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 k-esima se ivi si sceglie il numero k.

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:

$$P(A_0) = \left(1 - \frac{1}{n}\right)^n, \qquad P(A_k) = \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k}.$$

Notiamo che se si definisce $X =$"numero di concordanze", appare più adeguato esprimere le probabilità richieste come $P(X = 0) = P(A_0)$ e $P(X = k) = P(A_k)$.

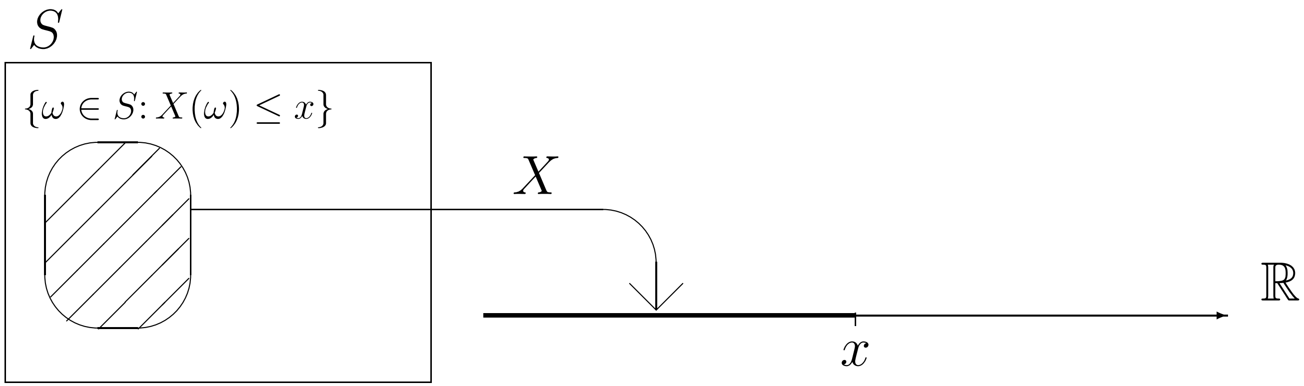
Nella tabella sono riportati i valori delle probabilità di X per varie scelte di n.

	$n=2$	$n=3$	$n=4$	$n=5$
$P(X=0)$	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

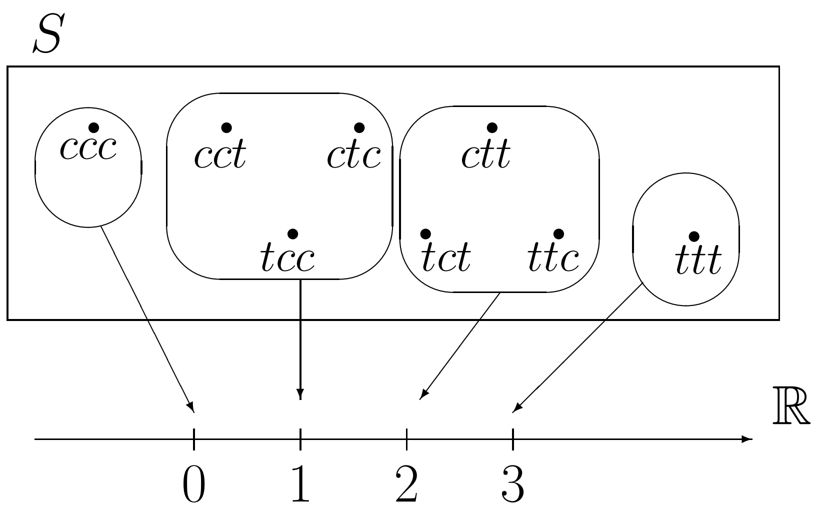
$$\{\omega \in S \colon X(\omega) \le x\} \in \mathcal{F} \qquad \text{per ogni } x \in \mathbb{R}.$$



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. Nel lancio di 3 monete non truccate sia Y il numero di teste che si ottengono; Y è una variabile aleatoria che assume i valori $0, 1, 2, 3$ con le seguenti probabilità:

$$
\begin{aligned}
P(Y=0) &= P(\{ccc\}) = \frac{1}{8} \\
P(Y=1) &= P(\{cct, ctc, tcc\}) = \frac{3}{8} \\
P(Y=2) &= P(\{ctt, tct, ttc\}) = \frac{3}{8} \\
P(Y=3) &= P(\{ttt\}) = \frac{1}{8}
\end{aligned}
$$



Poiché Y deve assumere uno tra i valori $0, 1, 2, 3$ abbiamo

$$
P\left(\bigcup_{i=0}^{3} \{Y=i\}\right) = \sum_{i=0}^{3} P(Y=i) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1.
$$

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

$P(X = i)$	$P(X = 17)$	$P(X = 18)$	$P(X = 19)$	$P(X = 20)$
$\dfrac{\binom{i-1}{2}}{\binom{20}{3}}$	$\dfrac{\binom{16}{2}}{\binom{20}{3}} = \dfrac{2}{19}$	$\dfrac{\binom{17}{2}}{\binom{20}{3}} = \dfrac{34}{285}$	$\dfrac{\binom{18}{2}}{\binom{20}{3}} = \dfrac{51}{380}$	$\dfrac{\binom{19}{2}}{\binom{20}{3}} = \dfrac{3}{20}$

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

$$P(Y = k) = \binom{3}{k} \left(\frac{1}{5}\right)^k \left(\frac{4}{5}\right)^{3-k}, \qquad k = 0, 1, 2, 3.$$

Pertanto la probabilità richiesta è

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \left(\frac{4}{5}\right)^3 = 1 - 0{,}512 = 0{,}488.$$

Esempio. Si lancia ripetutamente una moneta truccata, che in un singolo lancio dà testa con probabilità p, fino a che non appaia testa per la prima volta oppure si siano fatti n lanci. Se X denota il numero totale di volte che lanciamo la moneta, allora X è una variabile aleatoria che assume valori $1, 2, \ldots, n$ con probabilità

$$\begin{aligned} P(X=1) &= P(T_1) = p \\ P(X=k) &= P(C_1 \cap C_2 \cap \ldots \cap C_{k-1} \cap T_k) = (1-p)^{k-1}\, p \qquad (k = 2, 3, \ldots, n-1) \\ P(X=n) &= P(C_1 \cap C_2 \cap \ldots \cap C_n) + P(C_1 \cap C_2 \cap \ldots \cap C_{n-1} \cap T_n) = (1-p)^{n-1}, \end{aligned}$$

dove C_i e T_i rappresentano rispettivamenente la fuoriuscita di croce e testa al lancio i-esimo, e dove si è fatto uso dell'indipendenza nei lanci. Verifichiamo che

$$\begin{aligned} & P\left(\bigcup_{k=1}^{n} \{X=k\}\right) = \sum_{k=1}^{n} P(X=k) = \sum_{k=1}^{n-1} (1-p)^{k-1}\, p + (1-p)^{n-1} \\ & = p \sum_{r=0}^{n-2} (1-p)^r + (1-p)^{n-1} = p \left[\frac{1-(1-p)^{n-1}}{1-(1-p)}\right] + (1-p)^{n-1} = 1, \end{aligned}$$

avendo posto $r = k-1$ e avendo notato che $\sum_{r=0}^{m} c^r = (1-c^{m+1})/(1-c)$ 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. $\lim_{x\to\infty} F(x) = 1$.
3. $\lim_{x\to-\infty} F(x) = 0$.
4. $F(x)$ è continua a destra, ossia $\lim_{h\to 0^+} F(x+h) = F(x)$ 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 b_n è una successione di reali che cresce verso ∞, allora gli eventi $\{X \leq b_n\}$ formano una successione crescente di eventi il cui limite è $\{X < \infty\}$. Quindi, per la proprietà di continuità della probabilità, si ha

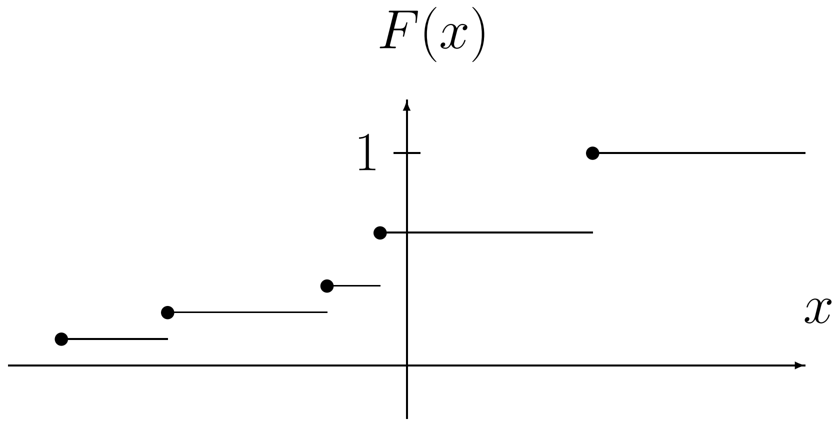
$$\lim_{n\to\infty} P(X \leq b_n) = P(X < \infty) = 1.$$

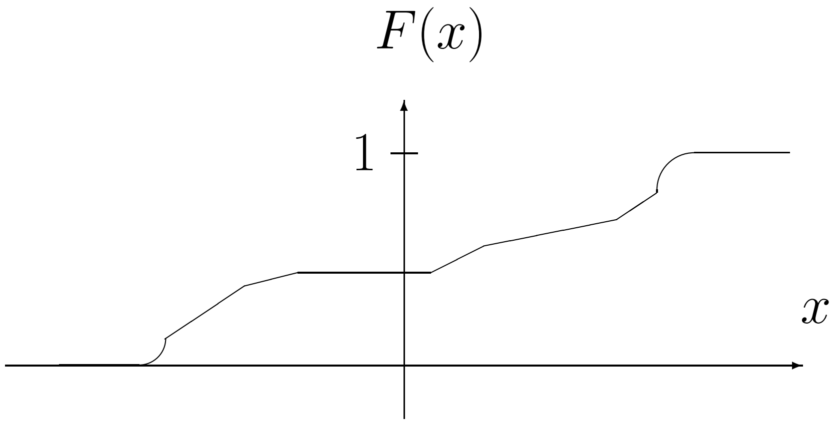
La Proprietà 3 si dimostra in modo analogo alla Proprietà 2, ed è lasciato come esercizio. Per dimostrare la Proprietà 4 notiamo che se x_n, $n \geq 1$, è una successione decrescente che converge a $x \in \mathbb{R}$, allora $\{X \leq x_n\}$, $n \geq 1$, costituisce una successione di eventi decrescente il cui limite coincide con $\{X \leq x\}$. Quindi, in conclusione, la proprietà di continuità fa sì che

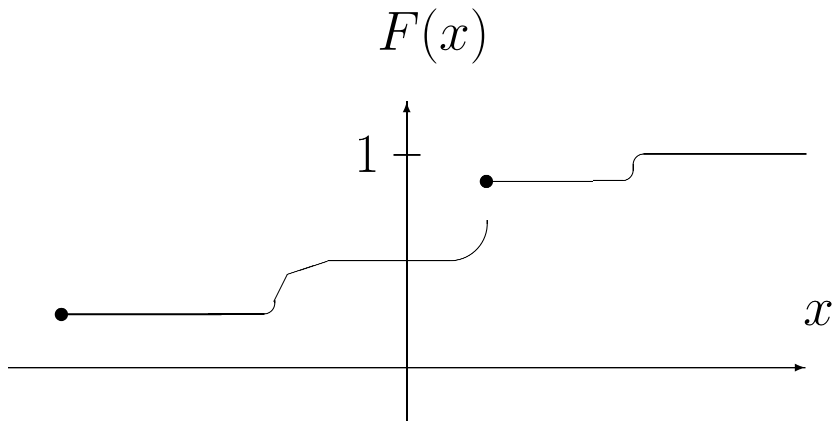
$$\lim_{n\to\infty} P(X \leq x_n) = P(X \leq x).$$

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)

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x < 2 \\ 1, & x \ge 2 \end{cases}$$

(ii)

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{2}, & 0 < x < 1 \\ 1, & x \ge 1 \end{cases}$$

(iii)

$$F(x) = \begin{cases} 0, & x \le -1 \\ \dfrac{x+1}{2}, & -1 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

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) $P(X < b) = \lim_{n\to\infty} F(b_n)$ per ogni $b \in \mathbb{R}$ fissato e per ogni successione crescente b_n, $n \geq 1$, che converge a b.

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 b_n, $n \geq 1$, è una successione crescente che converge a b, allora dalla proprietà di continuità otteniamo:

$$P(X < b) = P\left(\lim_{n\to\infty} \{X \leq b_n\}\right) = \lim_{n\to\infty} P(X \leq b_n) = \lim_{n\to\infty} F(b_n).$$

Una funzione di distribuzione $F(x)$ non è necessariamente continua. Infatti, limite sinistro e limite destro di $F(\cdot)$ in $x \in \mathbb{R}$ non necessariamente coincidono, essendo

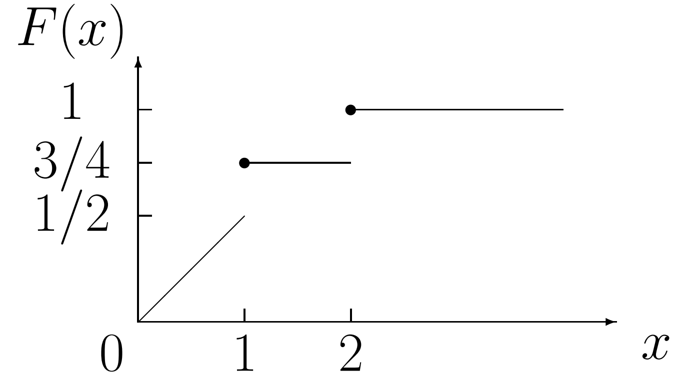
$$F(x^-) := \lim_{h \to 0^-} F(x+h) \leq \lim_{h \to 0^+} F(x+h) = F(x).$$

Se la disuguaglianza precedente è soddisfatta come uguaglianza, allora $F(\cdot)$ è continua in x e risulta $P(X = x) = F(x) - F(x^-) = 0$.

Altrimenti, $F(\cdot)$ è discontinua in x e risulta $P(X = x) = F(x) - F(x^-) > 0$.

Esempio. Sia X una variabile aleatoria avente funzione di distribuzione

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



Calcolare (a) $P(X < 2)$, (b) $P(X = 1)$, (c) $P(X > 1/2)$, (d) $P(1 < X \leq 3)$, (e) $P(1 \leq X \leq 2)$, (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) $P(X > 1/2) = 1 - P(X \leq 1/2) = 1 - F(1/2) = 1 - 1/4 = 3/4$;

(d) $P(1 < X \leq 3) = F(3) - F(1) = 1 - 3/4 = 1/4$;

(e) $P(1 \leq X \leq 2) = P(X \leq 2) - P(X < 1) = F(2) - F(1^-) = 1 - 1/2 = 1/2$;

(f) $P(X > 1|X > 1/2) = \frac{P(X>1)}{P(X>1/2)} = \frac{1-P(X\leq 1)}{3/4} = \frac{1-F(1)}{3/4} = \frac{1-3/4}{3/4} = \frac{1/4}{3/4} = \frac{1}{3}$.

4.3 Variabili aleatorie discrete

Una variabile aleatoria che possa assumere al più un'infinità numerabile di valori è detta *discreta*. Per una variabile aleatoria discreta X, 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

$$p(x_i) \geq 0 \qquad i = 1, 2, \ldots,$$

$$p(x) = 0 \qquad \text{altrimenti.}$$

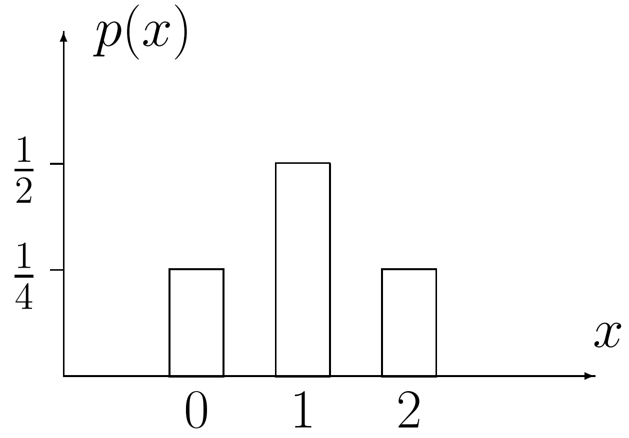
Poiché X deve assumere almeno uno dei valori x_i, abbiamo che

$$\sum_{i=1}^{\infty} p(x_i) = \sum_{i=1}^{\infty} P(X = x_i) = P\left(\bigcup_{i=1}^{\infty} \{X = x_i\}\right) = 1.$$

Può essere utile rappresentare la densità discreta in forma grafica ponendo i valori x_i in ascissa e $p(x_i)$ in ordinata. Per esempio, se la densità discreta di X è

$$p(0) = \frac{1}{4}, \qquad p(1) = \frac{1}{2}, \qquad p(2) = \frac{1}{4},$$

graficamente si ha



La densità discreta consente di calcolare la probabilità che la variabile aleatoria assuma valori in un sottoinsieme qualsiasi B di $\mathbb{R}$; ad esempio nell'intervallo $[a, b]$:

$$P(X \in B) = \sum_{k:\, x_k \in B} p(x_k), \qquad P(a \leq X \leq b) = \sum_{k:\, a \leq x_k \leq b} p(x_k).$$

Esempio. Sia X la variabile aleatoria discreta che descrive il numero di bit pari a $\mathbf{1}$ in un vettore booleano, di lunghezza n, scelto a caso.

(a) Calcolare $P(X=k)$, per $k=0,1,\ldots,n$.

(b) Verificare che $\sum_{k=0}^{n} P(X=k)=1$.

(c) Calcolare $P(X \geq 1)$.

Soluzione. (a) Dividendo il numero di vettori booleani con k bit pari a $\mathbf{1}$ per il numero di vettori booleani di lunghezza n, si trae

$$P(X=k)=\frac{\binom{n}{k}}{2^n}, \qquad k=0,1,\ldots,n.$$

(b) Pertanto,

$$\sum_{k=0}^{n} P(X=k)=\frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}=\frac{1}{2^n}\cdot 2^n=1.$$

(c) Si ha infine

$$P(X \geq 1)=1-P(X=0)=1-\frac{1}{2^n}.$$

Esempio. Sia $p(k) = c\,\dfrac{\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$, dove λ è una costante positiva. Si calcoli
(a) $P(X = 0)$,
(b) $P(X > 2)$.

Soluzione. 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 $p(k) = e^{-\lambda}\,\dfrac{\lambda^k}{k!}$, $k \geq 0$, e pertanto

(a) $P(X = 0) = p(0) = e^{-\lambda}\,\lambda^0/0! = e^{-\lambda}$,
(b) $P(X > 2) = 1 - P(0 \leq X \leq 2) = 1 - p(0) - p(1) - p(2) = 1 - e^{-\lambda}\left(1 + \lambda + \dfrac{\lambda^2}{2}\right)$.

Esempio. Determinare la densità della variabile aleatoria discreta X, che descrive il massimo che si ottiene lanciando 2 dadi.

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 $p(k) = P(X = k)$ 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 poiché

$$\boxed{\sum_{k=1}^{n} k = \frac{n(n+1)}{2}}$$

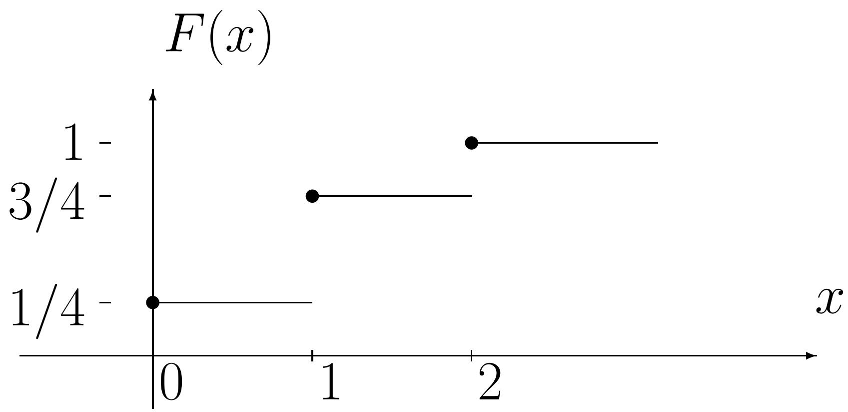
si trae

$$\sum_{k=1}^{6} p(k) = \sum_{k=1}^{6} \frac{2k-1}{36} = \frac{1}{18}\sum_{k=1}^{6} k - \frac{1}{6} = \frac{1}{18}\,\frac{6\cdot 7}{2} - \frac{1}{6} = 1.$$

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 $p(x_i)$. Quindi $F(x)$ può essere così espressa in termini della densità discreta:

$$F(x) = \sum_{k:\, x_k \le x} p(x_k) = \begin{cases} 0, & x < x_1 \\ p(x_1), & x_1 \le x < x_2 \\ p(x_1) + p(x_2), & x_2 \le x < x_3 \\ p(x_1) + p(x_2) + p(x_3), & x_3 \le x < x_4 \\ \vdots \end{cases}$$

Ad esempio, se X ha densità discreta $p(0) = \frac{1}{4}$, $p(1) = \frac{1}{2}$, $p(2) = \frac{1}{4}$, si ha



$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{1}{4} & 0 \le x < 1 \\ \dfrac{3}{4} & 1 \le x < 2 \\ 1 & x \ge 2. \end{cases}$$

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 euro se non esce mai testa. Indicando con 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 $p(k) = P(X = k)$, risulta

$$p(1) = P(T_1) = \frac{1}{2}, \qquad p(2) = P(\overline{T_1} \cap T_2) = \frac{1}{4},$$

$$p(3) = P(\overline{T_1} \cap \overline{T_2} \cap T_3) = \frac{1}{8}, \qquad p(-c) = P(\overline{T_1} \cap \overline{T_2} \cap \overline{T_3}) = \frac{1}{8},$$

Pertanto: $F(x) = 0$ per $x < -c$, $F(x) = 1/8$ per $-c \leq x < 1$, $F(x) = 5/8$ per $1 \leq x < 2$, $F(x) = 7/8$ per $2 \leq x < 3$, $F(x) = 1$ per $x \geq 3$. 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 la 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 $p(k) = \frac{1}{6}$ per $k = 1, 2, \ldots, 6$ otteniamo

$$E(X) = \sum_{k=1}^{6} k \cdot p(k) = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2} = 3{,}5.$$

Esempio. Dato un evento A, definiamo la funzione indicatrice I_A di A come

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

Si calcoli $E(I_A)$.

Soluzione. Poiché $p(1) = P(I_A = 1) = P(A)$ e $p(0) = P(I_A = 0) = P(\overline{A})$, abbiamo

$$E(I_A) = 0 \cdot p(0) + 1 \cdot p(1) = p(1) = P(A).$$

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, D_1 e D_2, ed è libero di scegliere in che ordine rispondere. Se risponde per prima alla domanda 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 D_i, $i = 1, 2$. Se la probabilità di $E_i = \{\text{conosce la risposta di } D_i\}$ 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 D_i $(i = 1, 2)$; si ha

$$\begin{array}{ll} P(X_1 = 0) = 1 - P_1 & \qquad P(X_2 = 0) = 1 - P_2 \\ P(X_1 = V_1) = P_1\,(1 - P_2) & \qquad P(X_2 = V_2) = P_2\,(1 - P_1) \\ P(X_1 = V_1 + V_2) = P_1\,P_2 & \qquad P(X_2 = V_1 + V_2) = P_1\,P_2 \end{array}$$

con guadagno atteso $E[X_1] = V_1\,P_1\,(1 - P_2) + (V_1 + V_2)\,P_1\,P_2$ se risponde prima a D_1, mentre risulta $E[X_2] = V_2\,P_2\,(1 - P_1) + (V_1 + V_2)\,P_1\,P_2$ se risponde prima a D_2.
È più vantaggioso rispondere prima a D_1 se $V_1\,P_1\,(1 - P_2) \geq V_2\,P_2\,(1 - P_1)$, ovvero se

$$\frac{V_1\,P_1}{1 - P_1} \geq \frac{V_2\,P_2}{1 - P_2}.$$

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. 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 $P(X = 36) = \frac{36}{120}$, $P(X = 40) = \frac{40}{120}$, $P(X = 44) = \frac{44}{120}$ e quindi

$$E(X) = 36 \cdot \frac{36}{120} + 40 \cdot \frac{40}{120} + 44 \cdot \frac{44}{120} = \frac{4832}{120} = 40{,}2667.$$

Il numero medio di studenti presenti su un autobus è $120/3 = 40$, che quindi è minore di $E(X)$. Questo fenomeno si verifica perché: più studenti sono presenti su un singolo autobus e più sarà probabile che lo studente scelto a caso provenga proprio da quello. Così si assegna peso maggiore agli autobus con più studenti.

Più in generale, se vi sono k autobus con $n_1, n_2, \ldots, n_k$ studenti, si ha

$$E(X) = \sum_{i=1}^{k} n_i \cdot \frac{n_i}{\sum_{j=1}^{k} n_j} = \frac{\sum_{i=1}^{k} n_i^2}{\sum_{j=1}^{k} n_j} \qquad \left(\geq \frac{\sum_{i=1}^{k} n_i}{k} \right)$$

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. Descriviamo il miglioramento in percentuale prodotto dal farmaco sperimentale mediante una variabile aleatoria X discreta tale che

$$P(X = -1) = 0{,}75 \qquad P(X = 5) = 0{,}15 \qquad P(X = 50) = 0{,}10$$

Il miglioramento medio prodotto dal farmaco è del 5%, essendo

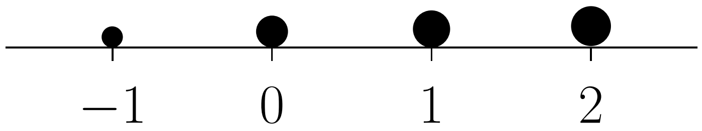
$$E(X) = \sum_x x \cdot P(X = x) = -1 \cdot 0{,}75 + 5 \cdot 0{,}15 + 50 \cdot 0{,}10 = 5.$$

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.

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, $E(X) = \sum_{x=-1}^{2} x \cdot p(x) = -1 \cdot 0{,}10 + 0 \cdot 0{,}25 + 1 \cdot 0{,}30 + 2 \cdot 0{,}35 = 0{,}9.$

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

$$p(j)=P(X=j)=\frac{\binom{j-1}{0}\binom{1}{1}\binom{n-j}{k-1}}{\binom{n}{k}}=\frac{\binom{n-j}{k-1}}{\binom{n}{k}}, \qquad j=1,2,\ldots,n-k+1.$$

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

$$
\begin{aligned}
E(X) &= \sum_j j\,p(j) = \frac{1}{\binom{n}{k}} \sum_{j=1}^{n-k+1} j \binom{n-j}{k-1} = \frac{1}{\binom{n}{k}} \sum_{j=1}^{n-k+1} \sum_{h=1}^{j} \binom{n-j}{k-1} \\
&= \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \sum_{j=h}^{n-k+1} \binom{n-j}{k-1} = \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \sum_{r=k-1}^{n-h} \binom{r}{k-1},
\end{aligned}
$$

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

$$p(j) = P(X = j) = \frac{\binom{n-j}{0}}{\binom{n}{1}} = \frac{1}{n}, \qquad j = 1, 2, \dots, n.$$

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 n 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

$$p(j) = P(Y = j) = p(1-p)^{j-1}, \qquad j = 1, 2, \ldots,$$

dove $p = k/n$. Si ha pertanto

$$E(Y) = p \sum_{j=1}^{\infty} j(1-p)^{j-1} = p \sum_{j=1}^{\infty} \sum_{k=1}^{j} (1-p)^{j-1} = p \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} (1-p)^{j-1}.$$

Ponendo $n = j - k$ si ha

$$E(Y) = p \sum_{k=1}^{\infty} (1-p)^{k-1} \sum_{n=0}^{\infty} (1-p)^{n} = p \, \frac{1}{p} \, \frac{1}{p} = \frac{1}{p} = \frac{n}{k}.$$

Notiamo che, per $k \le n$, risulta $E(X) = \dfrac{n+1}{k+1} \le \dfrac{n}{k} = E(Y)$.

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 $g(X)$.

Essendo $g(X)$ stessa una variabile aleatoria discreta, avrà una densità discreta che possiamo determinare conoscendo quella di X. Ricavata la densità discreta di $g(X)$ possiamo calcolare $E[g(X)]$ 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

$$P(Y=1)=P(X^2=1)=P(X=\pm 1)=P(X=1)+P(X=-1)=0{,}5$$

$$P(Y=0)=P(X^2=0)=P(X=0)=0{,}5$$

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 $0{,}5=E(X^2)\neq [E(X)]^2=[-1\cdot 0{,}2+0\cdot 0{,}5+1\cdot 0{,}3]^2=[0{,}1]^2=0{,}01$.

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 $E(X^2)=\displaystyle\sum_i x_i^2\,p(x_i)=(-1)^2\cdot 0{,}2+0^2\cdot 0{,}5+1^2\cdot 0{,}3=0{,}5.$

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

$$E[g(X)]=\sum_i g(x_i)\,p(x_i)=\sum_{k=1}^{6}\min\{k,4\}\,p(k)=\frac{1}{6}\,(1+2+3+4+4+4)=3.$$

Esempio. Sia X la variabile aleatoria tale che $P(X=0)=1-p$ e $P(X=1)=p$, $0\le p\le 1$. Calcolare il minimo di $\psi(a):=E[(X-a)^2]$, per $a\in\mathbb{R}$.

Soluzione. Risulta $\psi(a)=E[(X-a)^2]=\sum_{k=0}^{1}(k-a)^2\,p(k)=a^2\,(1-p)+(1-a)^2\,p$ e quindi $\psi'(a)=2a(1-p)-2(1-a)p=0$ per $a=p$. Quindi il minimo di $\psi(a)$ è p.

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

$$E[g(X)]=\sum_i (a\,x_i+b)\,p(x_i)=a\sum_i x_i\,p(x_i)+b\sum_i p(x_i)=a\,E(X)+b,$$

avendo fatto uso dell'identità $\sum_i p(x_i)=1$.

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 valore atteso di X è anche il momento di ordine 1 di X.

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.

Soluzione. Facendo uso del teorema del binomio si ha

$$E[Y^n]=E[(aX+b)^n]=E\left[\sum_{k=0}^{n}\binom{n}{k}a^kX^kb^{n-k}\right]$$

e pertanto, per la proprietà di linearità,

$$E[Y^n]=\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}E[X^k].$$

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, pari a 0, ma è presente maggiore dispersione nei valori di Y piuttosto che in quelli di W e nei valori di Z rispetto a 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 $E(X) = \mu$; la varianza di X, denotata con $\mathrm{Var}(X)$, è così definita:

$$\mathrm{Var}(X) = E[(X-\mu)^2] = \sum_{x:\, p(x)>0} (x-\mu)^2\, p(x).$$

Una formula alternativa per la varianza 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.

Esempio. Calcolare $\mathrm{Var}(X)$, dove X rappresenta l'esito del lancio di un dado.
Soluzione. Essendo $p(k) = \frac{1}{6}$ per $k = 1, 2, \ldots, 6$ e $E(X) = \frac{7}{2}$ otteniamo

$$\mathrm{Var}(X) = E(X^2) - \mu^2 = \sum_{k=1}^{6} k^2 \cdot p(k) - \left(\frac{7}{2}\right)^2 = \frac{1}{6} \cdot 91 - \frac{49}{4} = \frac{35}{12} \approx 2{,}9167.$$

Una variabile aleatoria X si dice *degenere* se esiste un reale x_0 tale che $P(X = x_0) = 1$. In tal caso la funzione di distribuzione di X è data da

$$F(x) = \begin{cases} 0, & x < x_0 \\ 1, & x \geq x_0. \end{cases}$$

Se X è una variabile aleatoria degenere tale che $P(X = x_0) = 1$ allora è facile ricavare che $E[X] = x_0$ e $\mathrm{Var}(X) = 0$.

Notiamo che 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.

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 $E(a\,X+b)=a\,E(X)+b$, si ha

$$\begin{aligned}\mathrm{Var}(a\,X+b) &= E[\{a\,X+b-E(a\,X+b)\}^2]=E[\{a\,X+b-a\,E(X)-b\}^2] \\ &= a^2\,E[\{X-E(X)\}^2]=a^2\,\mathrm{Var}(X).\end{aligned}$$

Si noti che risulta $\mathrm{Var}(a\,X+b)\geq \mathrm{Var}(X)$ se e solo se $|a|\geq 1$.

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, è detta *deviazione standard* di X. Cioè

$$\sigma_X=\sqrt{\mathrm{Var}(X)}=\sqrt{E[(X-\mu)^2]}.$$

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

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

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 in forma compatta:

$$p(x) = p^x \, (1 - p)^{1-x}, \qquad x = 0, 1.$$

Notiamo che per una variabile aleatoria di Bernoulli X la funzione di distribuzione è

$$F(x) = \sum_{k \leq x} p(k) = \sum_{k \leq x} p^k \, (1-p)^{1-k} = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

e il momento di ordine n è dato da

$$E[X^n] = \sum_{x=0}^{1} x^n \, p(x) = \sum_{x=0}^{1} x^n \, p^x \, (1-p)^{1-x} = p.$$

Ne segue che valore atteso e varianza sono

$$E[X] = p$$

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = p(1-p).$$

Notiamo che $\mathrm{Var}(X) = 0$ se $p = 0$ e $p = 1$, mentre $\mathrm{Var}(X)$ è massima per $p = 1/2$.

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

$$p(x) = \binom{n}{x} p^x\,(1-p)^{n-x}, \qquad x = 0, 1, \ldots, n \qquad (0 \le p \le 1).$$

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 che descrive il numero di volte che esce testa in 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

$$P(X=0) = \binom{5}{0}\left(\frac{18}{37}\right)^0\left(1-\frac{18}{37}\right)^5 \approx 0{,}0357$$

$$\begin{aligned} P(X\geq 2) &= 1-P(X=0)-P(X=1) \\ &= 1-\binom{5}{0}\left(\frac{18}{37}\right)^0\left(1-\frac{18}{37}\right)^5-\binom{5}{1}\left(\frac{18}{37}\right)^1\left(1-\frac{18}{37}\right)^4 \approx 0{,}7952. \end{aligned}$$

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	0,3061
$p(4)$	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:

$$E(X^k) = \sum_{x\colon p(x)>0} x^k \cdot p(x) = \sum_{i=0}^{n} i^k \cdot \binom{n}{i} p^i\,(1-p)^{n-i} = \sum_{i=1}^{n} i^k \cdot \binom{n}{i} p^i\,(1-p)^{n-i}.$$

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 $n\,p\,(1-p)$.

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

$$p(k) = P(X = k) = \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k}, \qquad 0 \leq k \leq n.$$

Quindi, si ha

$$E[X] = n\,p = n\,\frac{1}{n} = 1, \qquad Var[X] = n\,p\,(1-p) = n\,\frac{1}{n}\left(1 - \frac{1}{n}\right) = 1 - \frac{1}{n}.$$

Ad esempio, per $n = 5$ risulta:

$p(0)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$
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 $\mathrm{Var}(X) = n\,p\,(1-p)$. Pertanto, ricordando la proprietà di linearità del valore atteso e la proprietà della varianza, si ha

$$E(F_n) = E\left(\frac{X}{n}\right) = \frac{1}{n}\,E(X) = \frac{1}{n}\,n\,p = p,$$

$$\mathrm{Var}(F_n) = \mathrm{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2}\,\mathrm{Var}(X) = \frac{1}{n^2}\,n\,p\,(1-p) = \frac{p\,(1-p)}{n}.$$

Osserviamo, in particolare, che $E(F_n)$ è costante in n, mentre $\mathrm{Var}(F_n)$ è decrescente e tende a 0 quando n tende a $+\infty$.

Si noti, inoltre, che F_n è una variabile aleatoria discreta che assume valori $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$.

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

$$\frac{P(X=k)}{P(X=k-1)} = \frac{\dfrac{n!}{k!\,(n-k)!}\,p^k\,(1-p)^{n-k}}{\dfrac{n!}{(k-1)!\,(n-k+1)!}\,p^{k-1}\,(1-p)^{n-k+1}} = \frac{(n-k+1)\,p}{k\,(1-p)}.$$

Quindi $P(X = k) \geq P(X = k-1)$ se e solo se

$$(n-k+1)\,p \geq k\,(1-p)$$

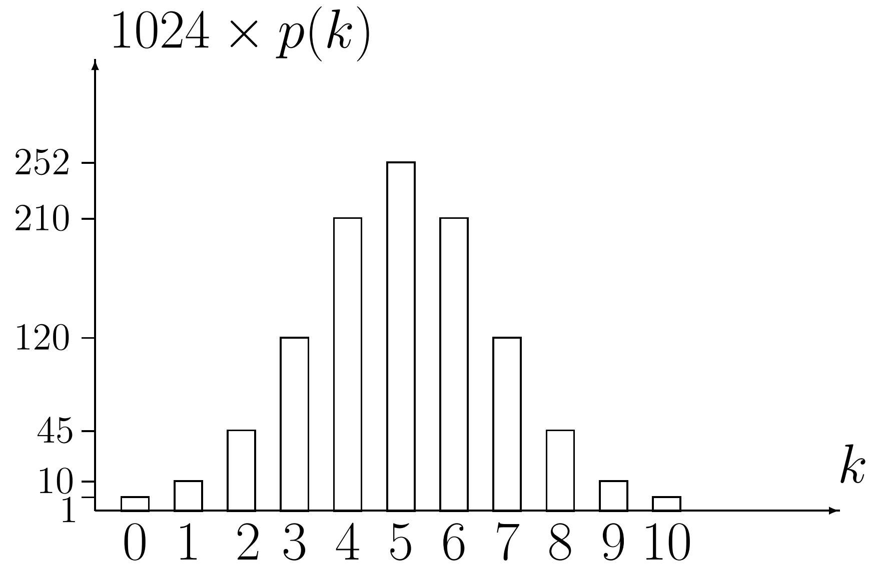
ossia

$$k \leq (n+1)p.$$

Nel caso di una variabile binomiale di parametri $(10, \frac{1}{2})$ la densità discreta

$$p(k) = \binom{10}{k} \left(\frac{1}{2}\right)^{10}$$

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

$$P(X = k) = P(Y = n - k), \qquad k = 0, 1, \ldots, n.$$

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

$$\sum_{k=0}^{\infty} p(k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda}\, e^{\lambda} = 1,$$

essendo

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}.$$

La variabile aleatoria di Poisson può essere utilizzata 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. Sia $\lambda = n\,p$; allora

$$\begin{aligned} P(Y=k) &= \frac{n!}{k!\,(n-k)!}\,p^k\,(1-p)^{n-k} \\ &= \frac{(n)_k}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!}\cdot\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}\cdot\frac{(1-\lambda/n)^n}{(1-\lambda/n)^k}. \end{aligned}$$

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

$$P(Y=k)=\binom{n}{k}p^k\,(1-p)^{n-k}\approx\frac{\lambda^k}{k!}\,e^{-\lambda}\qquad(k=0,1,\ldots).$$

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 $\lambda=5$ e quindi

$$P(Y=k)=\binom{90}{k}\left(\frac{1}{18}\right)^k\left(1-\frac{1}{18}\right)^{90-k}\qquad\approx\frac{(5)^k}{k!}\,e^{-5};$$

$$P(Y=0)=\left(1-\frac{1}{18}\right)^{90}=0{,}0058\qquad\approx e^{-5}=0{,}0067.$$

Se $n=900$ e $p=1/180$ si ha ancora $\lambda=5$ 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 $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-1/2} \approx 0{,}393$.

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.
Soluzione. La probabilità desiderata è $\binom{10}{0}(0{,}1)^0(0{,}9)^{10} + \binom{10}{1}(0{,}1)^1(0{,}9)^9 = 0{,}7361$, mentre l'approssimazione di Poisson fornisce $\frac{1^0}{0!}e^{-1} + \frac{1^1}{1!}e^{-1} = 2\,e^{-1} = 0{,}7358$.

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. Pensiamo che le richieste di stampa giungono da un grande numero n di utenti, ognuno dei quali ha probabilità $3{,}2/n$ di fare una richiesta al minuto; allora, per l'approssimazione di Poisson, $P(X \leq 2) = e^{-3{,}2} + 3{,}2\,e^{-3{,}2} + \frac{(3{,}2)^2}{2}\,e^{-3{,}2} \approx 0{,}3799$.

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

$$P(X = n) = (1 - p)^{n-1} p, \qquad n = 1, 2, \ldots$$

Invero, si ha $X = n$ se le prime $n-1$ prove siano state un insuccesso e l'n-esima prova un successo. La formula si ricava anche per l'ipotesi di indipendenza tra le prove.

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

$$\begin{aligned}
\text{(a)} \qquad P(X=n) &= (1-p)^{n-1}\,p; \\
\text{(b)} \qquad P(X>k) &= p\sum_{n=k+1}^{\infty}(1-p)^{n-1} = p\,(1-p)^k\sum_{j=0}^{\infty}(1-p)^j \\
&= p\,(1-p)^k\,\frac{1}{1-(1-p)} = (1-p)^k,
\end{aligned}$$

avendo posto $j = n-k-1$ e avendo ricordato che, per $-1 < c < 1$,

$$\sum_{j=0}^{\infty} c^j = \frac{1}{1-c}.$$

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$^{\text{o}}$ lancio sapendo che nei primi 3 lanci non esce testa.

Soluzione. Essendo $p = 1/2$, per la proprietà di assenza di memoria si ha:

$$P(X > 5 | X > 3) = P(X > 2) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

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:

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

In particolare, se si ripete una serie di prove indipendenti, ognuna con probabilità di successo p, fino a che non si verifica il primo successo, allora il numero atteso di prove è uguale a $1/p$. Ad esempio, per un fissato valore nel lancio di un dado è $p = \frac{1}{6}$ e quindi $E(X) = 6$; per il singolo estratto nel gioco del lotto è $p = \frac{1}{18}$ e quindi $E(X) = 18$.

Esempio. Calcolare la probabilità che una variabile aleatoria geometrica X assuma valore non maggiore del suo valore atteso, se $1/p$ è intero. Valutare tale probabilità quando $p \to 1^-$ e quando $p \to 0^+$.

Soluzione. Ricordando che $E(X) = 1/p$ e che $P(X \le k) = 1 - (1-p)^k$ si ha

$$P\{X \le E(X)\} = 1 - (1-p)^{1/p}.$$

Pertanto

$$\lim_{p\to 1^-} P\{X \le E(X)\} = 1 - \lim_{p\to 1^-} (1-p)^{1/p} = 1,$$

$$\lim_{p\to 0^+} P\{X \le E(X)\} = 1 - \lim_{p\to 0^+} (1-p)^{1/p} = 1 - \lim_{x\to\infty} \left(1 - \frac{1}{x}\right)^x = 1 - e^{-1} \approx 0{,}632.$$

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

$$P(X=k)=\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} \qquad k=0,1,\ldots,n.$$

Una variabile aleatoria dotata di tale densità discreta per opportuni valori n, m, N è detta variabile aleatoria *ipergeometrica*.

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:

$$p(k) = P(X=k) = \frac{\binom{5}{k}\binom{3}{4-k}}{\binom{8}{4}}.$$

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. 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

$$\begin{aligned} P(X=0) &= P(X=0|B)\,P(B) + P(X=0|\overline{B})\,P(\overline{B}) \\ &= \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} \cdot \frac{3}{10} + \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} \cdot \frac{7}{10} = \frac{6\cdot 5\cdot 4}{10\cdot 9\cdot 8}\cdot\frac{3}{10} + \frac{9\cdot 8\cdot 7}{10\cdot 9\cdot 8}\cdot\frac{7}{10} \\ &= 0{,}05 + 0{,}49 = 0{,}54 \end{aligned}$$

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 la legge di X è approssimata da una legge binomiale:

$$P(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} = \frac{\frac{(m)_k}{k!} \cdot \frac{(N-m)_{n-k}}{(n-k)!}}{\frac{(N)_n}{n!}} = \binom{n}{k} \frac{(m)_k \cdot (N-m)_{n-k}}{(N)_n}$$

Per m grande risulta $(m)_k = m(m-1)\cdots(m-k-1) \approx m^k$ e pertanto

$$P(X = k) \approx \binom{n}{k} \frac{m^k\,(N-m)^{n-k}}{N^n} = \binom{n}{k} \left(\frac{m}{N}\right)^k \left(\frac{N-m}{N}\right)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

con $p = \dfrac{m}{N}$, e con m e N 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 di X è data da: $F(x) = 0$ per $x < 1$, $F(x) = k/N$ per $k \leq x < k + 1$ $(k = 1, 2, \ldots, n - 1)$, $F(x) = 1$ 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

$$E(X) = \sum_{k=1}^{N} k\, P(X = k) = \frac{1}{N} \sum_{k=1}^{N} k = \frac{N+1}{2}.$$

Analogamente, poiché $\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}$, 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

$$\begin{aligned} \mathrm{Var}(X) &= \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} \\ &= \frac{N+1}{12}\,[2(2N+1) - 3(N+1)] = \frac{N+1}{12}\,(N-1) \\ &= \frac{N^2-1}{12}. \end{aligned}$$

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 $\lim\limits_{h\to 0} f(h)/h = 0$. [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 $0 \le t_0 < t_1 < t_2 < \ldots < t_n$ e degli interi non negativi $k_1, k_2, \ldots, k_n$ gli eventi $E_1, E_2, \ldots, E_n$ sono indipendenti.

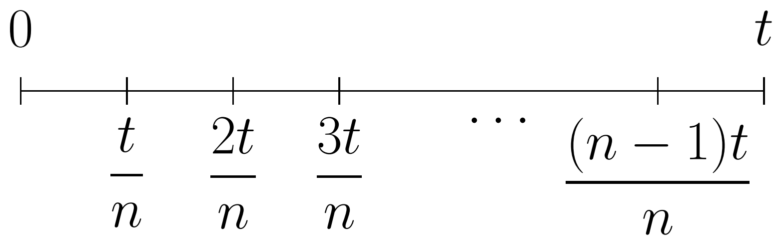
Dalle ipotesi 1 e 2 segue la seguente condizione:

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,
$$P(B) \le t\left[\frac{o(t/n)}{t/n}\right].$$

$\forall t > 0$ 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

$$\begin{aligned} P(A) &= P\{k \text{ degli } n \text{ sottointervalli contengono 1 evento} \\ &\qquad \text{e gli altri } n-k \text{ ne contengono } 0\} \\ &= \binom{n}{k} p^k (1-p)^{n-k}, \end{aligned}$$

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 $p = \frac{\lambda t}{n} + o\left(\frac{\lambda t}{n}\right)$. 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 $\lim_{n\to\infty} p = 0$ 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.a) 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.b) 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.c) 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.d) 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.e) 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.