = 1$ se l'i-esima anatra non è colpita, e 0 altrimenti, per $i = 1, 2, \ldots, n$, così $X = X_1 + \ldots + X_n$ dà il numero totale di anatre che non sono colpite. Quindi il numero atteso richiesto, per la proprietà di linearità, è dato da

$$E[X] = E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n].$$

Ogni cacciatore colpirà l'i-esima anatra con probabilità $\frac{1}{n} \cdot p$, ($\frac{1}{n}$ per scegliere l'i-esima anatra, e p per il successo nel colpire), sicché per l'indipendenza si ha

$$E[X_i] = P\{X_i = 1\} = \left(1 - \frac{p}{n}\right)^n \qquad \text{e quindi} \qquad E[X] = n\left(1 - \frac{p}{n}\right)^n.$$

Nota. La frazione attesa di anatre che non sono colpite è

$$E\left[\frac{X}{n}\right] = \frac{E[X]}{n} = \left(1 - \frac{p}{n}\right)^n \to e^{-p} \quad \text{per} \quad n \to \infty.$$

Esempio. Algoritmo quick-sort. Supponiamo di disporre di un insieme di n valori distinti $x_1, x_2, \ldots, x_n$ e di volerli ordinare in modo crescente (sort). Una procedura efficace al riguardo è l'algoritmo di quick-sort, definito come segue. Se $n = 2$ l'algoritmo confronta due valori e li mette in ordine appropriato. Se $n > 2$, uno degli elementi è scelto a caso, ad esempio x_i, e si confrontano gli altri valori con x_i. I valori inferiori a x_i vengono messi tra parentesi graffe alla sinistra di x_i, mentre quelli superiori a x_i vengono messi alla destra di x_i. L'algoritmo si ripete per i valori interni alle parentesi graffe e continua finché i valori non sono tutti ordinati.
Supponiamo ad esempio di voler ordinare i seguenti 10 numeri distinti

$$5, 9, 3, 10, 11, 14, 8, 4, 17, 6.$$

Iniziamo scegliendo un numero a caso, ad esempio 10 (ogni numero può essere scelto con probabilità 1/10). Confrontando gli altri valori con tale numero si ottiene

$$\{5, 9, 3, 8, 4, 6\}, 10, \{11, 14, 17\}$$

Rivolgiamo ora l'attenzione al sottoinsieme $\{5, 9, 3, 8, 4, 6\}$ e scegliamo a caso uno dei suoi valori, ad esempio il 6. Confrontando ognuno dei valori tra parentesi con 6 e mettendo a sinistra di esso i valori più piccoli di 6 e a destra quelli più grandi di 6 si ottiene

$$\{5, 3, 4\}, 6, \{9, 8\}, 10, \{11, 14, 17\}.$$

Considerando ora la parentesi a sinistra e scegliendo 4 per i successivi confronti, si giunge a

$$\{3\}, 4, \{5\}, 6, \{9, 8\}, 10, \{11, 14, 17\}.$$

Si continua finché non ci sono più sottoinsiemi che contengono più di un elemento.

Sia X il numero di confronti che servono all'algoritmo di quick-sort per ordinare n numeri distinti. Si ha allora che $E[X]$ è una misura dell'efficienza dell'algoritmo.
Per calcolare $E[X]$, esprimiamo X come somma di altre variabili aleatorie

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} I(i, j),$$

dove per $1 \leq i < j \leq n$, $I(i,j)$ è uguale ad 1 se i e j sono prima o poi confrontati direttamente mentre è uguale a 0 altrimenti. Quindi si ha

$$
\begin{aligned}
E[X] = E[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} I(i,j)] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[I(i,j)] \\
&= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P\{i \,\mathrm{e}\, j \,\text{sono confrontati tra loro}\}.
\end{aligned}
$$

Per determinare la probabilità che i e j non siano mai confrontati, osserviamo che i valori $\{i, i+1, \ldots, j-1, j\}$ sono inizialmente nella stessa parentesi e vi rimarranno se il numero scelto per il primo confronto non è compreso tra i e j. Ad esempio se il numero da confrontare con gli altri è strettamente maggiore di j, allora tutti i valori $\{i, i+1, \ldots, j-1, j\}$ andranno in una parentesi a sinistra del numero scelto, mentre se esso è strettamente inferiore a i, allora tali valori andranno messi in una parentesi a destra.

Tutti i valori $\{i, i+1, \ldots, j-1, j\}$ rimarranno quindi all'interno della stessa parentesi finché uno di essi è scelto per effettuare i confronti. Il valore scelto viene poi confrontato con gli altri compresi tra i e j. Se non è né i, né j, allora, dopo il confronto, i andrà in una parentesi alla sua sinistra e j in una parentesi alla sua destra. D'altro lato, se il valore scelto nell'insieme $\{i, i+1, \ldots, j-1, j\}$ è uguale a i o j, ci sarà un confronto diretto tra i e j.

Supponendo allora che il valore scelto sia uno dei $j-i+1$ valori tra i e j, la probabilità che si tratti di i o di j è pari a $2/(j-i+1)$. Si può quindi concludere che

$$P\{i \,\mathrm{e}\, j \,\text{sono confrontati tra loro}\} = \frac{2}{j-i+1},$$

da cui si ricava

$$E[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1}.$$

Per ottenere l'ordine di grandezza di $E[X]$ per n grande, approssimiamo le somme con degli integrali. Ora

$$\begin{aligned}\sum_{j=i+1}^{n} \frac{2}{j-i+1} &\approx \int_{i+1}^{n} \frac{2}{x-i+1}\mathrm{d}x = 2\log(x-i+1)\Big|_{i+1}^{n} \\ &= 2\log(n-i+1) - 2\log(2) \approx 2\log(n-i+1).\end{aligned}$$

Quindi

$$\begin{aligned}E[x] &\approx \sum_{i=1}^{n-1} 2\log(n-i+1) \approx 2\int_{1}^{n-1} \log(n-x+1)\mathrm{d}x = 2\int_{2}^{n} \log(y)\mathrm{d}y \\ &= 2(y\log(y) - y)\Big|_{2}^{n} \approx 2n\log(n).\end{aligned}$$

Pertanto, per n grande, l'algoritmo di quick-sort richiede, in media, approssimativamente $2n\,log(n)$ confronti per ordinare n valori distinti.

7.3 Covarianza, Varianza di una somma e correlazioni

Proposizione. Se X e Y sono indipendenti, e se g e h sono due funzioni, allora

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

Dimostrazione. Supponiamo che X ed Y siano variabili discrete con distribuzione congiunta $p(x, y)$. Allora risulta

$$\begin{aligned} E[g(X)h(Y)] &= \sum_x \sum_y g(x)h(y)p(x, y) = \sum_x \sum_y g(x)h(y)p_X(x)p_Y(y) \\ &= \sum_x g(x)p_X(x) \sum_y h(y)p_Y(y) = E[g(X)]E[h(Y)], \end{aligned}$$

avendo sfruttato la relazione d'indipendenza $p(x, y) = p_X(x)p_Y(y)$, valida $\forall x, y$.

Come il valore atteso e la varianza di una singola variabile forniscono delle informazioni sulla variabile aleatoria, così la covarianza tra due variabili aleatorie fornisce un'informazione sulla relazione tra le due variabili.

Definizione. La covarianza tra X ed Y, indicata con $Cov(X,Y)$, è definita da

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

Sviluppando i termini e usando la proprietà di linearità, si ricava

$$\begin{aligned} Cov(X,Y) &= E[XY - E[X]Y - XE[Y] + E[X]E[Y]] \\ &= E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y]. \end{aligned}$$

– Se $Cov(X,Y) > 0$ le variabili aleatorie X ed Y si dicono *positivamente correlate*; in tal caso entrambe tendono ad assumere mediamente valore maggiore, o minore, della propria media.

– Al contrario, se $Cov(X,Y) < 0$ le variabili aleatorie X ed Y si dicono *negativamente correlate*; in tal caso se una tende ad assumere mediamente valore maggiore della propria media, l'altra tende ad assumere valore minore della propria media.

– Se $Cov(X,Y) = 0$ le variabili aleatorie X ed Y si dicono *scorrelate*, ovvero *non correlate*, ed in tal caso sono debolmente legate tra loro.

La seguente proposizione mostra la relazione tra variabili indipendenti e scorrelate.

Proposizione. Se X e Y sono variabili aleatorie indipendenti allora $Cov(X,Y) = 0$.
Dimostrazione. Usando $g(x) = x$ e $h(y) = y$ nella Proposizione precedente, si ha $E[XY] = E[X]\,E[Y]$ e quindi $Cov(X,Y) = E[XY] - E[X]\,E[Y] = 0$.

Il prossimo esempio mostra che il viceversa in generale è falso. Ossia, se X e Y sono variabili aleatorie scorrelate, ciò non implica che X e Y siano indipendenti.

Esempio. Consideriamo un gioco in cui si lancia un dado; valgono le seguenti regole: se esce 1 il primo giocatore perde 3 euro, se esce 2 il secondo giocatore perde 3 euro, se esce 3 il primo giocatore vince 1 euro, se esce 4 il secondo giocatore vince 1 euro, se esce 5 il primo giocatore vince 2 euro, se esce 6 il secondo giocatore vince 2 euro. Indicando con X e Y la vincita del primo e del secondo giocatore, rispettivamente, stabilire se tali variabili aleatorie sono indipendenti e se sono scorrelate.
Soluzione. Dalle regole esposte si ha la seguente tabella per i valori di X e Y:

ω	1	2	3	4	5	6
$X(\omega)$	-3	0	1	0	2	0
$Y(\omega)$	0	-3	0	1	0	2

Si ricavano così la densità congiunta e le densità discrete marginali di X e Y:

$x \setminus y$	-3	0	1	2	$p_X(x)$
-3	0	$1/6$	0	0	$1/6$
0	$1/6$	0	$1/6$	$1/6$	$1/2$
1	0	$1/6$	0	0	$1/6$
2	0	$1/6$	0	0	$1/6$
$p_Y(y)$	$1/6$	$1/2$	$1/6$	$1/6$	1

Si vede facilmente che X e Y non sono indipendenti, essendo $p(0,0) = 0 \neq p_X(0)\,p_Y(0) = \frac{1}{2} \cdot \frac{1}{2}$

Inoltre, si ricava che X e Y sono identicamente distribuite, essendo

$$p_X(t) = p_Y(t) \quad \text{per ogni} \;\; t = -3, 0, 1, 2.$$

Si ha quindi

$$E[X] = E[Y] = \sum_{x=-3}^{2} x\, p_X(x) = -3 \cdot \frac{1}{6} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} = 0,$$

$$E[XY] = \sum_{x=-3}^{2} \sum_{y=-3}^{2} x\, y\, p(x,y) = (-3) \cdot 0 \cdot \frac{1}{6} + 0 \cdot (-3) \cdot \frac{1}{6} + \ldots + 0 \cdot 2 \cdot \frac{1}{6} = 0,$$

pertanto

$$Cov(X,Y) = E[XY] - E[X]\,E[Y] = 0.$$

Le variabili aleatorie X e Y si dicono *scambiabili* se

$$p(x,y) = p(y,x) \quad \text{per ogni } x, y.$$

Si vede facilmente che se X e Y sono scambiabili, allora X e Y sono identicamente distribuite. Infatti, sommando su tutti i valori di x, dalla relazione $p(x,y) = p(y,x)$ si ha

$$p_Y(y) = \sum_x p(x,y) = \sum_x p(y,x) = p_X(y) \quad \text{per ogni } y,$$

cosicché X e Y sono identicamente distribuite.

Osservazione. Notiamo che se $Cov(X,Y) \neq 0$ allora X e Y non sono indipendenti.

Esempio. Nel lancio di una moneta ripetuto 3 volte, sia X il numero di volte che esce testa e sia Y il numero di variazioni, ossia quanti lanci danno risultati diversi dal lancio precedente. Calcolare $Cov(X,Y)$.

Soluzione. Abbiamo già ricavato la densità congiunta di (X,Y), da cui si trae che X e Y non sono indipendenti; inoltre si ricava:

$x\backslash y$	0	1	2	$p_X(x)$
0	1/8	0	0	1/8
1	0	1/4	1/8	3/8
2	0	1/4	1/8	3/8
3	1/8	0	0	1/8
$p_Y(y)$	1/4	1/2	1/4	1

$$E(X) = 0\cdot\frac{1}{8} + 1\cdot\frac{3}{8} + 2\cdot\frac{3}{8} + 3\cdot\frac{1}{8} = \frac{3}{2}$$

$$E(Y) = 0\cdot\frac{1}{4} + 1\cdot\frac{1}{2} + 2\cdot\frac{1}{4} = 1$$

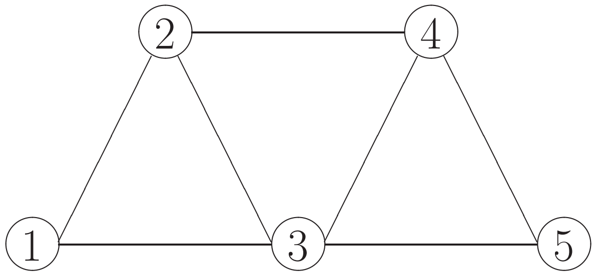
$$E(XY) = 1\cdot 1\cdot\frac{1}{4} + 1\cdot 2\frac{1}{8} + 2\cdot 1\cdot\frac{1}{4} + 2\cdot 2\cdot\frac{1}{8} = \frac{3}{2}$$

e pertanto: $Cov(X,Y) = E(XY) - E(X)\,E(Y) = \dfrac{3}{2} - \dfrac{3}{2}\cdot 1 = 0.$

Notiamo anche che

$$X \sim Bin\left(3,\frac{1}{2}\right) \;\Rightarrow\; E[X] = \frac{3}{2} \qquad \text{e} \qquad Y \sim Bin\left(2,\frac{1}{2}\right) \;\Rightarrow\; E[Y] = \frac{2}{2} = 1.$$

Esempio. Consideriamo l'esperimento che consiste nello scegliere a caso 2 dei 7 archi del grafo in figura. Sia X il numero di archi scelti collegati al nodo 3, e sia Y il numero di archi scelti collegati al nodo 1. Determinare $Cov(X,Y)$, $P(X=Y)$, $P(X+Y=2)$.



Soluzione. Dall'esame dalla struttura del grafo, essendo $\binom{7}{2} = \frac{(7)_2}{2!} = \frac{7\cdot 6}{2} = 21$ si ha

$x \setminus y$	0	1	2	$p_X(x)$
0	1/21	2/21	0	3/21
1	6/21	5/21	1/21	12/21
2	3/21	3/21	0	6/21
$p_Y(y)$	10/21	10/21	1/21	1

$$E[X] = 0\,\frac{3}{21} + 1\,\frac{12}{21} + 2\,\frac{6}{21} = \frac{24}{21} = \frac{8}{7}$$

$$E[Y] = 0\,\frac{10}{21} + 1\,\frac{10}{21} + 2\,\frac{1}{21} = \frac{12}{21} = \frac{4}{7}$$

$$E[X\,Y] = 1\cdot 1\,\frac{5}{21} + 1\cdot 2\,\frac{1}{21} + 2\cdot 1\,\frac{3}{21} = \frac{13}{21}, \quad Cov(X,Y) = \frac{13}{21} - \frac{8}{7}\cdot\frac{4}{7} = -\frac{5}{147} = -0{,}034$$

$$P(X=Y) = \sum_{x=y} p(x,y) = \frac{6}{21} < P(X+Y=2) = \sum_{x+y=2} p(x,y) = \frac{8}{21}.$$

Proposizione. La covarianza possiede le seguenti proprietà:

(i) $Cov(X, Y) = Cov(Y, X)$,

(ii) $Cov(X, X) = \mathrm{Var}(X)$,

(iii) $Cov(aX, Y) = a\, Cov(X, Y)$, per ogni costante reale a,

(iv) $Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$.

Dimostrazione. La (i) e la (ii) seguono immediatamente dalla definizione covarianza, $Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$. Per la proprietà di linearità del valore medio, $E(aX) = a\, E(X)$, si ha la (iii):

$$Cov(aX, Y) = E[(aX - E(aX))(Y - E(Y))] = a\, Cov(X, Y).$$

Per dimostrare la (iv), ossia la proprietà di additività della covarianza, poniamo $\mu_i = E[X_i]$ e $v_j = E[Y_j]$. Allora

$$E\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} \mu_i, \qquad E\Big[\sum_{j=1}^{m} Y_j\Big] = \sum_{j=1}^{m} v_j.$$

Inoltre

$$
\begin{aligned}
Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) &= E\left[\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i\right)\left(\sum_{j=1}^{m} Y_j - \sum_{j=1}^{m} v_j\right)\right] \\
&= E\left[\sum_{i=1}^{n}(X_i - \mu_i)\sum_{j=1}^{m}(Y_j - v_j)\right] \\
&= E\left[\sum_{i=1}^{n}\sum_{j=1}^{m}(X_i - \mu_i)(Y_j - v_j)\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} E[(X_i - \mu_i)(Y_j - v_j)],
\end{aligned}
$$

dove l'ultima uguaglianza segue dalla proprietà di linearità (in quanto il valore atteso di una somma di variabili aleatorie è uguale alla somma dei valori attesi).

Esercizio. Mostrare la proprietà di linearità $Cov(aX + b, Y) = a\,Cov(X, Y)$, valida per ogni scelta delle costanti reali a, b.

Proposizione.

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} Cov(X_i, X_j).$$

Dimostrazione. Dalla precedente proposizione, posto $Y_j = X_j$, $j = 1, \ldots, n$, si ha

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) &= Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) \\
&= \sum_{i=j} Cov(X_i, X_j) + \sum_{i\neq j} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i\neq j} Cov(X_i, X_j).
\end{aligned}$$

La tesi segue immediatamente ricordando che $Cov(X_i, X_j) = Cov(X_j, X_i)$, e quindi

$$\sum_{i\neq j} Cov(X_i, X_j) = \sum_{i<j} Cov(X_i, X_j) = \sum_{i>j} Cov(X_i, X_j) = 2\sum_{i<j} Cov(X_i, X_j).$$

(La *matrice di covarianza* è una matrice quadrata, $n \times n$, avente $Cov(X_i, X_j)$ come elemento (i, j); quindi è simmetrica, con le varianze sulla diagonale principale).

Ovviamente, se $X_1, \ldots, X_n$ sono a due a due scorrelate, cioè se $Cov(X_i, X_j) = 0$ per ogni $i \neq j$, risulta:

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}(X_i).$$

Notiamo che dalla proposizione precedente si ha, per $a_1, \ldots, a_n$ costanti arbitrarie,

$$\text{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \text{Var}(X_i) + 2 \sum_{i<j} a_i a_j Cov(X_i, X_j).$$

Da ciò, nel caso $n = 2$, per $a_1 = 1$ e $a_2 = 1$ si trae

$$\text{Var}\,(X+Y) = \text{Var}\,(X) + \text{Var}\,(Y) + 2Cov(X,Y),$$

mentre per $a_1 = 1$ e $a_2 = -1$ si ha

$$\text{Var}\,(X-Y) = \text{Var}\,(X) + \text{Var}\,(Y) - 2Cov(X,Y).$$

Esempio. Siano $X_1, \ldots, X_n$ variabili indipendenti ed identicamente distribuite con varianza σ^2. Determinare $\mathrm{Var}(\overline{X})$, dove $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ è la media campionaria.

Soluzione. Per l'indipendenza delle variabili $X_1, \ldots, X_n$ si ha

$$\mathrm{Var}(\overline{X}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2}\, n\sigma^2 = \frac{\sigma^2}{n}.$$

Esempio. Siano $X_1, \ldots, X_n$ variabili indipendenti ed identicamente distribuite con valore atteso μ e varianza σ^2. Le variabili $X_i - \overline{X}$ sono chiamate deviazioni, mentre

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \overline{X})^2$$

è chiamata varianza campionaria. Determinare $E[S^2]$.

Soluzione. Consideriamo la seguente identità algebrica:

$$\begin{aligned}(n-1)S^2 &= \sum_{i=1}^{n}(X_i - \mu + \mu - \overline{X})^2\\ &= \sum_{i=1}^{n}(X_i-\mu)^2 + \sum_{i=1}^{n}(\overline{X}-\mu)^2 - 2(\overline{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)\\ &= \sum_{i=1}^{n}(X_i-\mu)^2 + n(\overline{X}-\mu)^2 - 2(\overline{X}-\mu)n(\overline{X}-\mu)\\ &= \sum_{i=1}^{n}(X_i-\mu)^2 - n(\overline{X}-\mu)^2\end{aligned}$$

Prendendo i valori attesi di entrambi i membri dell'uguaglianza precedente si ottiene

$$(n-1)E[S^2] = \sum_{i=1}^{n} E[(X_i-\mu)^2] - nE[(\overline{X}-\mu)^2] = n\sigma^2 - n\mathrm{Var}(\overline{X}) = (n-1)\sigma^2,$$

dove si è utilizzato il fatto che $E[\overline{X}] = \mu$ e $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$ nell'uguaglianza finale. Dividendo per $n-1$ si ottiene che $E[S^2] = \sigma^2$.

Esempio. Calcolare la varianza di una variabile aleatoria binomiale X di parametri n e p.

Soluzione. Dato che una tale variabile rappresenta il numero di successi in n prove indipendenti quando il successo in una prova ha probabilità p, possiamo scrivere

$$X = X_1 + \cdots + X_n,$$

dove le X_i sono variabili di Bernoulli indipendenti tali che

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esima prova è un successo,} \\ 0 & \text{altrimenti.} \end{cases}$$

Ricordando che i momenti di una variabile di Bernoulli sono $E[X_i^n] = p$, si ha

$$\mathrm{Var}(X_i) = E[X_i^2] - (E[X_i])^2 = p - p^2 = p\,(1-p).$$

Pertanto, dalla proprietà di linearità si ottiene infine

$$\mathrm{Var}(X) = \mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = n\,p\,(1-p).$$

Esempio. Varianza del numero di accoppiamenti. Calcolare la varianza di X, il numero di persone che seleziona il proprio cappello tra N.

Soluzione Utilizzando la rappresentazione di X usata in un esempio precedente, si ha che $X = X_1 + X_2 + \cdots + X_N$, dove

$$X_i = \begin{cases} 1 & \text{se la persona } i \text{ prende il suo cappello,} \\ 0 & \text{altrimenti.} \end{cases}$$

Si ha poi

$$\mathrm{Var}(X) = \sum_{i=1}^{N} \mathrm{Var}(X_i) + 2 \sum_{i<j} Cov(X_i, X_j).$$

Essendo $P(X_i = 1) = 1/N$, si ha

$$E[X_i] = \frac{1}{N}, \qquad \mathrm{Var}(X_i) = \frac{1}{N}\left(1 - \frac{1}{N}\right) = \frac{N-1}{N^2}.$$

Inoltre

$$X_iX_j = \begin{cases} 1 & \text{se l'}i\text{-esima e la }j\text{-esima persona selezionano il loro cappello,} \\ 0 & \text{altrimenti.} \end{cases}$$

Pertanto si ha

$$E[X_iX_j] = \sum_{u=0}^{1}\sum_{v=0}^{1} u\,v\,p(u,v) = p(1,1) = P(X_i = 1)\,P(X_j = 1|X_i = 1) = \frac{1}{N}\,\frac{1}{N-1}$$

da cui segue

$$Cov(X_i, X_j) = \frac{1}{N(N-1)} - \left(\frac{1}{N}\right)^2 = \frac{N-(N-1)}{N^2(N-1)} = \frac{1}{N^2(N-1)}$$

e quindi

$$\mathrm{Var}(X) = \frac{N-1}{N} + 2\binom{N}{2}\frac{1}{N^2(N-1)} = \frac{N-1}{N} + \frac{1}{N} = 1$$

In conclusione sia media che varianza del numero di accoppiamenti sono pari a 1.

Il coefficiente di correlazione di due variabili aleatorie X e Y, indicata con $\rho(X,Y)$, è così definita, se le varianze $\mathrm{Var}(X)$ e $\mathrm{Var}(Y)$ sono positive e finite:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} = \frac{Cov(X,Y)}{\sigma_X\,\sigma_Y},$$

dove $\sigma_X = \sqrt{\mathrm{Var}(X)}$ e $\sigma_Y = \sqrt{\mathrm{Var}(Y)}$ sono le deviazioni standard di X e Y.

Proposizione. Risulta

$$-1 \leq \rho(X,Y) \leq 1.$$

Dimostrazione. Denotando con σ_X^2 e σ_Y^2 le varianze di X e Y, si ha

$$0 \leq \mathrm{Var}\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = \mathrm{Var}\left(\frac{X}{\sigma_X}\right) + \mathrm{Var}\left(\frac{Y}{\sigma_Y}\right) + 2\,Cov\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right)$$

$$= \frac{\mathrm{Var}(X)}{\sigma_X^2} + \frac{\mathrm{Var}(Y)}{\sigma_Y^2} + 2\,\frac{Cov(X,Y)}{\sigma_X\,\sigma_Y} = 2\,[1 + \rho(X,Y)]$$

da cui segue $\rho(X,Y) \geq -1$. D'altronde risulta

$$0 \leq \mathrm{Var}\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = \frac{\mathrm{Var}(X)}{\sigma_X^2} + \frac{\mathrm{Var}(Y)}{\sigma_Y^2} - 2\,\frac{Cov(X,Y)}{\sigma_X\,\sigma_Y} = 2\,[1 - \rho(X,Y)]$$

e quindi si ha anche $\rho(X,Y) \leq 1$.

Ricordiamo che se $\mathrm{Var}(Z) = 0$ allora Z è degenere, ossia costante con probabilità 1. Quindi, dalla precedente dimostrazione segue che

$$\rho(X,Y) = 1 \quad \Leftrightarrow \quad \mathrm{Var}\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 0 \quad \Leftrightarrow \quad P\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = c\right) = 1.$$

Quindi risulta

$$\frac{Y}{\sigma_Y} = \frac{X}{\sigma_X} - c, \quad \text{ossia} \quad Y = \frac{\sigma_Y}{\sigma_X} X - c\,\sigma_Y,$$

pertanto esiste una relazione lineare tra le variabili aleatorie:

$$Y = aX + b, \qquad \text{con } \; a = \frac{\sigma_Y}{\sigma_X} > 0.$$

Similmente, si ha

$$\rho(X,Y) = -1 \quad \Leftrightarrow \quad \mathrm{Var}\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 0 \quad \Leftrightarrow \quad P\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y} = c\right) = 1,$$

pertanto

$$\frac{Y}{\sigma_Y} = -\frac{X}{\sigma_X} + c, \quad \text{ossia} \quad Y = -\frac{\sigma_Y}{\sigma_X} X + c\,\sigma_Y,$$

cosicché anche in questo caso c'è una relazione lineare tra le variabili aleatorie:

$$Y = aX + b, \qquad \text{con } \; a = -\frac{\sigma_Y}{\sigma_X} < 0.$$

In sintesi, se $Y = aX + b$ allora $\rho(X,Y) = \pm 1$, a seconda del segno di a. Infatti, poiché $Cov(X, aX+b) = a\,\mathrm{Var}(X)$, e $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$, si ha

$$\rho(X,Y) = \rho(X, aX+b) = \frac{a\,\mathrm{Var}(X)}{\sqrt{\mathrm{Var}(X)\,a^2\,\mathrm{Var}(X)}} = \frac{a}{|a|} = \begin{cases} 1, & a > 0 \\ -1, & a < 0. \end{cases}$$

ll coefficiente di correlazione è una misura del grado di linearità tra X e Y. Un valore di $\rho(X,Y)$ vicino a $+1$ o a -1 indica un alto livello di linearità tra X e Y, mentre un valore vicino a 0 indica un'assenza di tale linearità.

Si noti che $\rho(X,Y)$ ha lo stesso segno di $Cov(X,Y)$:
un valore positivo di $\rho(X,Y)$ indica che Y tende a crescere, in media, quando X cresce, mentre $\rho(X,Y) < 0$ indica che Y tende a decrescere, in media, quando X cresce;
se $\rho(X,Y) = 0$ allora X e Y sono scorrelate.

Dalla definizione segue che il coefficiente di correlazione

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \qquad \text{è adimensionale.}$$

Infatti, $Cov(X,Y)$ ha le dimensioni di $X \cdot Y$, così come $\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}$.

Esercizio Nel lancio di un dado ripetuto n volte, sia X il numero di volte che esce 6 e Y il numero di volte che esce 5. Calcolare il coefficiente di correlazione di (X, Y).

Soluzione. Definiamo le variabili aleatorie di Bernoulli di parametro $p = 1/6$:

$$X_i = \begin{cases} 1 & \text{se esce 6 al lancio } i\text{-esimo} \\ 0 & \text{altrimenti,} \end{cases} \qquad (i = 1, 2, \ldots, n),$$

$$Y_j = \begin{cases} 1 & \text{se esce 5 al lancio } j\text{-esimo} \\ 0 & \text{altrimenti,} \end{cases} \qquad (j = 1, 2, \ldots, n),$$

per cui risulta

$$E(X_i) = E(Y_j) = p = \frac{1}{6}, \qquad \mathrm{Var}(X_i) = \mathrm{Var}(Y_j) = p(1-p) = \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{36}.$$

Le variabili X_i e Y_j sono indipendenti se $i \neq j$, ma non lo sono se $i = j$, quindi

$$E(X_i Y_j) = \begin{cases} 0 & \text{se } i = j, \quad \text{essendo } P(X_i Y_i = 0) = 1 \\ E(X_i)\, E(Y_j) = \dfrac{1}{36} & \text{se } i \neq j, \quad \text{essendo } X_i \text{ e } Y_j \text{ indipendenti.} \end{cases}$$

Pertanto risulta

$$\mathrm{Cov}(X_i, Y_j) = E(X_i Y_j) - E(X_i) E(Y_j) = \begin{cases} -\dfrac{1}{36} & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases}$$

Notiamo che

$$X = X_1 + \ldots + X_n \sim Bin\left(n, \frac{1}{6}\right), \qquad Y = Y_1 + \ldots + Y_n \sim Bin\left(n, \frac{1}{6}\right),$$

ossia X e Y sono la somma di n variabili di Bernoulli indipendenti e identicamente distribuite (i.i.d.) e pertanto hanno distribuzione binomiale di parametri n e $p = 1/6$, cosicché

$$\mathrm{Var}(X) = \mathrm{Var}(Y) = np(1-p) = n\,\frac{5}{36}.$$

Inoltre, dalle proprietà della covarianza segue

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, Y_j) = -\frac{n}{36},$$

Infine si ottiene che X e Y sono negativamente correlate:

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{-n/36}{n\,5/36} = -\frac{1}{5} = -0{,}2 < 0.$$

Notiamo che $\rho(X,Y)$ non dipende da n, a differenza di $\mathrm{Cov}(X,Y)$.

Notiamo che (X, Y) ha distribuzione multinomiale di parametri $(n; \frac{1}{6}, \frac{1}{6}, \frac{4}{6})$, cioè

$$P(X = x, Y = y) = \frac{n!}{x!\, y!\, (n - x - y)!} \left(\frac{1}{6}\right)^x \left(\frac{1}{6}\right)^y \left(\frac{4}{6}\right)^{n-x-y}$$

per $x, y = 0, 1, \ldots, n$ e $x + y \le n$. Quindi risulta

$$\begin{aligned}
E[XY] &= \sum_{x=0}^{n} \sum_{y=0}^{n-x} x\, y\, \frac{n!}{x!\, y!\, (n - x - y)!} \left(\frac{1}{6}\right)^x \left(\frac{1}{6}\right)^y \left(\frac{4}{6}\right)^{n-x-y} \\
&= \sum_{x=1}^{n} \sum_{y=1}^{n-x} \frac{n!}{(x-1)!\, (y-1)!\, (n - x - y)!} \left(\frac{1}{6}\right)^x \left(\frac{1}{6}\right)^y \left(\frac{4}{6}\right)^{n-x-y} \\
&= \frac{n(n-1)}{36} \sum_{x=1}^{n} \sum_{y=1}^{n-x} \frac{(n-2)!}{(x-1)!\, (y-1)!\, (n - x - y)!} \left(\frac{1}{6}\right)^{x-1} \left(\frac{1}{6}\right)^{y-1} \left(\frac{4}{6}\right)^{n-x-y} \\
&= \frac{n(n-1)}{36}
\end{aligned}$$

(La somma della distribuzione multinomiale di parametri $(n - 2; \frac{1}{6}, \frac{1}{6}, \frac{4}{6})$ vale 1). Pertanto

$$Cov(X, Y) = E[XY] - E[X]\, E[Y] = \frac{n(n-1)}{36} - \frac{n}{6} \cdot \frac{n}{6} = -\frac{n}{36}.$$

Esempio Siano X e Y tali che $p(0,0) = p(1,1) = p$ e $p(1,0) = p(0,1) = \frac{1}{2} - p$.

(i) Determinare i valori ammissibili di p.

(ii) Stabilire per quali valori di p si ha che X e Y sono indipendenti.

(iii) Calcolare $\rho(X,Y)$ e studiarlo al variare di p.

Soluzione. Dalle condizioni $p(x,y) \geq 0 \ \forall x,y$ segue $p \geq 0$ e $\frac{1}{2} - p \geq 0$, mentre $\sum_x \sum_y p(x,y) = 1$ per ogni $p \in \mathbb{R}$; quindi i valori ammissibili di p sono

$$0 \leq p \leq \frac{1}{2}.$$

(ii) Dalla seguente tabella segue che $p(x,y) = p_X(x)\, p_Y(y)$ per ogni x e y se e solo se $p = 1/4$.

Inoltre notiamo che X e Y sono scambiabili.

$x \setminus y$	0	1	$p_X(x)$
0	p	$1/2 - p$	$1/2$
1	$1/2 - p$	p	$1/2$
$p_Y(y)$	$1/2$	$1/2$	1

Poiché X e Y sono entrambe variabili aleatorie di Bernoulli di parametro $1/2$, si ha

$$E[X] = E[Y] = \frac{1}{2}, \qquad \mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}.$$

Inoltre si ha

$$E[X\,Y] = \sum_{x=0}^{1} \sum_{y=0}^{1} x\,y\,p(x,y) = p(1,1) = p,$$

da cui segue

$$Cov(X,Y) = E[XY] - E[X]\,E[Y] = p - \frac{1}{4}.$$

Quindi il coefficiente di correlazione è dato da

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{p-\frac{1}{4}}{\frac{1}{4}} = 4\,p - 1, \qquad 0 \le p \le \frac{1}{2}.$$

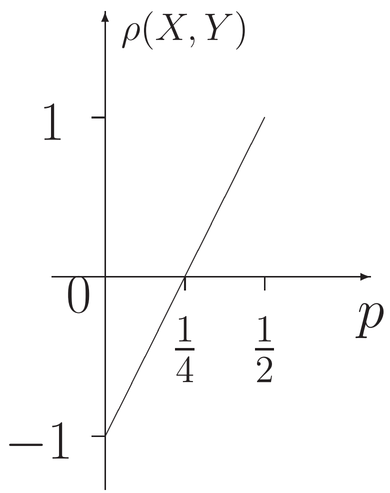
Grafico di $\rho(X,Y) = 4\,p-1$, per $0 \le p \le \frac{1}{2}$;

risulta:

$\rho(X,Y) = -1$ per $p = 0$,

$\rho(X,Y) = 0$ per $p = 1/4$,

$\rho(X,Y) = 1$ per $p = 1/2$.



Per $p = 0$ si ha $\rho(X, Y) = -1$
ed infatti in tal caso
$P(Y = 1 - X) = p(0, 1) + p(1, 0) = 1$,
ossia $Y = aX + b$, con $a = -1$ e $b = 1$.

$x \setminus y$	0	1	$p_X(x)$
0	0	1/2	1/2
1	1/2	0	1/2
$p_Y(y)$	1/2	1/2	1

Per $p = 1/2$ si ha $\rho(X, Y) = 1$
ed infatti in tal caso
$P(Y = X) = p(0, 0) + p(1, 1) = 1$,
ossia $Y = aX + b$, con $a = 1$ e $b = 0$.

$x \setminus y$	0	1	$p_X(x)$
0	1/2	0	1/2
1	0	1/2	1/2
$p_Y(y)$	1/2	1/2	1

Esercizi per casa

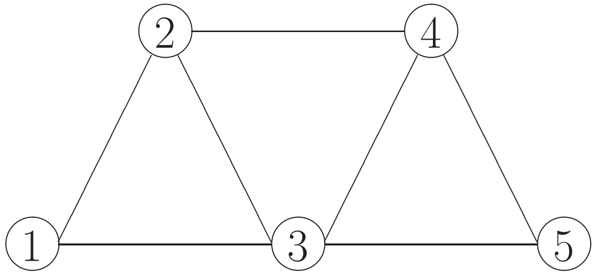
7.1) Un esperimento consiste nel lanciare tre monete. Le prime due monete sono non truccate, mentre la terza è truccata nel senso che in mostra testa con probabilità 1/4. Sia X il numero volte che esce testa, e sia Y il numero di volte che si registrano delle variazioni.
(i) Determinare le funzioni di probabilità congiunta $p(x, y)$ e marginali $p_X(x)$ e $p_Y(y)$.
(ii) Stabilire se X e Y sono indipendenti.
(iii) Calcolare il coefficiente di correlazione $\rho(X, Y)$ e commentare il risultato ottenuto.

7.2) Da un'urna contenente 2 biglie nere, 2 biglie bianche e 1 biglia rossa si effettuano 3 estrazioni a caso senza reinserimento. Denotiamo con X il numero di biglie nere estratte, e con Y il numero di biglie bianche estratte. Determinare
(i) $E(X + Y)$ e $\text{Var}(X + Y)$,
(ii) $E(X - Y)$ e $\text{Var}(X - Y)$.

7.3) Siano X e Y variabili aleatorie congiuntamente distribuite.
(i) Calcolare $Cov(X + Y, X - Y)$.
(ii) Stabilire le condizioni affinché tale quantità sia positiva, nulla, negativa.

7.4) Siano X e Y variabili aleatorie di Bernoulli di parametro 1/4 e 1/2, rispettivamente, e sia $P(X = Y) = 1/4$.
(i) Si può affermare che X e Y sono indipendenti?
(ii) Calcolare il coefficiente di correlazione $\rho(X, Y)$ e commentare il risultato ottenuto.

7.5) Consideriamo l'esperimento che consiste nello scegliere a caso due dei 5 nodi del grafo in figura.



Sia $X = 1$ se i due nodi scelti sono collegati direttamente da un arco, e sia $X = 0$ altrimenti. Indichiamo con Y quanti sono i nodi scelti che sono numerati con un numero dispari.

(i) Determinare le funzioni di probabilità congiunta $p(x, y)$ e marginali $p_X(x)$ e $p_Y(y)$.

(ii) Stabilire se X e Y sono indipendenti.

(iii) Calcolare il coefficiente di correlazione $\rho(X, Y)$ e commentare il risultato ottenuto.

(iv) Calcolare $p_{Y|X}(y|0)$ e $p_{Y|X}(y|1)$ per $y = 0, 1, 2$.

(v) Ricavare $E(Y|X = x)$ e $\mathrm{Var}(Y|X = x)$ per $x = 0, 1$.

Calcolo delle Probabilità e Statistica Matematica – A.A. 2009/10

CAPITOLO 8 – Teoremi Limite

8.1 Introduzione

8.2 La disuguaglianza di Chebyshev e la legge debole dei grandi numeri

8.3 Il teorema del limite centrale

8.1 Introduzione

I risultati più significativi del calcolo delle probabilità sono certamente i teoremi limite.

Di questi, i più importanti sono quelli che vengono classificati come *leggi dei grandi numeri* e come *teoremi del limite centrale*.

In generale, un teorema viene considerato una legge dei grandi numeri quando stabilisce delle condizioni sotto le quali la media di una successione di variabili aleatorie converge (in qualche senso) alla media attesa.

D'altra parte, i teoremi del limite centrale sono caratterizzati dal determinare condizioni sotto le quali la somma di un grande numero di variabili aleatorie (opportunamente standardizzata) ha la distribuzione di probabilità che tende a essere normale.

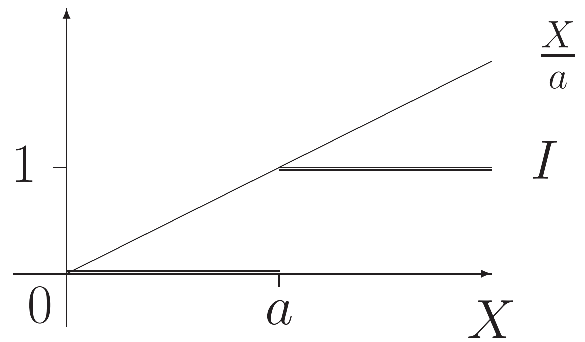
8.2 La disuguaglianza di Chebyshev e la legge debole dei grandi numeri

Proposizione. Disuguaglianza di Markov. Se X è una variabile aleatoria che assume solo valori non negativi, allora per ogni numero reale $a > 0$,

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Dimostrazione. Per $a > 0$, sia

$$I = \mathbf{1}_{\{X \geq a\}} = \begin{cases} 1 & \text{se } X \geq a \\ 0 & \text{altrimenti.} \end{cases}$$



Ricordando che $X \geq 0$, si ha: $I \leq \dfrac{X}{a}$. Infatti:

se $X < a$ si ha: $I = 0 \leq \dfrac{X}{a}$; se $X \geq a$ si ha: $I = 1 \leq \dfrac{X}{a}$.

Dalle proprietà del valore atteso segue: $E[I] \leq \dfrac{E[X]}{a}$.

Per la variabile di Bernoulli I risulta $E[I] = P(I = 1) = P(X \geq a)$, quindi si ha

$$P(X \geq a) \leq \frac{E(X)}{a}. \qquad \text{(Tale disuguaglianza è significativa per } a > E(X).)$$

Proposizione. Disuguaglianza di Chebyshev. Se X è una variabile aleatoria di media μ e varianza σ^2, entrambe finite, allora per ogni numero reale $k > 0$,

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

ovvero

$$P\{|X-\mu| < k\} \geq 1 - \frac{\sigma^2}{k^2}.$$

Dimostrazione. Essendo $(X-\mu)^2$ una variabile aleatoria non negativa, possiamo applicare la disuguaglianza di Markov, con $a = k^2$:

$$P(X \geq a) \leq \frac{E(X)}{a} \qquad \text{diventa} \qquad P\{(X-\mu)^2 \geq k^2\} \leq \frac{E[(X-\mu)^2]}{k^2}.$$

Essendo $(X-\mu)^2 \geq k^2$ se e solo se $|X-\mu| \geq k$, si ricava infine:

$$P\{|X-\mu| \geq k\} \leq \frac{E[(X-\mu)^2]}{k^2} = \frac{\sigma^2}{k^2}.$$

Notiamo che $P\{|X-\mu| \geq k\} = 1 - P\{|X-\mu| < k\} \leq \frac{\sigma^2}{k^2}$, da cui segue

$$P\{\mu - k < X < \mu + k\} = P\{-k < X - \mu < k\} = P\{|X-\mu| < k\} \geq 1 - \frac{\sigma^2}{k^2}.$$

Esempio. Si supponga che il numero di studenti che si iscrive ad un corso di lingua inglese è descritto da una variabile aleatoria X di media $\mu = 50$.
(a) Cosa si può dire su $P(X \geq 75)$?
(b) Cosa si può dire su $P(40 < X < 60)$ se $\sigma^2 = 25$?

Soluzione. (a) Facendo uso della disuguaglianza di Markov, con $a = 75$, si ha

$$P(X \geq 75) \leq \frac{E(X)}{75} = \frac{50}{75} = \frac{2}{3}.$$

(b) Ricordiamo che dalla disuguaglianza di Chebyshev segue

$$P\{|X - \mu| < k\} \geq 1 - \frac{\sigma^2}{k^2}.$$

Quindi, essendo $\mu = 50$ e $\sigma^2 = 25$, per $k = 10$ si ottiene

$$P(40 < X < 60) = P\{|X - 50| < 10\} = P\{|X - \mu| < k\} \geq 1 - \frac{\sigma^2}{k^2} = 1 - \frac{25}{100} = \frac{3}{4}.$$

Esempio. Sia X uniformemente distribuita su $(0, 10)$. Allora, poiché

$$E[X] = \frac{10}{2} = 5, \qquad \mathrm{Var}(X) = \frac{(10)^2}{12} = \frac{25}{3},$$

dalla disuguaglianza di Chebyshev segue

$$P\{|X-5| > 4\} \le \frac{25}{3 \cdot 16} = \frac{25}{48} \approx 0{,}52.$$

Ricordando che $F(x) = \dfrac{x}{10}$ per $0 \le x \le 10$, il valore esatto è

$$\begin{aligned} P\{|X-5| > 4\} &= P\{-(X-5) > 4\} + P\{X-5 > 4\} = P(X < 1) + P(X > 9) \\ &= F(1) + 1 - F(9) = \frac{1}{10} + 1 - \frac{9}{10} = \frac{2}{10} = 0{,}2. \end{aligned}$$

Esempio. Sia X una variabile normale di media μ e varianza σ^2. Risulta

$$P\{\mu - h\,\sigma < X < \mu + h\,\sigma\} = P\Big(-h < \frac{X-\mu}{\sigma} < h\Big) = \Phi(h) - \Phi(-h) = 2\,\Phi(h) - 1,$$

essendo $\dfrac{X-\mu}{\sigma}$ normale standard, ed essendo $\Phi(-h) = 1 - \Phi(h)$.
Invece dalla disuguaglianza di Chebyshev, per $k = h\,\sigma$, si ha

$$P\{\mu - h\,\sigma < X < \mu + h\,\sigma\} = P\{|X-\mu| < h\,\sigma\} \ge 1 - \frac{\sigma^2}{(h\,\sigma)^2} = 1 - \frac{1}{h^2}.$$

	$h = 1$	$h = 1{,}5$	$h = 2$	$h = 2{,}5$	$h = 3$
$2\,\Phi(h) - 1$	0,6826	0,8664	0,9544	0,9876	0,9974
$1 - 1/h^2$	0	0,5555	0,75	0,84	0,8888

Teorema. La legge debole dei grandi numeri. Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie i.i.d., ognuna con media finita $E[X_i] = \mu$. Allora, posto

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

per ogni $\epsilon > 0$ si ha

$$\lim_{n\to\infty} P\{|\overline{X}_n - \mu| \geq \epsilon\} = 0.$$

Equivalentemente, per ogni $\epsilon > 0$ risulta

$$\lim_{n\to\infty} P(\mu - \epsilon < \overline{X}_n < \mu + \epsilon) = \lim_{n\to\infty} P\{|\overline{X}_n - \mu| < \epsilon\} = 1.$$

Dimostrazione. Assumiamo inoltre che $\mathrm{Var}(X_i) = \sigma^2$ sia finita. Poiché $E[\overline{X}_n] = \mu$ e $\mathrm{Var}[\overline{X}_n] = \sigma^2/n$, dalla disuguaglianza di Chebyshev si ha, per ogni $\epsilon > 0$

$$P\{|\overline{X}_n - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{n} \frac{1}{\epsilon^2}.$$

Procedendo al limite si ha

$$\lim_{n\to\infty} P\{|\overline{X}_n - \mu| \geq \epsilon\} \leq \lim_{n\to\infty} \frac{\sigma^2}{n} \frac{1}{\epsilon^2} = 0,$$

da cui segue immediatamente l'asserto.

8.3 Il teorema del limite centrale

Vedremo ora uno dei più importanti risultati della teoria della probabilità. Esso, in breve, asserisce che la somma di un grande numero di variabili aleatorie indipendenti ha distribuzione approssimativamente normale. La dimostrazione è omessa per brevità.

Ricordiamo che se le variabili aleatorie $X_1, \ldots, X_n$ sono indipendenti e identicamente distribuite (i.i.d.) con valore atteso $E[X_i] = \mu$ e varianza $\mathrm{Var}[X_i] = \sigma^2$, allora si ha

$$E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] = n\mu,$$

$$\mathrm{Var}[X_1 + \cdots + X_n] = \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n] = n\sigma^2.$$

Pertanto,

$$Z_n = \frac{X_1 + \cdots + X_n - E[X_1 + \cdots + X_n]}{\sqrt{\mathrm{Var}(X_1 + \cdots + X_n)}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

è una variabile aleatoria standard.

Teorema. Il teorema del limite centrale. Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie i.i.d., ognuna di media μ e varianza $\sigma^2 > 0$ finite. Allora, posto

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}},$$

per ogni $a \in \mathbb{R}$ si ha: $\lim_{n\to\infty} P\{Z_n \le a\} = \Phi(a) \qquad \left(= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\, dx\right)$.

In altri termini, la somma standardizzata

$$Z_n = \frac{X_1 + \cdots + X_n - E[X_1 + \cdots + X_n]}{\sqrt{\mathrm{Var}(X_1 + \cdots + X_n)}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tende ad avere distribuzione normale standard quando $n \to \infty$.

Valgono quindi le seguenti approssimazioni, per n grande:

$$P(Z_n \le a) = P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a\right) \approx \Phi(a),$$

$$P(a \le Z_n \le b) = P\left(a \le \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le b\right) \approx \Phi(b) - \Phi(a).$$

Notiamo che dalla seguente approssimazione, valida per n grande,

$$P(Z_n \le a) = P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a\right) \approx \Phi(a),$$

si ricava

$$P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a\right) = P\left(X_1 + \cdots + X_n \le n\mu + a\sigma\sqrt{n}\right) \approx \Phi(a).$$

Ponendo

$$x = n\mu + a\sigma\sqrt{n}$$

si ricava

$$a = \frac{x - n\mu}{\sigma\sqrt{n}}.$$

Pertanto, per n grande si ha

$$P\left(X_1 + \cdots + X_n \le x\right) \approx \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right).$$

Analogamente, per n grande risulta

$$P\left(y \le X_1 + \cdots + X_n \le x\right) \approx \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{y - n\mu}{\sigma\sqrt{n}}\right).$$

Esempio. Si genera a caso una sequenza di $n = 10.000$ bit $X_1, \dots, X_n$, ossia variabili aleatorie di Bernoulli indipendenti di parametro $1/2$. Posto

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

determinare un limite inferiore e un'approssimazione per

$$P(0{,}49 < \overline{X}_n < 0{,}51).$$

Soluzione. Ricordiamo che $\mu = E[X_i] = 1/2$ e $\sigma^2 = \mathrm{Var}(X_i) = 1/4$, pertanto

$$E(\overline{X}_n) = \mu = \frac{1}{2}, \qquad \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n} = \frac{1}{4n} = \frac{1}{4\cdot 10^4}.$$

Dalla disuguaglianza di Chebyshev si trae il limite inferiore

$$P\{|\overline{X}_n - \mu| < 0{,}01\} \geq 1 - \frac{\mathrm{Var}(\overline{X}_n)}{(0{,}01)^2} = 1 - \frac{1}{4\cdot 10^4}\,\frac{1}{(0{,}01)^2} = 1 - \frac{1}{4} = \frac{3}{4},$$

ossia

$$P(0{,}49 < \overline{X}_n < 0{,}51) = P\{|\overline{X}_n - \mu| < 0{,}01\} \geq 0{,}75.$$

Notiamo che, essendo $n = 10.000$ si ha

$$\begin{aligned} P(0{,}49 < \overline{X}_n < 0{,}51) &= P\Big(0{,}49 < \frac{X_1 + \ldots + X_n}{n} < 0{,}51\Big) \\ &= P(4.900 < X_1 + \ldots + X_n < 5.100) \\ &= P(4.900{,}5 < X_1 + \ldots + X_n < 5.099{,}5). \end{aligned}$$

Nell'ultima uguaglianza si è effettuata una correzione di continuità, per poi approssimare la somma di variabili discrete con una variabile continua. Poiché risulta

$$E[X_1 + \ldots + X_n] = n\mu = 5.000 \qquad \sqrt{\mathrm{Var}(X_1 + \ldots + X_n)} = \sigma\sqrt{n} = \frac{1}{2}\,100 = 50$$

standardizzando si ha

$$P(0{,}49 < \overline{X}_n < 0{,}51) = P\left(\frac{4.900{,}5 - n\mu}{\sigma\sqrt{n}} < \frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} < \frac{5.099{,}5 - n\mu}{\sigma\sqrt{n}}\right)$$

$$= P\left(\frac{4.900{,}5 - 5.000}{50} < Z_n < \frac{5.099{,}5 - 5.000}{50}\right) = P(-1{,}99 < Z_n < 1{,}99)$$
$$\approx \Phi(1{,}99) - \Phi(-1{,}99) = 2\,\Phi(1{,}99) - 1 = 2 \cdot 0{,}9767 - 1 = 0{,}9534$$

avendo usato l'approssimazione basata sul teorema del limite centrale.

Esercizio. Cosa cambia nell'esempio precedente se $n = 1.000.000$?

Esempio. Si supponga che il numero di studenti che si iscrive ad un corso universitario è descritto da una variabile aleatoria di Poisson X di media $\mu = 100$. Qual è la probabilità che si iscrivano almeno 120 studenti?

Soluzione. Essendo $P(X=k) = e^{-\mu}\dfrac{(\mu)^k}{k!}$, non è agevole determinare il valore esatto

$$P(X \geq 120) = \sum_{k=120}^{\infty} P(X=k) = e^{-100} \sum_{k=120}^{\infty} \frac{(100)^k}{k!} = 1 - e^{-100} \sum_{k=0}^{119} \frac{(100)^k}{k!}.$$

Ricordando che X può essere riguardata come somma di $n = 100$ variabili aleatorie di Poisson indipendenti di media 1 e varianza 1, si ha $E[X] = 100$ e $\mathrm{Var}(X) = 100$. Quindi possiamo adoperare l'approssimazione basata sul teorema del limite centrale:

$$\begin{aligned} P(X \geq 120) &= P(X \geq 119{,}5) \qquad \text{(la correzione di continuità)} \\ &= P\left(\frac{X-100}{\sqrt{100}} \geq \frac{119{,}5-100}{\sqrt{100}}\right) = P\left(Z \geq 1{,}95\right) \\ &\approx 1 - \Phi(1{,}95) = 1 - 0{,}9744 = 0{,}0256 \end{aligned}$$

essendo $Z = \dfrac{X - E[X]}{\sqrt{\mathrm{Var}(X)}} = \dfrac{X-100}{\sqrt{100}}$ normale standard.

Esempio. Determinare la probabilità approssimata che lanciando a caso 10 dadi la somma X dei valori ottenuti sia compresa tra 30 e 40, estremi inclusi.

Soluzione. Indicando l'esito del lancio i-esimo con X_i, questa è una variabile uniforme discreta con $E[X_i] = \frac{N+1}{2} = \frac{7}{2}$ e $\text{Var}(X_i) = \frac{N^2-1}{12} = \frac{35}{12}$. Pertanto risulta $X = X_1 + \ldots + X_{10}$, e quindi

$$E[X] = 10 \cdot \frac{7}{2} = 35, \qquad \text{Var}(X) = 10 \cdot \frac{35}{12} = \frac{350}{12} = 29{,}1666.$$

L'approssimazione basata sul teorema del limite centrale dà quindi:

$$\begin{aligned}
P(30 \le X \le 40) &= P(29{,}5 \le X \le 40{,}5) \qquad \text{(la correzione di continuità)} \\
&= P\left(\frac{29{,}5-35}{\sqrt{29{,}1666}} \le \frac{X-35}{\sqrt{29{,}1666}} \le \frac{40{,}5-35}{\sqrt{29{,}1666}}\right) \\
&= P\left(-1{,}02 \le Z \le 1{,}02\right) \\
&\approx \Phi(1{,}02) - \Phi(-1{,}02) \\
&= 2\,\Phi(1{,}02) - 1 = 2 \cdot 0{,}8461 - 1 = 0{,}6922
\end{aligned}$$

essendo $Z = \frac{X - E[X]}{\sqrt{\text{Var}(X)}} = \frac{X-35}{\sqrt{29{,}1666}}$ normale standard.

Esempio. Siano X_i, $i = 1, \ldots, 24$, variabili aleatorie indipendenti, con distribuzione uniforme su $(0, 1)$. Determinare in modo approssimato i valori di

$$P\left(\sum_{i=1}^{24} X_i > 14\right), \qquad P\left(8 < \sum_{i=1}^{24} X_i < 16\right).$$

Soluzione. Poiché $E[X_i] = 1/2$ e $\mathrm{Var}(X_i) = 1/12$, posto $X = \sum_{i=1}^{24} X_i$ si ha $E[X] = 12$ e $\mathrm{Var}(X) = 2$. Quindi, dal teorema del limite centrale si trae:

$$P(X > 14) = P\left(\frac{X-12}{\sqrt{2}} > \frac{14-12}{\sqrt{2}}\right) \approx 1 - \Phi(1{,}41) = 1 - 0{,}9207 = 0{,}0793$$

ed inoltre

$$\begin{aligned} P(8 < X < 16) &= P\left(\frac{8-12}{\sqrt{2}} \le \frac{X-12}{\sqrt{2}} \le \frac{16-12}{\sqrt{2}}\right) \\ &= P\left(-2{,}83 \le Z \le 2{,}83\right) \\ &\approx \Phi(2{,}83) - \Phi(-2{,}83) \\ &= 2\,\Phi(2{,}83) - 1 = 2 \cdot 0{,}9977 - 1 = 0{,}9954 \end{aligned}$$

essendo $Z = \dfrac{X - E[X]}{\sqrt{\mathrm{Var}(X)}} = \dfrac{X-12}{\sqrt{2}}$ normale standard.

Teorema di DeMoivre-Laplace. Sia S_n il numero di successi che si realizzano in n prove indipendenti, in ognuna delle quali il successo ha probabilità p, e sia Z una variabile normale standard. Per $a < b$ si ha:

$$\lim_{n\to+\infty} P\left\{a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right\} = P\left(a \le Z \le b\right) = \Phi\big(b\big) - \Phi\big(a\big).$$

Il teorema di DeMoivre-Laplace afferma che, se n è grande, la standardizzata di una variabile aleatoria binomiale di parametri n e p ha approssimativamente la stessa distribuzione di una variabile aleatoria normale standard.

Esempio. Sia X il numero di volte che esce testa lanciando $n = 40$ volte una moneta equa ($p = 0{,}5$). Determinare $P(X = 20)$. Ricavarne poi un'approssimazione normale e confrontare il risultato con la soluzione esatta. Cosa cambia se $p = 0{,}4$?

Soluzione. Per $p = 0{,}5$ si ha $E(X) = 20$ e $\mathrm{Var}(X) = 10$; il risultato esatto è

$$P(X = 20) = \binom{40}{20} (0{,}5)^{40} \simeq 0{,}1254.$$

Usiamo l'approssimazione normale. Mentre la variabile aleatoria binomiale è discreta a valori interi, la normale è una variabile continua ed è essenziale scrivere $P(X = i)$ come $P(i - 1/2 < X < i + 1/2)$ prima di applicare l'approssimazione normale. Si ha:

$$P(X = 20) = P\big(19{,}5 \le X < 20{,}5\big) = P\Big(\frac{19{,}5 - 20}{\sqrt{10}} < \frac{X - 20}{\sqrt{10}} < \frac{20{,}5 - 20}{\sqrt{10}}\Big)$$

$$\simeq P\Big(-0{,}16 < \frac{X - 20}{\sqrt{10}} < 0{,}16\Big) = \Phi\big(0{,}16\big) - \Phi\big(-0{,}16\big) = 2\,\Phi\big(0{,}16\big) - 1 = 0{,}1272.$$

Per $p = 0{,}4$ il risultato esatto e l'approssimazione normale sono:

$$P(X = 20) = \binom{40}{20} (0{,}4)^{20} (0{,}6)^{20} \simeq 0{,}0554,$$

$$P(X = 20) = P\Big(\frac{19{,}5 - 16}{\sqrt{9{,}6}} < \frac{X - 16}{\sqrt{9{,}6}} < \frac{20{,}5 - 16}{\sqrt{9{,}6}}\Big) \approx \Phi\big(1{,}45\big) - \Phi\big(1{,}13\big) = 0{,}0557.$$