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

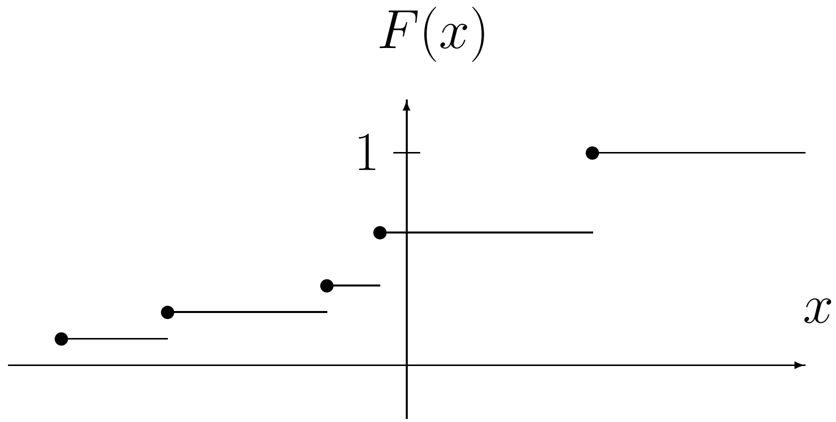
$$\lim_{n\to\infty} P(X \leq b_n) = P(X < \infty) = 1.$$

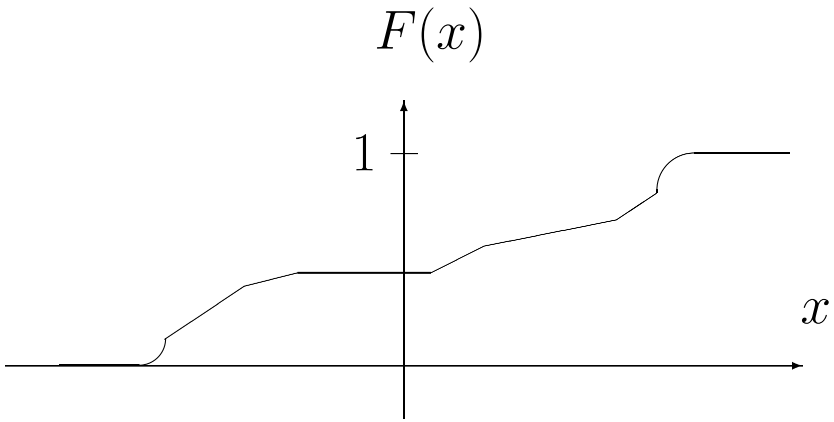
La Proprietà 3 si dimostra in modo analogo alla Proprietà 2, ed è lasciato come esercizio. Per dimostrare la Proprietà 4 notiamo che se x_n, $n \geq 1$, è una successione decrescente che converge a $x \in \mathbb{R}$, allora $\{X \leq x_n\}$, $n \geq 1$, costituisce una successione di eventi decrescente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

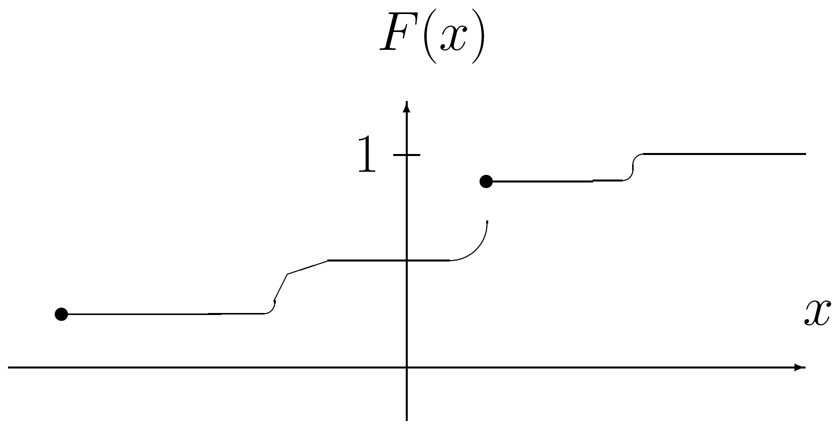
$$\lim_{n\to\infty} P(X \leq x_n) = P(X \leq x).$$

Proprietà di $F(x) = P(X \leq x)$:

1. $F(x)$ è non decrescente.
2. $\lim\limits_{x \to \infty} F(x) = 1$.
3. $\lim\limits_{x \to -\infty} F(x) = 0$.
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)

$$F(x) = \begin{cases} 0, & x \leq -1 \\ \frac{x+1}{2}, & -1 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

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

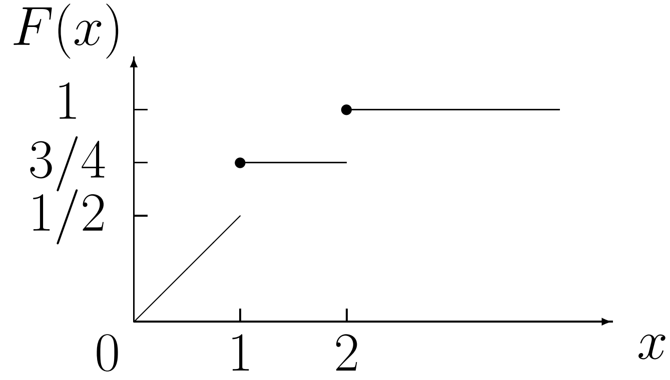
$$F(x^-) := \lim_{h \to 0^-} F(x+h) \leq \lim_{h \to 0^+} F(x+h) = F(x).$$

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

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



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

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

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

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

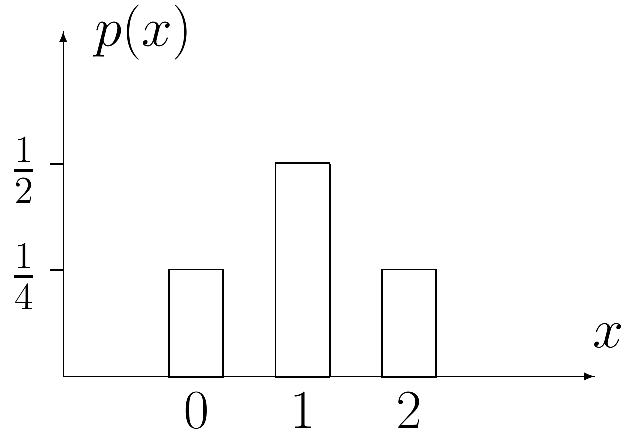
Poiché X deve assumere almeno uno dei valori x_i, abbiamo che

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

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 **1** in un vettore booleano, di lunghezza n, scelto a caso.
(a) Calcolare $P(X = k)$, per $k = 0, 1, \dots, n$.
(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 **1** per il numero di vettori booleani di lunghezza n, si trae

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

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

Soluzione. Per determinare c, imponendo che sia $\sum_{k=0}^{\infty} p(k) = 1$ abbiamo

$$1 = c \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = c\, e^{\lambda} \qquad \Longrightarrow \qquad c = e^{-\lambda},$$

avendo ricordato che $\boxed{\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x}$. Si ha quindi $p(k) = e^{-\lambda}\,\frac{\lambda^k}{k!}$, $k \geq 0$, e pertanto

(a) $P(X = 0) = p(0) = e^{-\lambda}\,\lambda^0/0! = e^{-\lambda}$,

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

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

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