Variabili aleatorie

Variabili aleatorie e variabili statistiche

- Nelle prime lezioni, abbiamo visto il concetto di "variabile statistica":
- Un "oggetto" o evento del mondo reale veniva associato a una certa "proprietà", qualitativa o quantitativa a seconda della scala di misura adottata;
- Quello di variabile aleatoria, anche se più formalizzato matematicamente, è un concetto analogo.

Che cos'è una variabile aleatoria?

Ω: spazio campione (insieme di tutti i risultati possibili dell'esperimento).

Gli eventi elementari, ω, dello spazio campione Ω sono entità concrete come persone, molecole, monete, dadi, carte, ..., e pertanto possiedono vari "attributi" misurabili.

Che cos'è una variabile aleatoria?

Consideriamo, ad esempio, una "popolazione" composta da *n* persone:

$$\Omega = \{ \omega_1, \omega_2, \omega_3, \dots \omega_n \}$$

Che cos'è una variabile aleatoria?

Esperimento aleatorio: estrazione casuale di una persona da questa popolazione.

Variabile aleatoria "punteggio in matematica della popolazione"

$$\omega \in \Omega \quad \rightarrow \; M\,(\omega),$$

dove $M\,(\omega)$ denota il "punteggio in matematica di ω".

Che cos'è una variabile aleatoria?

Allo stesso modo, possiamo indicare l'altezza, il peso e il reddito di ω con le funzioni:

$$\omega \rightarrow \; A\,(\omega)$$

$$\omega \rightarrow \; P\,(\omega)$$

$$\omega \rightarrow \; R\,(\omega)$$

Che cos'è una variabile aleatoria?

I "numeri" $Q\ (\omega)$, $A\ (\omega)$, $P\ (\omega)$, $R\ (\omega)$ che abbiamo associato ai punti campione ω_i sono chiamati variabili aleatorie (o variabili casuali).

Definizione di V.A.

Una variabile aleatoria X è una funzione numerica di ω avente come **dominio Ω** e come **codominio** (= **immagine**) l'insieme dei **numeri reali**:

$$\omega \in \Omega$$
$$X(\omega) : \Omega \to \Re$$

Definizione

Una variabile aleatoria è dunque un numero che viene assegnato, mediante una determinata regola, a ciascun punto dello spazio campione, ovvero a ciascuno degli esiti possibili di un esperimento aleatorio.

Perchè una variabile si dice "aleatoria"?

Il termine "aleatorio" allude al fatto che ci occupiamo degli esiti possibili di un esperimento aleatorio, ovvero, di un esperimento il cui esito è incerto prima che dell'esecuzione dell'esperimento stesso.
Una volta che l'esperimento viene eseguito, il valore $X\ (\omega)$ risulta completamente determinato (si parla allora di realizzazione di una variabile aleatoria).

Variabili aleatorie discrete e continue

Una variabile aleatoria si dice ***discreta*** se può assumere un numero finito, o al più infinito numerabile, di valori;
si dice ***continua*** se può assumere tutti gli infiniti valori dell'asse reale **R**, oppure di un suo intervallo [a,b]

Variabili aleatorie discrete e continue

- Le variabili aleatorie sono DISCRETE se producono risposte numeriche che derivano da un processo di conteggio. Ad es. "Il numero dei componenti la famiglia", "il numero delle stanze diun'abitazione", ecc.
- Le variabili aleatorie sono CONTINUE se generano risposte che derivano da un processo di misurazione. Ad es. "l'altezza", "il reddito", "il fatturato", ecc.

Esempio

Un esperimento aleatorio consiste nel lancio di due monete. Sia *Y* il numero di volte in cui l'esito "testa" viene osservato in ciascuna prova dell'esperimento. Vogliamo identificare i punti campione di Ω, assegnare un valore *y* a ciascun punto campione e identificare i punti campione associati a ciascuno dei valori che *Y* può assumere.

Ciascuna prova dell'esperimento può essere indicata da una coppia ordinata di simboli che identificano l'esito del lancio della prima e della seconda moneta. Ad esempio, *TC* indicherà l'esito "testa" per il lancio della prima moneta e l'esito "croce" per il lancio della seconda moneta.

I quattro punti campione in Ω sono:

$$E_1 : \{TT\},\ E_2 : \{TC\},\ E_3 : \{CT\},\ E_4 : \{CC\}$$

Il valore da assegnare a *Y* in corrispondenza di ciascun punto campione di Ω dipende dal numero di volte in cui viene osservato l'esito "testa".

Pertanto, possiamo costruire la nostra v.a. nel modo che segue:

Esempio

$\omega_1 = \{TT\}$, $\omega_2 = \{TC\}$, $\omega_3 = \{CT\}$, $\omega_4 = \{CC\}$

$\Downarrow$ $\Downarrow$ $\Downarrow$ $\Downarrow$

$Y = 2$ $Y = 1$ $Y = 1$ $Y = 0$

Esempio

La variabile aleatoria Y può assumere tre valori (Y = 0, 1, 2), e ciascuno di tali valori corrisponde a un evento complesso. Tali eventi complessi sono descritti, rispettivamente, dai seguenti insiemi di punti campione:

$$\{Y = 0\} = \{E_4\}$$

$$\{Y = 1\} = \{E_2 oppure E_3\}$$

$$\{Y = 2\} = \{E_1\}$$

Osservazione

- Si noti che
- Mentre:

$$\omega \in \Omega$$

B_Ω = insieme delle parti di Ω

- Si ha che:

$$E_i \in B_\Omega$$

Notazione in una v.a.

Solitamente le variabili aleatorie sono indicate con le lettere maiuscole, mentre gli specifici valori che assumono vengono indicati dalle lettere minuscole.

Osservazione importante:

Osserviamo che, mentre la regola da adottare per la "creazione" di una v.a. è arbitraria, in quanto dipende da ciò che vogliamo che la v.a. interpreti o rappresenti, lo stesso **non è vero** per la determinazione della distribuzione di probabilità P[X=x] della stessa v.a., in quanto quest'ultima è legata alle probabilità degli eventi elementari P[ω] e, di conseguenza, deve rispettare i tre assiomi di Kolmogorov e tutte le regole probabilistiche che da essi conseguono.

Distribuzione di probabilità di una variabile aleatoria discreta

La probabilità che la variabile aleatoria discreta Y assuma il valore y, $P(Y = y)$, è definita come la somma delle probabilità di tutti i punti campione in Ω a cui viene assegnato il valore y.

Questa proprietà è richiesta dal terzo assioma di Kolmogorov (assioma di additività).

$P\ (Y = y)$ si può indicare sinteticamente con $p(y)$.

Nel nostro esempio:

$\omega_1 = \{TT\}$, $\omega_2 = \{TC\}$, $\omega_3 = \{CT\}$, $\omega_4 = \{CC\}$

$Y = 2$ $\quad Y = 1$ $\quad Y = 1$ $\quad Y = 0$

$$P(Y = 0) = P(\omega_4)$$

$$P(Y = 1) = P(\omega_2) + P(\omega_3)$$

$$P(Y = 2) = P(\omega_1)$$

$$P(Y = 0) = P(\omega_4)$$

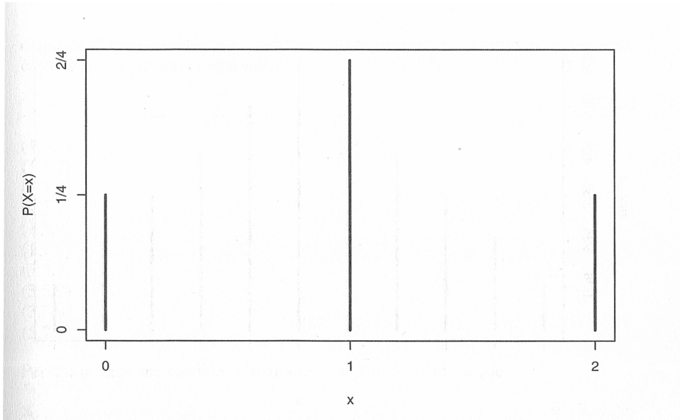
$$P(Y = 1) = P(\omega_2) + P(\omega_3)$$

$$P(Y = 2) = P(\omega_1)$$

Poiché i punti dello spazio campione ω_i hanno tutti la stessa probabilità, la distribuzione di probabilità di *Y* sarà:

y	***p(y)***
0	**1 / 4**
1	**2 / 4**
2	**1 / 4**

In termini grafici:



Distribuzione di probabilità di una variabile aleatoria discreta

La corrispondenza tra i numeri $\{y\}$ ed i rispettivi valori di probabilità $\{P\ (Y = y)\}$ definisce la distribuzione di probabilità (o funzione di probabilità) della variabile aleatoria discreta *Y.*

Proprietà di una distribuzione di probabilità discreta

Le probabilità $p(y) = P(Y = y)$ godono delle seguenti proprietà:

$$(i) \qquad P(Y = y_i) \geq 0$$

$$(ii) \qquad \sum_{\mathrm{i}} P(Y = y_i) = 1$$

$$\forall i = 1, 2, .., n$$

Osservazione

- Le proprietà (i) e (ii) derivano anch'esse dagli assiomi di Kolmogorov:
- P maggiore o uguale a 0;
- La probabilità dell'evento certo (somma delle probabilità dei singoli eventi) deve essere pari a 1.

Definizione:

- L'insieme dei valori numerici che possono essere assunti dalla v.a. Y:
- $S = \{y_1, y_2, \dots, y_i, \dots y_n\}$
- Tali da dare luogo a una probabilità positiva, è detto **supporto** della v.a. Y.

Un altro esempio: lancio simultaneo di due dadi

Lo spazio campione associato a questo esperimento è

$$S = \{(i, j) : 1 \le i \le 6; 1 \le j \le 6\}$$

ovvero, l'insieme di tutte le coppie ordinate (*i*, *j*) di numeri interi maggiori o uguali a 1 e minori o uguali a 6; *i* e *j* rappresentano, rispettivamente, l'esito del lancio di ciascuno dei due dadi.

Lo spazio campione S è pertanto formato da 36 eventi semplici:

36 eventi semplici:

$$E_1 = \{1,1\},\ E_2 = \{1,2\},\ E_3 = \{1,3\},\ E_4 = \{1,4\},\ E_5 = \{1,5\},\ E_6 = \{1,6\},$$
$$E_7 = \{2,1\},\ E_8 = \{2,2\},\ E_9 = \{2,3\},\ E_{10} = \{2,4\},\ E_{11} = \{2,5\},\ E_{12} = \{2,6\},$$
$$E_{13} = \{3,1\},\ E_{14} = \{3,2\},\ E_{15} = \{3,3\},\ E_{16} = \{3,4\},\ E_{17} = \{3,5\},\ E_{18} = \{3,6\},$$
$$E_{19} = \{4,1\},\ E_{20} = \{4,2\},\ E_{21} = \{4,3\},\ E_{22} = \{4,4\},\ E_{23} = \{4,5\},\ E_{24} = \{4,6\},$$
$$E_{25} = \{5,1\},\ E_{26} = \{5,2\},\ E_{27} = \{5,3\},\ E_{28} = \{5,4\},\ E_{29} = \{5,5\},\ E_{30} = \{5,6\},$$
$$E_{31} = \{6,1\},\ E_{32} = \{6,2\},\ E_{33} = \{6,3\},\ E_{34} = \{6,4\},\ E_{35} = \{6,5\},\ E_{36} = \{6,6\}.$$

Ai 36 eventi semplici corrispondono 11 possibili valori numerici (supporto della v.a. Y):

$$\{Y = 2\} = \{E_1\},$$
$$\{Y = 3\} = \{E_2, E_7\},$$
$$\{Y = 4\} = \{E_3, E_8, E_{13}\},$$
$$\{Y = 5\} = \{E_4, E_9, E_{14}, E_{19}\},$$
$$\{Y = 6\} = \{E_5, E_{10}, E_{15}, E_{20}, E_{25}\},$$
$$\{Y = 7\} = \{E_6, E_{11}, E_{16}, E_{21}, E_{26}, E_{31}\},$$
$$\{Y = 8\} = \{E_{12}, E_{17}, E_{22}, E_{27}, E_{32}\},$$
$$\{Y = 9\} = \{E_{18}, E_{13}, E_{28}, E_{33}\},$$
$$\{Y = 10\} = \{E_{24}, E_{29}, E_{34}\},$$
$$\{Y = 11\} = \{E_{30}, E_{35}\},$$
$$\{Y = 12\} = \{E_{36}\}.$$

Gli eventi semplici sono equiprobabili

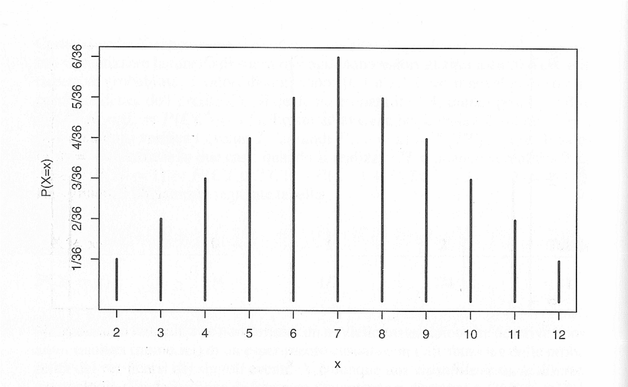
- Se i dadi non sono truccati, tutti gli eventi semplici dello spazio campione **S** sono equiprobabili.
- A ciascun evento semplice può quindi essere assegnata la stessa probabilità, ovvero 1/36.

Possiamo quindi costruire la nostra distribuzione di probabilità

x	combinazioni possibili						p(x)
2	1,1						**1/36**
3	1,2	2,1					**2/36**
4	2,2	3,1	1,3				**3/36**
5	2,3	3,2	4,1	1,4			**4/36**
6	3,3	4,2	2,4	5,1	1,5		**5/36**
7	3,4	4,3	5,2	2,5	6,1	1,6	**6/36**
8	4,4	5,3	3,5	6,2	2,6		**5/36**
9	6,3	3,6	5,4	4,5			**4/36**
10	5,5	6,4	4,6				**3/36**
11	5,6	6,5					**2/36**
12	6,6						**1/36**

Abbiamo sommato le probabilità degli eventi elementari.

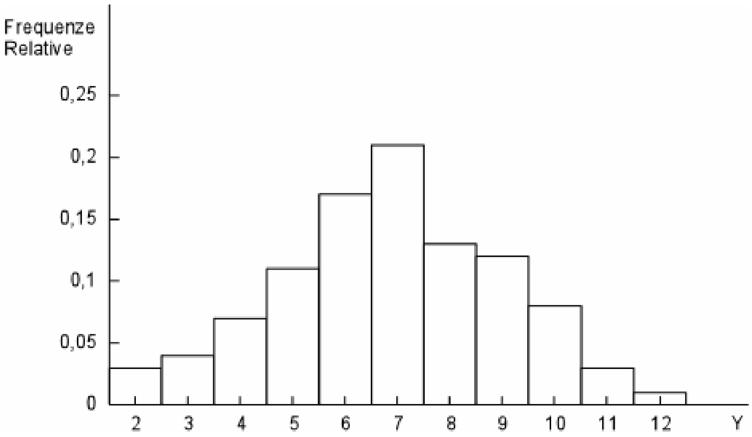
In termini grafici:



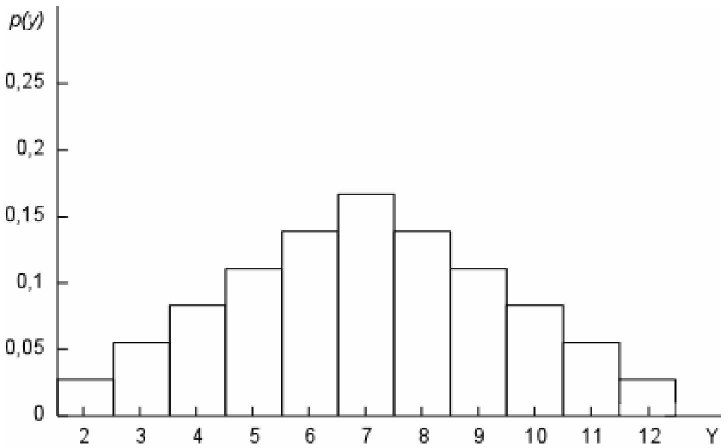
Cosa succede se lanciamo, empiricamente, i due dadi per 100 volte consecutive?

Y	Frequenze osservate in 100 lanci	Frequenze relative
$y_1 = 2$	**3**	0,03
$y_2 = 3$	**4**	0,04
$y_3 = 4$	7	0,07
$y_4 = 5$	**11**	0,11
$y_5 = 6$	**17**	0,17
$y_6 = 7$	**21**	0,21
$y_7 = 8$	**13**	0,13
$y_8 = 9$	**12**	0,12
$y_9 = 10$	**8**	0,08
$y_{10} = 11$	**3**	0,03
$y_{11} = 12$	**1**	0,01

La distribuzione empirica (con n=100 lanci)



La distribuzione teorica della v.a. Y= esito del lancio di due dadi



Osservazione

- La prima distribuzione (empirica) si può considerare una approssimazione della seconda (teorica);
- La distribuzione di probabilità di una v.a. fornisce *un* modello teorico *della distribuzione di frequenza di una popolazione empirica, di una popolazione reale*.

Funzione di ripartizione di una v.a.

In alcuni casi, può risultare utile calcolare la probabilità che **X** assuma un valore minore o uguale a x_k, cioè:

$$F(x_k) = P[X \leq x_k]$$

Questa funzione è detta **funzione di ripartizione di X**, e vale la relazione:

$$F(x_k) = p(x_1) + p(x_2) + \ldots + p(x_k) = \sum_{x_i \leq x_k} p(x_i)$$

Analogie con la "distribuzione cumulata" di una variabile statistica.

Proprietà della funzione di ripartizione

1. $F(-\infty) \to 0$
2. $F(x_n) = F(+\infty) \to 1$
3. $F(x_i) - F(x_{i-1}) = p(x_i)$

Dalla f.r. si può sempre ricavare la funzione di probabilità

Osservazione: la f.r. è utile, ad esempio, per calcolare la probabilità di un intervallo tra due numeri:
dati 2 numeri a e b, con a<b, si ha:
P (a<X<b) = P (X≤b) - P (X≤a) = F(b) - F(a)

Inoltre, si ha che:

La f.r. F(x) è una funzione reale monotòna non-decrescente tale che:

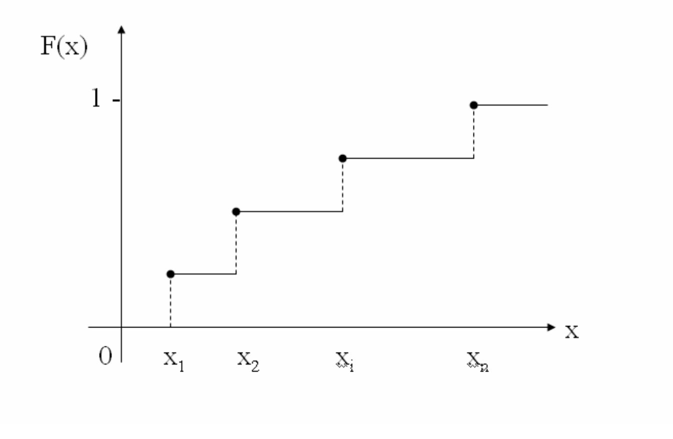
$$0 \le F(x) \le 1 \qquad \forall x$$

La rappresentazione analitica di F(x) è una “funzione a gradini”.

Una f.r. è una funzione a gradini tra 0 e 1, del tipo:

$$F(x)=\begin{cases}0 & \text{per } -\infty < x < x_1 \\ p(x_1) & \text{per } x_1 \le x < x_2 \\ p(x_1)+p(x_2) & \text{per } x_2 \le x < x_3 \\ p(x_1)+p(x_2)+\ldots+p(x_i) & \text{per } x_i \le x < x_{i+1} \\ p(x_1)+p(x_2)+\ldots+p(x_{n-1}) & \text{per } x_{n-1} \le x < x_n \\ 1 & \text{per } x_n \le x < \infty \end{cases}$$

In termini grafici:



Valori caratteristici di una v.a. discreta

Valore atteso di una variabile aleatoria discreta

Sia *Y* una variabile aleatoria discreta con una distribuzione di probabilità *p*(*y*). Il valore atteso o speranza matematica di *Y* è definito come:

$$E\left(Y\right)=\sum_{i=1}^{n} y_i P\left(Y=y_i\right)$$

$$Supporto=\left\{y_1, y_2, \ldots y_i \ldots . y_n\right\}$$

Esempio:
X = esito del lancio di un solo dado

Supponendo che il dado non sia truccato, la distribuzione di probabilità di *X* sarà uniforme: $\mathbf{P(X_i=x_i) = 1/6}$, con $i = 1, ..., 6$. Il valore atteso di *X* è pertanto:

$$E(X) = \sum xp(x) =$$

$$= 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = 3,5$$

Proprietà dell'operatore valore atteso

- Valore atteso della

somma di due variabili aleatorie

Proprietà 1

Il valore atteso della somma di due variabili aleatorie discrete, *X* e *Y* , è uguale alla somma dei rispettivi valori attesi:

$$E(X+Y)=E(X)+E(Y)$$

Dalla proprietà 1 derivano le seguenti proprietà (proprietà di linearità):

$$E(a)=a$$

$$E(aX)=aE(X)$$

$$E(aX+bY)=aE(X)+bE(Y)$$

$$E(X-Y)=E(X)-E(Y)$$

Esempio

Consideriamo l'esperimento consistente
nel lancio di due dadi (il nostro primo esempio).

Siano: X_1 l'esito prodotto dal lancio del primo dado e X_2 l'esito prodotto dal lancio del secondo dado. Sia poi *Y* = *X1 + X2*.

Il valore atteso di *Y sarà* pertanto:

$$E(X_1 + X_2) = E(X_1) + E(X_2) = 3{,}5 + 3{,}5 = 7$$

Che era in effetti il "valore medio" che avevamo trovato per i dati del nostro primo esempio.

Varianza di una variabile aleatoria discreta

La varianza di una variabile aleatoria discreta *X* è definita come il valore atteso di $(X - E(X))^2$.

$$Var(X) = E\left[\left(X - E(X) \right)^2 \right]$$

Cioè, in altri termini:

$$Var(X) = \sum_{i=1}^{n} \left(x_i - E(X) \right)^2 p(x_i)$$

Deviazione standard di una v.a.

$$\sigma(X) = \sqrt{Var(X)}$$

Esempio: lancio di un dado

Abbiamo già calcolato il valore atteso di questa v.a., che è pari a 3,5.

$$Var(X) = \sum_{i=1}^{n} \left(x_i - E(X)\right)^2 p(x_i)$$

Applichiamo la formula:

$$Var(X) = (1-3{,}5)^2 \frac{1}{6} + (2-3{,}5)^2 \frac{1}{6} + (3-3{,}5)^2 \frac{1}{6} +$$

$$+ (4-3{,}5)^2 \frac{1}{6} + (5-3{,}5)^2 \frac{1}{6} + (6-3{,}5)^2 \frac{1}{6} = \frac{35}{12} \cong 2{,}92$$

$$\sigma = \sqrt{\frac{35}{12}} \cong 1{,}71$$

Una formula alternativa per la varianza

La formula

$$Var(X) = \sum_x \left(x - E(X)\right)^2 p(x)$$

È equivalente alla formula:

$$Var(X) = E(X^2) - \left[E(X)\right]^2$$

Proprietà dell'operatore varianza

La varianza di una variabile aleatoria discreta *X* moltiplicata per una costante è uguale alla varianza della variabile aleatoria moltiplicata per la costante al quadrato:

$$Var(aX) = a^2 Var(X)$$

La varianza è un operatore quadratico.

Esempio

X_1 = {2, 3, 5, 6}

X_2 = {4, 6, 10, 12}

$Var\ (X_1) = 2{,}5$

$Var\ (X_2) = 10{,}0 = 2{,}5 \times 4 = 2{,}5 \times 2^2$

Proprietà dell'operatore varianza

La varianza di una variabile aleatoria discreta *X* non cambia se a ciascun valore *x* viene sommata una costante *a:*

$$Var(X + a) = Var(X)$$

Proprietà

Se *X* e *Y* sono due variabili aleatorie indipendenti, allora si ha che:

$$Var(X+Y) = Var(X) + Var(Y)$$

Differenza tra due variabili aleatorie indipendenti

Analogamente, si ha:

$$Var(X-Y) = Var(X) + Var(Y)$$

Variabile aleatoria continua

- X è una v.a. *continua* se il suo dominio Ω è un insieme infinito non numerabile, cioè con la "potenza del continuo". Quindi:
- 1. L'insieme dei valori x che possono essere assunti dalla v.a. X è infinito non numerabile (ad esempio coincide con $\Re$, o con un suo intervallo[a,b]);
- 2. Perdono di significato i singoli punti x ed è necessario procedere con riferimento ad intervalli.
- 3. Le probabilità associate alla v.a. X sono interpretate da una funzione detta "di densità".

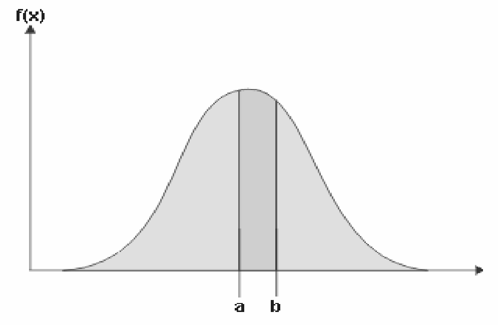
Variabile aleatoria continua

- In tale contesto, non è possibile elencare i valori che la v.a. X assume, associati alle rispettive probabilità, come accadeva nel caso discreto.
- Il problema viene superato associando a ciascun punto dell'intervallo in cui è definita X una **funzione matematica**, f(x), che non è la probabilità, ma è pari alla probabilità di un "intervallo infinitesimo";

$$f(x_0) = \lim_{\varepsilon \to 0} P\left[x_0 \le X \le x_0 + \varepsilon\right]$$

Questa funzione è detta **funzione di densità della v.a. X.**

Distribuzioni di probabilità continue



L'area sottesa dalla curva tra due valori (es. a-b) è la probabilità che la variabile aleatoria assuma valori compresi tra a e b.

Proprietà di una funzione di densità.

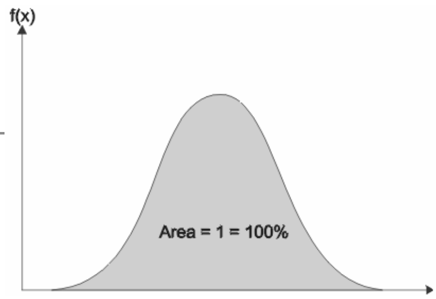
$$f(x) \geq 0 \qquad \forall x \in \mathbf{R}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Sono cioè soddisfatti i primi due assiomi di Kolmogorov.

Probabilità che X sia compresa tra a e b:

$$P(a < X < b) = \int_{a}^{b} f(x)dx$$



$$f(y) \geq 0$$

$$\int_{-\infty}^{\infty} f(y)\,dy = 1$$

Variabile aleatoria continua

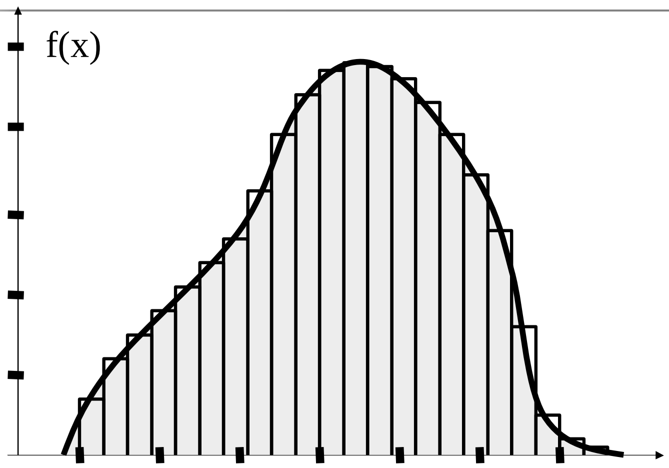
In una v.a. continua, probabilità diverse da zero possono quindi essere assegnate ad *intervalli di valori* della stessa variabile aleatoria continua *X*.

A ciascuno dei singoli valori puntuali che la variabile aleatoria continua può assumere, invece, è sempre associata una probabilità uguale a zero, $P(X = x) = 0$.

Variabile aleatoria continua

Quest'ultima affermazione è facilmente comprensibile se interpretiamo la probabilità in termini geometrici:
l'area sottesa alla funzione di densità *f(x)* nell'intervallo corrispondente ad un punto è necessariamente uguale a zero.

Variabile aleatoria continua: "limite" del caso discreto



Valore atteso per variabili continue

Variabili aleatorie discrete: somma

$$E[x] = \sum_{i=1}^{i_{\max}} x_i \, p\left(X = x_i\right)$$

Variabili aleatorie continue: integrale

$$E[x] = \int_{x_{\min}}^{x_{\max}} x f(x) dx$$

Varianza per variabili continue

Variabili aleatorie discrete: somma

$$Var[x] = \sum_{i=1}^{i_{\max}} [x_i - E(X)]^2 \, p(X = x_i)$$

Variabili aleatorie continue: integrale

$$Var[x] = \int_{x_{\min}}^{x_{\max}} [x - E(X)]^2 f(x) dx$$

Le proprietà degli operatori valore atteso e varianza sono le stesse, sia per v.a. discrete, sia per v.a. continue.