

Lezione 12: Variabili Casuali

Corso di Statistica
Facoltà di Economia
Università della Basilicata

Prof. Massimo Aria

- aria@unina.it

Variabile Casuale

Una **Variabile Casuale** (v.c.) è una regola (funzione reale) che associa ad E (evento elementare di Ω) uno ed un solo numero reale.

Spesso una v.c. non è altro che la traduzione numerica immediata degli eventi elementari, ma può essere anche una funzione più complessa definita sugli eventi di Ω .

Si pensi ad esempio alla prova "lancio di un dado".

Per convenzione, attribuiamo un numero ad ognuna delle facce del dado.

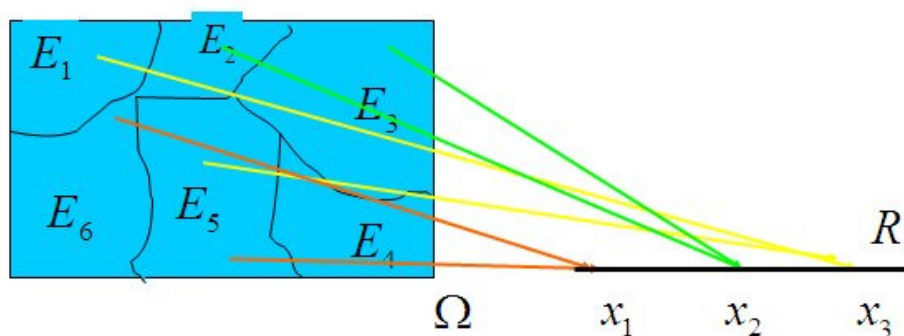
Oppure nelle carte da gioco, dove associamo un numero ad ognuna delle carte del mazzo (l'asso vale uno, il re vale 10, ecc.)

Ciò in quanto l'uso di numeri, in luogo della descrizione qualitativa di eventi, permette il ricorso ad una serie di strumenti di ausilio alla descrizione della prova (i.e. il calcolo del risultato atteso del lancio di un dado o di una misura della variabilità dei risultati).

Definizione della variabile casuale

Variabile Casuale

Per ogni Evento dello Spazio Campionari o
la v.c. X assume il valore reale x_i



Funzione di probabilità e Supporto

Ad ogni valore assunto dalla variabile casuale è possibile associare una probabilità utilizzando la misura di probabilità definita sui sottoinsiemi dello spazio campionario Ω .

"Si verifica l'evento E con probabilità $P(E)$ "

da cui

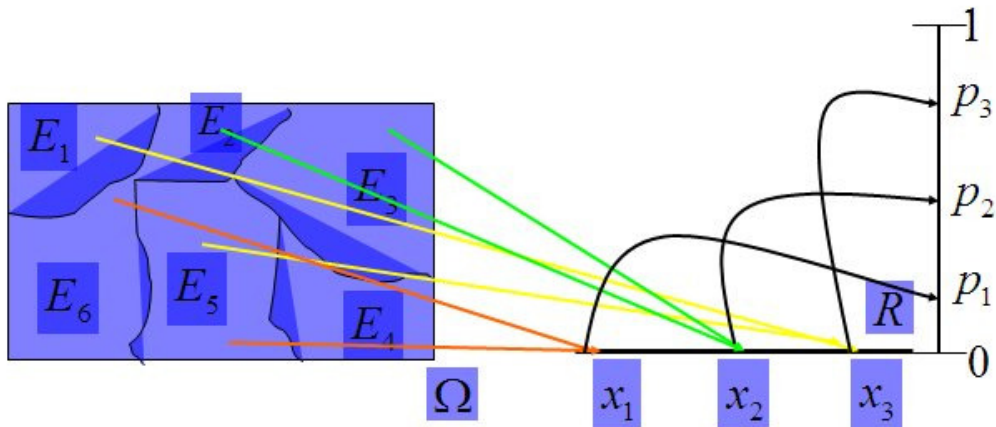
"La v.c. X assume il valore x con probabilità $P(x)$ "

La misura di probabilità definita sui valori di X prende il nome di **funzione di probabilità della v.c. X** .

Si indica invece come **supporto della v.c. X** l'insieme dei valori che X può assumere con una certa probabilità positiva in una prova specifica.

Definizione del supporto della variabile casuale

Supporto della v.c. X



Tipologia di variabili casuali

Una variabile casuale può essere:

• v.c. discreta

È definita su uno spazio campionario Ω discreto, cioè composto da un numero di eventi finito e numerabile. La definizione della funzione di probabilità della v.c. è agevole in quanto, essendo Ω discreto, è possibile enumerare tutti i possibili eventi, associare ad ognuno di essi un numero e quindi la probabilità che la v.c. assuma tale numero.

Alcune prove sono naturalmente discrete (numero di figli, lancio della moneta, ecc.) altre invece possono essere rese discrete da un particolare partizionamento di Ω (reddito superiore a 50.000€, sì o no?, durata del percorso Napoli-Roma, meno di 2 ore, sì o no?, ecc.)

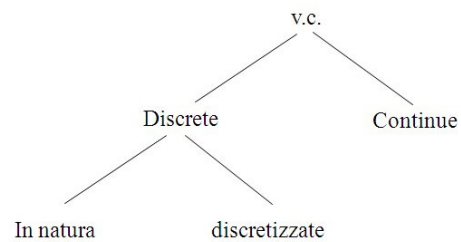
• v.c. continua

È invece definita su uno spazio campionario Ω continuo, dove il numero degli eventi non è finito.

In questo caso la definizione di una funzione di probabilità è possibile solo qualora questa è misurabile e presenta una inversa.

Tipi di variabili casuali

Il modello di probabilità è una esemplificazione della realtà e può essere descritto da una v.c.



Tipologia di variabili casuali

Variabile casuale discreta

Una variabile casuale discreta è un corrispondenza tra gli eventi di Ω ed un insieme discreto (finito o numerabile) di numeri reali.

Una v.c. discreta è nota se si conoscono i valori che può assumere e le rispettive probabilità.

Una sua rappresentazione grafica si ottiene ponendo sull'asse delle ascisse i valori della X e sulle ordinate le probabilità associate alle x.

Sia $P(X=x_i)=p_i$

Una v.c. discreta è ben definita se rispetta i postulati della probabilità:

1) $p_i \geq 0$ per ogni $i=1,2,\dots,k$ (positività)

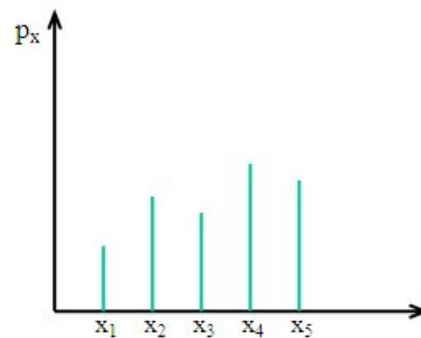
2) $\sum p_i = 1$ (certezza)

Schema di una v. c. discreta

Schematicamente

Esempio: lancio di un dado, di una moneta, estrazione di un numero al lotto (a k eventi sono associati k valori della V.C.)

Eventi	$p(E)$	v.c.X	p_x
E_1	p_1	x_1	p_{x_1}
E_2	p_2	x_2	p_{x_2}
E_3	p_3	x_3	p_{x_3}
\vdots	\vdots	\vdots	\vdots
E_k	p_k	x_k	p_{x_k}



Esempio: costruzione di una v. c. discreta

Esempio: consideriamo una famiglia con 3 figli

$$\Omega = \{ \overset{E_1}{MMM}, \overset{E_2}{MMF}, \overset{E_3}{MFM}, \overset{E_4}{FMM}, \overset{E_5}{MFF}, \overset{E_6}{FMF}, \overset{E_7}{FFM}, \overset{E_8}{FFF} \}$$
$$\frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8}$$

Definiamo la variabile casuale X = "numero dei figli maschi"

	X	p_i
$\{E_8\}$	0	1/8
$\{E_5 \cup E_6 \cup E_7\}$	1	3/8
$\{E_2 \cup E_3 \cup E_4\}$	2	3/8
$\{E_1\}$	3	1/8

Variabile casuale continua

Una **variabile casuale continua** X è una funzione misurabile e a valori reali che assegna ad ogni evento E , contenuto in Ω di uno spazio di probabilità continuo, un numero reale x appartenente a \mathbb{R} .

In altre parole la v.c. continua può assumere tutti i valori compresi in un dato intervallo reale.

Essa presenta una maggiore complessità rispetto alle v.c. discrete in quanto non è possibile elencare tutti i valori che la variabile casuale assume, essendo questi un'infinità.

Per ovviare a ciò si assegna la probabilità, non a singoli valori, ma ad intervalli di valori di \mathbb{R} sulla base della natura e della specificità della prova.

Funzione di ripartizione e di densità

Data una v.c. continua X , essa è nota se, per ogni numero reale x , è nota funzione $F(x)$ oppure la funzione $f(x)$.

Le due funzioni sono definite dalla seguente relazione:

$$F(x) = P(-\infty \leq X \leq x) = P(X \leq x) = \int_{-\infty}^x f(w) dw$$

F(x) è detta **funzione di ripartizione** della v.c. X . Essa esprime la probabilità che la v.c. X "assuma valori fino a x ".

La conoscenza di $F(x)$ è necessaria e sufficiente per calcolare qualunque probabilità su X (utilizzando i postulati e i teoremi della probabilità).

f(x) è detta **funzione di densità** della v.c. X . Essa esprime la densità di probabilità in un intervallo infinitesimale.

Proprietà della $F(x)$

- 1) $F(x)$ non è decrescente $x_1 < x_2 \Rightarrow F(x_1) < F(x_2)$
- 2) $\lim_{x \rightarrow -\infty} F(x) = 0$; $\lim_{x \rightarrow +\infty} F(x) = 1$;
- 3) $F(x)$ continua da destra $\lim_{x \rightarrow x_0^+} F(x) = F(x_0)$

Proprietà della funzione di ripartizione

Esistenza di una v.c. continua

Così come per le variabili discrete, una v.c. continua è ben definita se rispetta i postulati della probabilità:

1) *Positività*

$$f(x) \geq 0$$

2) *Certezza*

L'integrale di $f(x)$ nell'intervallo in cui è definita la variabile X è pari a 1.

Se X è una variabile casuale che assume valori in $[a, b]$, essa è definita se:

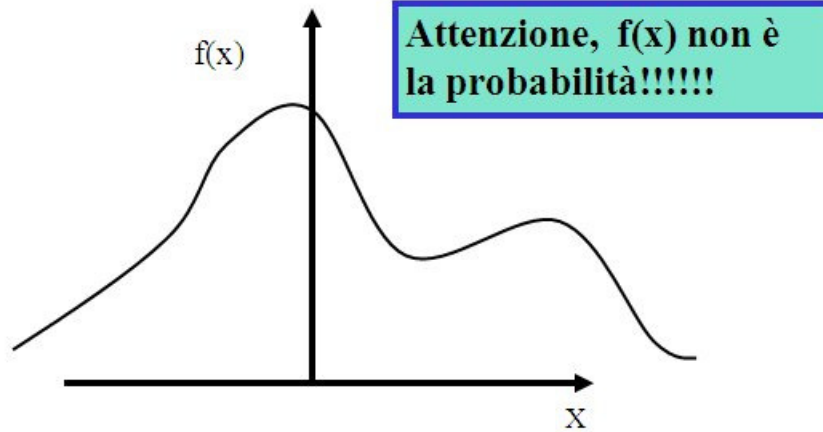
Esiste la $f(x)$ tale che per ogni $x_0 \in (a, b)$

$$P(x_0 \leq X \leq x_0 + dx) = f(x_0) dx$$

$$\left\{ \begin{array}{l} f(x) \geq 0 \\ \int_a^b f(x) dx = 1 \end{array} \right. \quad \text{Per ogni } x \text{ tale che } a < x < b$$

Esistenza di una v.c. continua

Rappresentazione di una v.c. continua

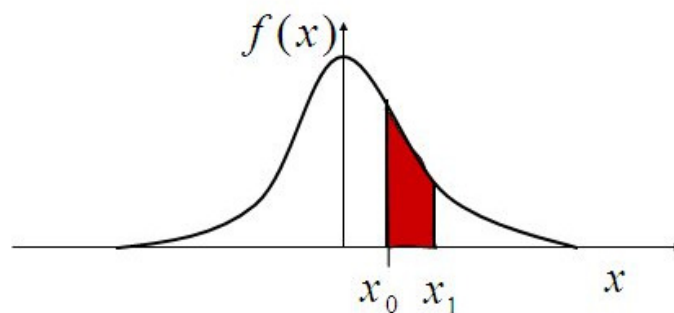


$f(x)$ è proporzionale (a meno di un infinitesimo) alla probabilità di un intervallo sufficientemente piccolo

Calcolo della probabilità di una v.c. continua

Calcolo delle probabilità per v.c. continue

$$\begin{aligned} P(x_0 \leq X \leq x_1) &= P(X \leq x_1) - P(X \leq x_0) = \\ &= F(x_1) - F(x_0) \end{aligned}$$



Valore atteso di una v.c.

Così come per le variabili statistiche, anche per le variabili casuali è possibile calcolare alcuni indici di sintesi che ne consentano la descrizione e il confronto.

In particolare si fa riferimento al valor medio e alla varianza di una v.c.

Si definisce **valore atteso** (*momento primo* o *valor medio*) di una v.c. X , la somma dei valori della X ponderati per le rispettive probabilità.

Nel caso di v.c. continue il concetto di somma è da "intendersi nel continuo" per cui l'operatore somma trova suo analogo nell'integrale.

Il valore atteso è indicato con il simbolo $E()$.

Valore atteso di una v.c.

$$E(X) = \sum_i x_i \cdot p_i \quad \text{v.c. discreta}$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx \quad \text{v.c. continua}$$

Valore atteso di una v.c.

Varianza di una v.c.

Si definisce **varianza** (*momento secondo*) di una v.c. X , la somma degli scarti al quadrato tra i valori x e il valor medio, ponderati per le rispettive probabilità.

Nel caso di v.c. continue il concetto di somma è da "intendersi nel continuo" per cui l'operatore somma trova suo analogo nell'integrale.

Varianza di una v.c.

$$\text{var}(X) = \sum_i [x_i - E(X)]^2 \cdot p_i \quad \text{v.c. discreta}$$

$$\text{var}(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 \cdot f(x) dx \quad \text{v.c. continua}$$

Varianza di una v.c.

Calcolo della media e della varianza di una v.c.

La variabile casuale X = "numero dei figli maschi"

x_i	p_i	$x_i \cdot p_i$	$[x_i - E(X)]$	$[x_i - E(X)]^2 \cdot p_i$
0	0,125	0	-1,5	0,28125
1	0,375	0,375	-0,5	0,09375
2	0,375	0,75	0,5	0,09375
3	0,125	0,375	1,5	0,28125
	1	1,5		0,75

$$E(X) = \sum_i x_i \cdot p_i = 1,5$$

$$\text{var}(X) = \sum_i [x_i - E(X)]^2 \cdot p_i = 0,75$$

Disuguaglianza di Cebicev

La **disuguaglianza di Cebicev** è uno dei risultati più notevoli del calcolo delle probabilità.

La disuguaglianza afferma che, per ogni v.c. X che $E(X)=\mu$ e $\text{Var}(X)=\sigma^2 < +\infty$, si ha:

$$\mathbf{P(|X - \mu| < \epsilon) \geq 1 - (1 / \epsilon^2)}$$

per ogni $\epsilon > 0$.

Ponendo $\epsilon = k\sigma$, la disuguaglianza si può esprimere nel seguente modo:

$$\mathbf{P(|X - \mu| < k) \geq 1 - (\sigma^2 / k^2)}$$

per ogni $k > 0$.

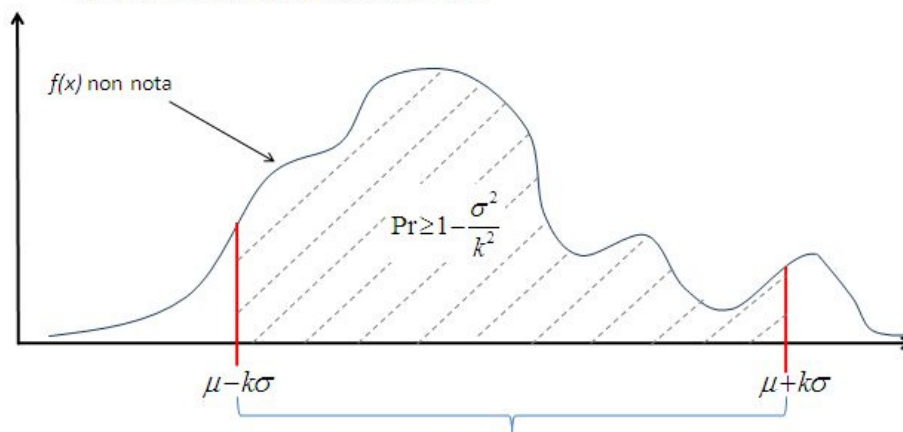
In altre parole, l'importanza di questo risultato risiede nel fatto che, noti i momenti primo e secondo di una v.c. X, è sempre possibile trovare un limite inferiore alla probabilità che la variabile assuma valori compresi in un intervallo $[\mu - \epsilon; \mu + \epsilon]$.

Ciò anche quando è ignota la funzione di probabilità della variabile casuale!!!

Disuguaglianza di Cebicev

Disuguaglianza di Cebicev

Sia X una v.c. con $E(X)=\mu$ e $\text{Var}(X)$ pari a σ^2 .



Anche se $f(x)$ è ignota, è possibile trovare un minimo alla probabilità che la v.c. assuma valori in un dato intervallo!

Il valore della probabilità e l'ampiezza dell'intervallo dipende dal valore di k e dalla varianza della v.c.

Nella prossima lezione

Nella prossima lezione si affronteranno i seguenti argomenti:

- alcune variabili casuali discrete notevoli
- v.c. Uniforme Discreta
- v.c. di Bernoulli
- v.c. Binomiale
- v.c. di Poisson