# Nozioni sulla Statistica per le Variabili Aleatorie Discrete - Sommario

X

*Nozioni statistiche applicate alle variabili aleatorie discrete.*

X

# 1. Definizione di Valor medio

## Definizione del Valore Medio

X

*Definizione di valore medio per v.a. discrete, per densità associate a v.a. discrete. Definizione di valore medio finito e centrato.*

X

# 1. Definizione di Valore Medio per v.a. discrete

#Definizione

**Definizione (valore medio per v.a. discreta).**

Sia $X : \Omega \longrightarrow \mathbb{R}$ una *variabile aleatoria discreta*, ovvero che assume un numero *al più numerabile*, associata alla densità $q$. Allora ha senso considerare il valore

$$\overline{X} = \sum_{j=1}^{N} \frac{p\{X = X_j\}x_j}{p(\Omega)} = \sum_{j=1}^{N} q(x_j)x_j = \sum_{x \in \mathbb{R}} q(x)x$$

Abbiamo $p(\Omega) = 1$, ovvero la sua misura è $1$.

Si indica il *valore medio* di una v.a. discreta con

$$E[X]$$

e sinonimi per questo possono essere la *media*, il *valore atteso* o la *speranza matematica*.

#Osservazione

Osservazione (buona posizione della definizione).

La definizione è ben posta. Consideriamo la condizione

$$\sum_{x\in\mathbb{R}} |x|q(x) < +\infty$$

Infatti, la condizione $\sum |x|q_x$ è soddisfatta *sempre* quando abbiamo un numero *finito* di valori.
Se invece ho una *quantità numerabile* di oggetti, posso spezzare la sommatoria in due, ottenendo

$$\sum_{x<0}(-x)q(x) + \sum_{x\geq 0} xq(x) = \sum_{x\in\mathbb{R}} xq(x)$$

In entrambi abbiamo delle *serie a termini positivi*, quindi evitiamo il caso in cui abbiamo una serie *indeterminata*, dato che abbiamo posto la prima condizione.

Allora le opzioni sono due: o la media diverge a $+\infty$ o ho *valore medio finito* ( $\sum |x|q(x) < +\infty$). Se invece vale che $E[X] = 0$ allora si dice che $X$ è *"centrata"*.

## 2. Definizione di Valor Medio Generalizzato

#Definizione

Definizione (valor medio per una v.a. qualsiasi).

Sia $X : \Omega \longrightarrow \mathbb{R}^+$ una *variabile aleatoria non-negativa*. Si definisce il *valor medio* di $X$ la quantità

$$E[X] := \sup\{Y : \Omega \longrightarrow \mathbb{R} | 0 \leq Y \leq X\}$$

ammesso che sia finita.

Oppure viene definita come l'*integrale di Lebesgue* definita come

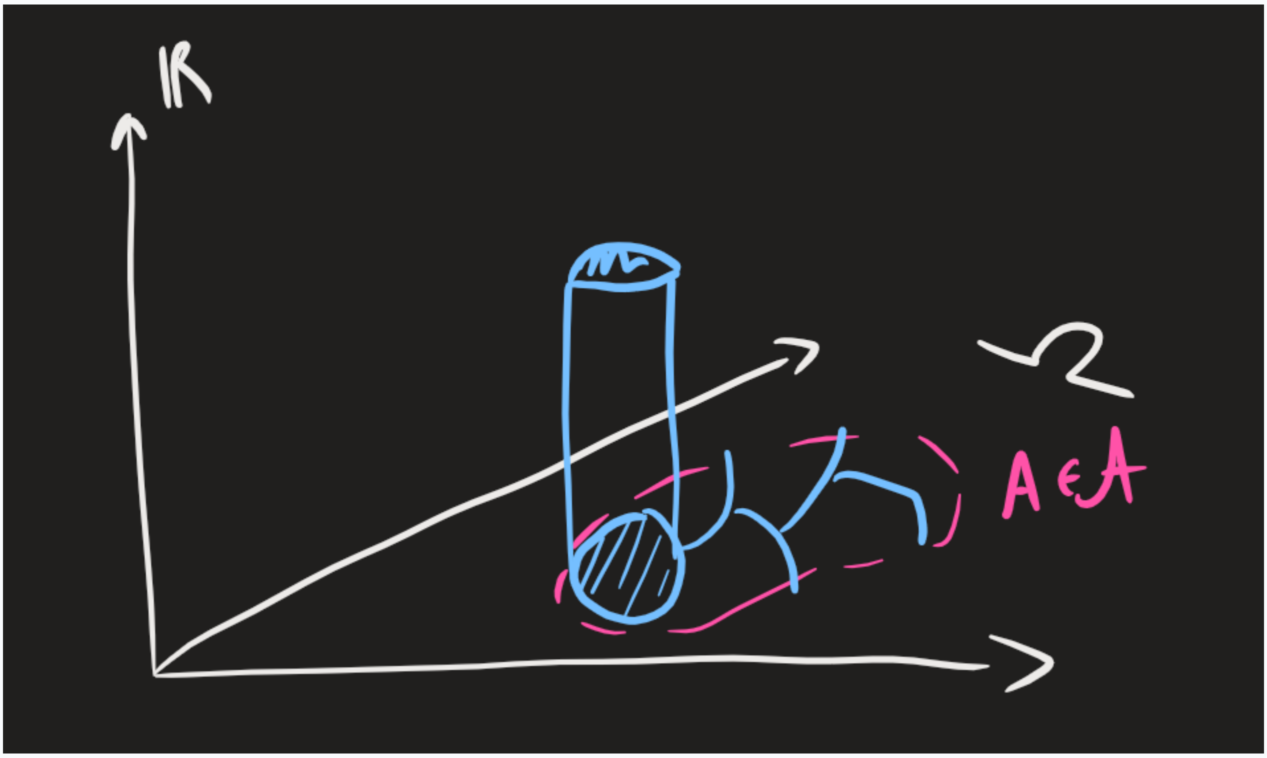
$$\boxed{E[X] := \int_\Omega X \text{ d}p}$$

con $p$ la misura su $\Omega$ (ammesso che sia sempre finita!).

Nel caso in cui abbiamo variabili aleatorie a *segno qualunque*, basta spezzare in $X^+$ e $X^-$ e vedere se entrambe le loro medie $E[X^+], E[X^-]$ sono finite. In tal caso, si definisce

$$E[X] := E[X^+] - E[X^-]$$

**FIGURA 2.1.** (*Lebesgue*)




## 2. Proprietà del Valor medio

### Proprietà del Valore Medio

X

*Proprietà (teoremi) sul valore medio. Valore medio delle densità composte, linearità del valore medio, monotonia del valore medio con corollario, confronto tra valori medi, prodotto tra valori medi di variabili aleatori indipendenti.*

X

## 0. Voci correlate

- Definizione del Valore Medio

## 1. Composizione del Valor Medio

#Teorema

> Teorema (composizione del valor medio).
>
> Sia $X = (X_1, \ldots, X_N)$ un *vettore aleatorio discreto* avente densità congiunta $q : \mathbb{R}^N \longrightarrow \mathbb{R}$. Sia $g : \mathbb{R}^N \longrightarrow \mathbb{R} \in \mathcal{C}^0$.

Poniamo la variabile aleatoria $Y$ definita come la composizione $Y := g \circ X$. Allora $Y$ ha *valore medio finito* (ovvero è definibile) se e solo se

$$\sum_{x \in \mathbb{R}^N} |g(x)| q(x) < +\infty$$

Se ha valore medio finito, allora vale che

$$E[Y] = \sum_{x \in \mathbb{R}^N} g(x) q(x)$$

In altre parole, al posto di dover la densità associata a $Y$, che potrebbe risultare ostica (o addirittura impossibile!) da trovare, dobbiamo solo calcolare il valore medio *mediante* la $g$ e la $q$, due funzioni già note.

#Dimostrazione

**DIMOSTRAZIONE** del Teorema 1 (composizione del valor medio)
Omessa. ■

Una conseguenza diretta di questo teorema è il seguente.

#Proposizione

Proposizione (calcolo di valori medi su spazi campionari).

Sia $(\Omega, p, \mathcal{A})$ uno *spazio di probabilità discreto* e $X$ una variabile aleatoria (in realtà funzione qualsiasi ma vabbè) su di esso avente media finita. Allora possiamo calcolare la sua media mediante lo spazio di probabilità

$$E[X] = \sum_{\omega \in \Omega} X(\omega) p(\{\omega\})$$

#Dimostrazione

**DIMOSTRAZIONE** della Proposizione 2 (calcolo di valori medi su spazi campionari)
Basta considerare che abbiamo spazi probabilistici discreti: dunque $\{X = x\}$ formano partizioni di $\Omega$. ■

## 2. Linearità del Valore Medio

Adesso vediamo una proprietà particolare del valore medio, che ce lo fa pensare come una specie di "operatore lineare".

#Proposizione

Proposizione (il valore medio è lineare).

Siano $X_1, X_2$ due v.a. discrete aventi valore medio finito. Allora valgono le seguenti proprietà:
i. *omogeneità*

$$\forall c \in \mathbb{R}, E[cX_1] = cE[X_1]$$

ii. *additività*

$$E[X_1 + X_2] = E[X_1] + E[X_2]$$

#Dimostrazione
**DIMOSTRAZIONE** della Proposizione 3 (il valore medio è lineare)
Bisogna considerare le composizioni $Y := c \cdot (X_1)$, $Z := (X_1) + (X_2)$ e applicare il teorema sulla composizione del valor medio. ■

## 3. Monotonia e Maggiorazione del Valor Medio

Vediamo le seguenti proprietà, che ci faranno pensare il valor medio come una *"specie di integrale"*, versione probabilistica.

#Proposizione

Proposizione (monotonia del valor medio).

Siano $X_1, X_2$ due v.a. discrete aventi valore medio finito. Allora vale la seguente implicazione

$$X_1 \leq X_2 \implies E[X_1] \leq E[X_2]$$

#Dimostrazione
**DIMOSTRAZIONE** della Proposizione 4 (monotonia del valor medio)
Basta porre $Y = X_2 - X_1$ (da cui si ha $Y \geq 0$), che per linearità ha valore medio finito. Quindi si ha

$$E[X_2] - E[X_1] =: E[Y] = \sum_{y \geq 0} yp\{Y = y\} \geq 0$$

che dimostra l'implicazione. ■

#Corollario

Corollario (minorazione del valor medio assoluto).

Sia $X$ una v.a. discreta avente valore medio finito. Allora anche $|X|$ ha valore medio finito e vale che

$$|E[X]| \leq E[|X|]$$

#Dimostrazione

**DIMOSTRAZIONE** del Corollario 5 (minorazione del valor medio assoluto)
Basta applicare la composizione su $g := |\cdot|$. ■

#Osservazione

Osservazione (similitudine con gli integrali).

Osserviamo che queste proprietà sono più o meno simili a quelle per *gli integrali di Riemann* (pagina). Infatti, abbiamo che formano uno *spazio vettoriale* (ovvero è l'integral-funktion è lineare), godono della monotonia, hanno minorazioni del valor assoluto.

Possiamo pensare dunque al *valor medio* come una specie di *"controparte"* per l'integrale. Ad esempio

$$\left|\int_I f\right| \leq \int_I |f| \sim |E[X]| \leq E[|X|]$$

## 4. Il confronto di valori medi

#Lemma

Lemma (confronto di valori medi).

Se $X, Y$ sono due *v.a. discrete* tali che

$$0 \leq X \leq Y$$

Allora vale la seguente implicazione:

$$E[Y] \leq +\infty \implies E[X] \leq +\infty$$

#Dimostrazione

**DIMOSTRAZIONE** del Lemma 7 (confronto di valori medi)
Si tratta di prendere la successione delle somme parziali delle variabili aleatorie $X_n$ definite come

$$X_n(\omega) = \begin{cases} X(\omega), X(\omega) \in \{x_1, \dots, x_n\} \\ 0, \text{altrimenti} \end{cases}$$

Per ipotesi iniziali abbiamo $X_n \leq Y$. Allora

$$E[X_n] \leq E[Y]$$

quindi la serie

$$E[X_n] := \sum_n xp\{X = x\}$$

converge nel caso $E[Y] < +\infty$; altrimenti diverge. ■

# 5. Il prodotto tra valori medi di v.a. indipendenti

Adesso vediamo come si comporta il valor medio rispetto al prodotto.

#Proposizione

Proposizione (prodotto tra valori medi di v.a. indipendenti).

Siano $X_1, \ldots, X_n$ delle *variabili aleatorie discrete* su di esso aventi *valore medio finito*.
Allora, se $X_1, \ldots, X_n$ sono *indipendenti*, allora vale che il prodotto $Y := \prod_n X_n$ ha valor medio finito e vale

$$\boxed{E[Y] = \prod_{j=1}^{n} E[X_j]}$$

#Dimostrazione

**DIMOSTRAZIONE** della Proposizione 8 (prodotto tra valori medi di v.a. indipendenti)
Basta applicare il teorema sulla composizione sulla composta $Y$. Sotto le ipotesi di indipendenza abbiamo

$$\sum_{x \in \mathbb{R}^N} |g(x)|q(x) = \sum_{x \in \mathbb{R}^N} \prod_{j=1}^{N} |x_j| q_j x_j$$

Posso *"far commutare il prodotto con la somma"*: infatti espandendo l'espressione abbiamo

$$\sum_{x_1 \in \mathbb{R}} \cdots \sum_{x_N \in \mathbb{R}} \prod_{j=1}^{N} |x_j| q_j x_j$$

quindi effettuando lo scambio abbiamo, fissando un $x_j$

$$\sum_{x \in \mathbb{R}^N} \prod_{j=1}^{N} |x_j| q_j x_j = \prod_{j=1}^{N} \sum_{x_j \in \mathbb{R}} \underbrace{|x_j| q_j x_j}_{\sim E[X_j]} < +\infty$$

che prova la tesi del primo punto. Ripetendo gli stessi passaggi senza il valor assoluto, abbiamo la tesi finale. ■

X

# 3. Definizione della Varianza e Deviazione Standard

## Definizione di Varianza e Deviazione Standard

X

*Idea della varianza. Definizione di varianza, deviazione standard, variabile aleatoria avente momento secondo finito.*

X

## 0. Voci correlate

- Definizione del Valore Medio
- Proprietà del Valore Medio
- Definizione di Variabile Aleatoria

## 1. Definizione di Varianza e Deviazione Standard

**IDEA.** Supponiamo di avere la varianza di una variabile aleatoria $E[X]$. Però questa informazione ci sembra piuttosto *"povera"*, vogliamo sapere di più di questa variabile aleatoria $X$: in particolare *come* si comporta rispetto alla sua media $E[X]$.
Un primo modo di quantificare quest'idea è di usare la *distanza euclidea* $|X - E[X]|$ e prendere il suo valor medio $E[|X - E[X]|]$; tuttavia presenta dei problemi, in quanto diventa difficile da gestire con valori assoluto.
Un altro approccio che è *"simile"* alla distanza euclidea è di prendere la loro differenza e di considerarne il *quadrato*; ovvero $(X - E[X])^2$. Certo, non è proprio l'idea iniziale ma si comporta similmente. Ad esempio, abbiamo $E[(X - E[X])^2] = 0 \iff X = 0$. Questa sarà l'idea di cui considereremo

#Definizione

**Definizione (varianza, deviazione standard).**

Sia $X$ una *variabile aleatoria discreta avente valor medio finito* $E[X] < +\infty$. Se $E[(X - E[X])^2] < +\infty$ (ovvero la quantità appena considerata ha valor medio finito), possiamo definire la sua varianza come

$$\boxed{\operatorname{var} X := E\left[(X - E[X])^2\right]}$$

Notiamo subito che la *varianza* è sempre positiva, dal momento che lo è pure il suo argomento (in azzurro).

Per non scomodarci con questo problema (infatti avvolte vogliamo sapere *"in che modo si discosta dalla media"*), prendiamo la sua radice e la chiamiamo *deviazione standard* (o *scarto quadratico medio*), definito come

$$\boxed{\sigma_X := \sqrt{\operatorname{var} X}}$$

#Osservazione

Osservazione (l'esistenza di varianza implica l'esistenza della media).

Notiamo che l'esistenza della varianza $\operatorname{var} X$ implica che esiste la media.

Infatti, $(X - E[X])^2 = X^2 - 2(XE[X]) + E[X]^2$, ci interessa sapere come si comporta $X$. Per il *teorema sulla composizione delle medie di variabili aleatorie* (1) deve seguire che

$$\sum_{x \in \mathbb{R}} x^2 q(x) < +\infty$$

Poiché $|x| \leq 1 + x^2$, si verifica subito che

$$\sum_{x \in \mathbb{R}} |x| q(x) \leq \sum_{x \in \mathbb{R}} (1 + x^2) q(x) \leq 1 + \sum_{x \in \mathbb{R}} x^2 q(x) < +\infty$$

## 2. Valore medio avente momento secondo finito

Notiamo che non vale il viceversa: se una variabile aleatoria ha valor medio, non è detto che la media del suo quadrato sia finito. Definiamo una classe di variabili aleatorie che hanno sia *media finita* che il *valor medio del suo quadrato* finito.

#Definizione

Definizione (valor medio avente momento secondo finito).

Si dice che una variabile aleatoria $X$ ha *valor medio finito* se valgono che

$$E[X], E[X^2] < +\infty$$

In altre parole è una variabile aleatoria su cui *possiamo definire* la *varianza*.

# 4. Proprietà della Varianza

Proprietà della Varianza

X

*Proprietà della varianza: comportamento rispetto allo scalamento e alla somma. Condizione equivalente per la nullità della varianza. Disuguaglianza di Čebyšëv. La somma di variabili aleatorie discrete aventi momento secondo finito ha momento secondo finito (con dimostrazione costruttiva).*

X

## 0. Voci Correlate

- Definizione di Varianza e Deviazione Standard
- Proprietà del Valore Medio

## 1. Comportamento della Varianza rispetto alle Operazioni

#Proposizione

Proposizione (prime proprietà).

Sia $X$ una *variabile aleatoria* avente momento secondo finito. Allora valgono le seguenti:

$$\begin{cases} \operatorname{var} X = E\left[X^2\right] - (E[X])^2 \\ \operatorname{var}(cX) = c^2 \operatorname{var} X \\ \operatorname{var}(X+c) = \operatorname{var} X \end{cases}$$

Nella prima ho semplicemente definito un *modo semplice* per scrivere la varianza, nella ho $c^2$, dal momento che *non posso avere valori negativi*, e nella terza rimane uguale perché tanto sto traslando tutto assieme, sia la varianza che la media.

#Dimostrazione

**DIMOSTRAZIONE** della Proposizione 1 (prime proprietà)
Sono conti, usando la *linearità del valor medio* (Proposizione 3 (il valore medio è lineare)). Si consiglia al lettore di farlo per esercizio. ■

## 2. Condizione Necessaria e Sufficiente per la Nullità della Varianza

Prima di passare ad un'altra proprietà della varianza, enunciamo il seguente lemma (una proprietà che è similmente presente per gli integrali).

#Lemma

> Lemma (Condizione necessaria per variabili aleatori costanti).
>
> Sia $X$ una *variabile aleatoria discreta* con *valor medio finito*. Allora vale che:
>
> - Se esiste un $c \in \mathbb{R}$ tale che $p(\{X=c\})=1$, cioè abbiamo la variabile aleatoria costante $X=c$ a meno di un insieme di probabilità nulla,
> - Allora se $X$ è *a segno costante* (quindi o $\geq 0$ o $\leq 0$) e la sua media è nulla ( $E[X]=0$), allora $p(\{X=0\})=1$ (ovvero il valore $c$ dev'essere $0$), a meno di un insieme di probabilità nulla.
>
> Ovvero se abbiamo una *variabile aleatoria costante con media nulla a segno costante*, la costante dev'essere *nulla*.

#Dimostrazione

**DIMOSTRAZIONE** del Lemma 2 (Condizione necessaria per variabili aleatori costanti)
Per $x \neq c$ abbiamo che l'evento $\{X=x\} \subseteq \mathcal{C}\{X=c\}$ (ovvero è incluso nel complementare degli eventi in cui $X$ è $c$). Di conseguenza vale che $p(\{X=x\})=1-1=0$. Ne segue dunque che la sua media è

$$E[X]=\sum_{x \in \mathbb{R}^{\pm}} xp\{X=x\}=cp\{X=c\}=c$$

Per ottenere la tesi basta imporre $E[X]=0$ e dunque ottenere che

$$E[X]=0 \iff c=0$$

Questo ci è garantito dal fatto che stiamo agendo su una *sommatoria a termini costanti*, ovvero non abbiamo *"oscillazioni"* tra numeri negativi e positivi. Si può dimostrare la tesi anche per assurdo, usando lo stesso fatto: infatti se ci fosse un $x \neq 0$ tale che $p(\{X=x\}) \neq 0$, allora varrebbe la sommatoria $\sum_x xp(\{X=x\})>0$, che sarebbe un assurdo per il punto di prima. ■

Da qui possiamo già avere l'idea della prossima proprietà: dato che la varianza è una *"specie di distanza"* di $X$ da $E[X]$, allora si pensa che se è nulla allora $X$ deve concentrarsi sul valore medio (e viceversa!).

#Proposizione

> Proposizione (condizione necessaria e sufficiente per la nullità della varianza).

> Sia $X$ una *variabile aleatoria avente momento secondo finito*.
> Allora sono equivalenti:
>
> $$\operatorname{var} X = 0 \iff \exists c \in \mathbb{R} : p(\{X = c\}) = 1 \implies E[X] = c$$

#Dimostrazione

**DIMOSTRAZIONE** della Proposizione 3 (condizione necessaria e sufficiente per la nullità della varianza)
" $\Longleftarrow$ ": Supponendo che $p(\{X = c\}) = 1$, per il lemma appena enunciato (Lemma 2 (Condizione necessaria per variabili aleatori costanti)) segue che $c = E[X]$. Poiché gli eventi $\{X = E[X]\}$ e $\{(X - E[X])^2 = 0\}$ sono uguali, abbiamo

$$p(\{X = E[X]\}) = p(\{(X - E[X])^2 = 0\}) = 1$$

ancora per il lemma, applicato su $(X - E[X])^2$, abbiamo che sua media (ovvero la varianza di $X$) è nulla:

$$E\left[(X - E[X])^2\right] =: \operatorname{var} X = 0$$

che prova la tesi.
" $\implies$ ": Se $\operatorname{var} X = 0$, abbiamo per definizione

$$E\left[(X - E[X])^2\right] = 0$$

essendo il termine $(X - E[X])^2$ sempre positivo, per lo stesso lemma abbiamo

$$p(\left\{(X - E[X])^2 = 0\right\}) = 1$$

ma allora questo significa

$$p(\{X - E[X] = 0\}) = 1 \iff p(\{X = E[X]\}) = 1$$

che prova la tesi. ■

## 3. Disuguaglianza di Čebyšëv

Vogliamo dare una *versione quantitativa* della proposizione appena enunciata: ovvero vogliamo considerare *intorni* ad $E[X]$.

#Proposizione

> Proposizione (disuguaglianza di Čebyšëv).
>
> Sia $X$ una *variabile aleatoria discreta* con *momento secondo finito*. Allora vale la seguente disuguaglianza:

$$\forall \varepsilon > 0, \boxed{p(\{|X - E[X]| > \varepsilon\}) \leq \frac{\operatorname{var} X}{\varepsilon^2}}$$

Ovvero abbiamo che la probabilità che $X$ disti dalla media $E[X]$ è controllata da un valore che decresce sempre. In termini matematici,

$$\lim_{\varepsilon \to 0^+} p(\{|X - E[X]| > \varepsilon\}) = 0$$

#Dimostrazione

**DIMOSTRAZIONE** della Proposizione 4 (disuguaglianza di Čebyšëv)
Prendiamo la variabile aleatoria $Y$ definita come

$$Y := \varepsilon^2 \chi_{\{|X - E[X]| > \varepsilon\}}$$

ovvero se $X$ dista da $E[X]$ *"troppo poco"*, è nullo; altrimenti ho il valore $\varepsilon^2$.
Ho dunque due casi

$$\begin{cases} \omega \in \{|X - E[X]| > \varepsilon\} \implies Y(\omega) = \varepsilon^2 \leq (X(\omega) - E[X])^2 \\ \omega \in \{|X - E[X]| \leq \varepsilon\} \implies Y(\omega) = 0 \leq (X(\omega) - E[X])^2 \end{cases}$$

In ogni caso, ho sempre che il valore $Y$ è controllato da $(X - E[X])^2$ in alto. Dunque per la *monotonia del valor medio* (Proposizione 4 (monotonia del valor medio)) ho

$$\operatorname{var} X := E[(X - E[X]^2)] \geq E[Y] = \varepsilon^2 \cdot \underbrace{p\{|X - E[X]| > \varepsilon\}}_{E[\chi_A] = p(\{A\})}$$

Dividendo per $\varepsilon^2$ ottengo la tesi. ■

# 4. La Varianza della Somma di Variabili Aleatorie Discrete

Prima di dire il *comportamento della varianza* rispetto alla *somma*, vediamo se tale domanda è ben posta: ovvero se la somma delle variabili aleatorie ha momento secondo finito.

#Lemma

> Lemma (la somma delle variabili aleatorie ha momento secondo finito ).
>
> Siano $X_1, \ldots, X_N$ delle *variabili aleatorie discreti aventi momento secondo finito*. Allora anche la variabile aleatoria somma $Y := \sum_n^N X_n$ ha *momento secondo finito*.

#Dimostrazione

**DIMOSTRAZIONE** del Lemma 5 (la somma delle variabili aleatorie ha momento secondo

finito)
Si dimostra per $N=2$, poiché si generalizza su $\mathbb{N}$ per induzione. Si tratta di vedere $Y^2$ come la composizione $g \circ X$ (dove $X=(X_1, X_2)$), con $g(x_1, x_2) := (x_1+x_2)^2$. Alla per il teorema sulla composizione delle variabili aleatorie (Teorema 1 (composizione del valor medio)) si ha

$$(*) \quad \sum_{(x_1,x_2)\in\mathbb{R}^2} (x_1+x_2)^2 q(x)$$

Il nostro obbiettivo è di provare che questa è convergente. Definendo $q_1, q_2$ le densità associate a $X_1, X_2$ si ha

$$(**) \sum_{x\in\mathbb{R}} x^2 q_1(x) + \sum_{x\in\mathbb{R}} x^2 q_2(x) < +\infty$$

Quindi si tratta di ricavare (o maggiorare) da $(*)$ a $(**)$. Possiamo usare la maggiorazione $(x_1+x_2)^2 = x_1^2 + 2x_1x_2 + x_2^2 \leq 2(x_1^2+x_2^2)$ (per convincerci di questo basti pensare al quadrato del binomio). Allora si ha la catena delle disuguaglianze

$$\begin{aligned}\sum_{(x_1,x_2)\in\mathbb{R}^2} (x_1+x_2)^2 q(x) \leq & \sum_{(x_1,x_2)\in\mathbb{R}^2} 2(x_1^2+x_2^2) q(x_1,x_2) \\ & = 2 \cdot \sum_{(x_1,x_2)\in\mathbb{R}^2} x_1^2 q(x_1,x_2) + x_2^2 q(x_1 x_2)\end{aligned}$$

Adesso si tratta di spezzare la sommatoria in $\mathbb{R}^2$, dandoci

$$2 \cdot \sum_{(x_1,x_2)\in\mathbb{R}^2} x_1^2 q(x_1,x_2) + x_2^2 q(x_1 x_2) = 2\left(\underbrace{\sum_{x_1\in\mathbb{R}} x_1^2 q_1(x_1) + \sum_{x_2\in\mathbb{R}} x_2^2 q_2(x_2)}_{<+\infty}\right)$$

che prova la tesi. ■

#Osservazione

> Osservazione (calcolo della varianza).
>
> Abbiamo solamente dimostrato che *possiamo* calcolarci la varianza, ma non l'abbiamo ancora fatto. Adesso lo facciamo per $N=2$. In particolare, analizziamo la relazione tra la varianza della *somma* e delle *componenti individuali*.
>
> Notiamo preliminarmente che
>
> $$2X_1X_2 = (X_1+X_2)^2 - X_1^2 - X_2^2$$

Dato che tutti i membri a destra hanno *secondo momento finito*, lo ha pure $X_1 X_2$. Adesso per calcolare la varianza basta sfruttare la linearità di $E$

$$\begin{aligned}(X_1 + X_2 - E[X_1 + X_2])^2 &= (X_1 - E[X_1] + X_2 - E[X_2])^2 \\ &= (X_1 - E[X_1])^2 + (X_2 - E[X_2])^2 \\ &+ 2(E[X_1 X_2] - E[X_1]E[X_2])\end{aligned}$$

Portando tutto alla varianza, abbiamo

$$\boxed{\operatorname{var}(X_1 + X_2) = \operatorname{var} X_1 + \operatorname{var} X_2 + 2(E[X_1 X_2] - E[X_1]E[X_2])}$$

Notiamo che *queste due* non corrispondono perfettamente; infatti c'è l'eccesso alla fine. Diamo un significato a questo con la nozione di *covarianza*.

X

## 5. Definizione della Covarianza

### Definizione di Covarianza

X

*Definizione di covarianza. Definizione di variabili aleatorie scorrelate. Condizione sufficiente per variabili aleatorie scorrelate.*

X

## 0. Voci correlate

- Proprietà della Varianza

## 1. Definizione di Covarianza

**PREAMBOLO.** Dal calcolo della *covarianza* della somma di *due variabili aleatorie* $X, Y$ (Osservazione 6 (calcolo della varianza)) abbiamo ottenuto il risultato finale

$$\operatorname{var}(X_1 + X_2) = \operatorname{var} X_1 + \operatorname{var} X_2 + 2(\color{red}{E[X_1 X_2] - E[X_1]E[X_2]})$$

Notiamo che c'è questa parte in rosso che non è *eliminabile*: questo ci fa dire subito che *non sempre* vale la linearità sulla varianza. Ma cosa ci fa dire questa entità? Cosa significa? Cos'è? Diamone una definizione.

#Definizione

Definizione (covarianza di due variabili aleatorie).

Date due *discrete variabili aleatorie* $X_1, X_2$ aventi *momento secondo finito*, definiamo la *covarianza* di queste variabili aleatorie come

$$\operatorname{cov}(X_1, X_2) := E[X_1X_2] - E[X_1]E[X_2]$$

A seconda del valore assunto dalla *covarianza*, possiamo classificare il comportamento di due variabili aleatorie.

#Definizione

Definizione (variabili aleatorie scorrelate (o non correlate)).

Date due *variabili aleatorie discrete* $X_1, X_2$ aventi *momento secondo finito*, si dice che queste sono *scorrelate* se vale che la loro covarianza è nulla

$$\operatorname{cov}(X_1, X_2) = 0$$

# 2. Legame tra Variabili Aleatorie Scorrelate e Indipendenti

Poniamo questa osservazione sotto forma di proposizione

#Proposizione

Proposizione (indipendenza implica scorrelatezza).

Due *variabili aleatorie aventi momento secondo finito* $X_1, X_2$, se queste *sono indipendenti* allora essi sono *scorrelate*. Tuttavia *non* vale il viceversa!

#Dimostrazione

**DIMOSTRAZIONE** della Proposizione 3 (indipendenza implica scorrelatezza)
Basta usare il fatto che con l'*indipendenza* il prodotto *commuta con la media* (Proposizione 8 (prodotto tra valori medi di v.a. indipendenti)). Infatti abbiamo

$$E[X_1X_2] = E[X_1]E[X_2] \implies E[X_1X_2] - E[X_1]E[X_2] = \operatorname{cov}(X_1, X_2) = 0$$

che dimostra la tesi. Per dimostrare che non vale il viceversa, vedere l'esempio a seguito (Esempio 4 (non vale il viceversa)). ■

#Esempio

Esempio (non vale il viceversa).

Consideriamo una variabile aleatoria tale che

$$P\{X=-1\}=P\{X=0\}=P\{X=1\}$$

Poniamo poi $Y=X^2$, con

$$P\{Y=0\}=\frac{1}{3}, P\{Y=1\}=\frac{2}{3}$$

Dato che $E[X]=0$ e $X^3=X$, abbiamo che la covarianza tra $X, Y$ è nulla:

$$E[XY]=E[X^3]=E[X]=0 \implies \operatorname{cov}(X,Y)=0$$

Tuttavia queste non *sono indipendenti*! Infatti supponendo tale ci sarebbe l'assurdo per cui

$$P\{X=1, Y=0\}=P\{X=1, X^3=0\}=0$$

che contraddice col fatto che

$$P\{X=1\}P\{Y=0\}=\frac{1}{9}\neq 0$$

## 6. Proprietà della Covarianza

### Proprietà della Covarianza

X

*Proprietà della covarianza. Simmetria, bilinearità e stima.*

X

## 0. Voci correlate

- Definizione di Covarianza
- Proprietà della Varianza

## 1. Simmetria, bilinearità e condizione necessaria per varianza nulla

#Proposizione

Proposizione (prime proprietà della covarianza).

La covarianza gode delle proprietà della *simmetria* e *bilinearità*. Ovvero, date $X, Y$ delle variabili aleatorie aventi momento secondo finito e dati dei parametri $a, b \in \mathbb{R}$ abbiamo

$$
\begin{aligned}
& \operatorname{cov}(X, Y)=\operatorname{cov}(Y, X) \\
& \operatorname{cov}(a X+b Y, Z)=a \operatorname{cov}(X, Z)+b \operatorname{cov}(Y, Z)
\end{aligned}
$$

Inoltre abbiamo una *condizione sufficiente* per la *nullità della covarianza*. Infatti

$$
\operatorname{var} X=0 \Longrightarrow \operatorname{cov}(X, Y)=0
$$

#Dimostrazione

**DIMOSTRAZIONE** della Proposizione 1 (prime proprietà della covarianza)

Le prime due si dimostrano direttamente mediante per *definizione*. La terza si dimostra usando una della proprietà della varianza (Proposizione 3 (condizione necessaria e sufficiente per la nullità della varianza)), ovvero che la varianza è nulla se e solo se vale

$$
p\{X=E[X]\}=1
$$

poiché l'evento $\{X=E[X]\}$ è equivalente a $\{X-E[X]=0\}$ ed è incluso in $\{(X-E[X])(Y-E[Y])=0\}$ (la seconda parte non ci interessa, perché tanto la prima parte è tutta $\Omega$), abbiamo

$$
p\{(X-E[X])(Y-E[Y])=0\}=1
$$

Dunque, considerando $Z$ la variabile aleatoria $Z:=(X-E[X])(Y-E[Y])$ e applicandoci la stessa proprietà si ha

$$
E[Z]=0 \Longleftrightarrow E[(X-E[X])(Y-E[Y])]=: \operatorname{cov}(X, Y)=0
$$

che prova la tesi. ■

#Osservazione

Osservazione (la covarianza delle stesse variabili aleatorie).

Se prendiamo la *covarianza* di due stesse variabili aleatorie, ovvero calcoliamo $\operatorname{cov}(X, X)$ otteniamo la varianza stessa. Infatti con i calcoli si dimostra che

$$
\operatorname{cov}(X, X)=\operatorname{var}(X)
$$

## 2. Stima della Covarianza

Enunciamo una proprietà utile, dal punto di vista numerico (ovvero una versione *quantitativa* della terza proprietà)

Proposizione (la stima della covarianza).

Siano $X, Y$ delle *v.a. discrete aventi secondo momento finito*. Allora vale la stima

$$|\operatorname{cov}(X,Y)| \leq \sigma_x \sigma_y = \sqrt{\operatorname{var} X \operatorname{var} Y}$$

## 3. Applicazione delle Proprietà

#Osservazione

Osservazione (la generalizzazione della varianza delle somme).

Per generalizzare il risultato sulla varianza delle somme (1) basta considerare raggruppare *tutti* i termini in *due gruppi* e poi considerare le loro covarianze (per far ciò bisognerà usare la proprietà della covarianza!). Ad esempio abbiamo

$$\operatorname{var}(X+Y+Z) = \operatorname{var}(X+Y) + \operatorname{var} Z + 2\operatorname{cov}(X+Y, Z)$$

da cui, sfruttando la bilinearità, si ha

$$\ldots = \operatorname{var} X + \operatorname{var} Y + \operatorname{var} Z + 2(\operatorname{cov}(X,Y) + \operatorname{cov}(X,Z) + \operatorname{cov}(Y,Z))$$

Possiamo generalizzare il risultato per induzione su $N$, con la formula

$$\operatorname{var}\left(\sum_{n=1}^{N} X_n\right) = \sum_{n=1}^{N} \operatorname{var} X_n + \sum_{(k,j)\in\{1,\ldots,N\}^2 : k\neq j} \operatorname{cov}(X_k, X_j)$$

Notare che sulla seconda espressione non ci serve prendere la metà dato che la covarianza gode della *riflessività*.

X

## 7. Medie e Varianze Notevoli

### Definizione di Varianza e Deviazione Standard

X

*Idea della varianza. Definizione di varianza, deviazione standard, variabile aleatoria avente momento secondo finito.*

X

# 0. Voci correlate

- Definizione del Valore Medio
- Proprietà del Valore Medio
- Definizione di Variabile Aleatoria

# 1. Definizione di Varianza e Deviazione Standard

**IDEA.** Supponiamo di avere la varianza di una variabile aleatoria $E[X]$. Però questa informazione ci sembra piuttosto *"povera"*, vogliamo sapere di più di questa variabile aleatoria $X$: in particolare *come* si comporta rispetto alla sua media $E[X]$.
Un primo modo di quantificare quest'idea è di usare la *distanza euclidea* $|X - E[X]|$ e prendere il suo valor medio $E[|X - E[X]|]$; tuttavia presenta dei problemi, in quanto diventa difficile da gestire con valori assoluto.
Un altro approccio che è *"simile"* alla distanza euclidea è di prendere la loro differenza e di considerarne il *quadrato*; ovvero $(X - E[X])^2$. Certo, non è proprio l'idea iniziale ma si comporta similmente. Ad esempio, abbiamo $E[(X - E[X])^2] = 0 \iff X = 0$. Questa sarà l'idea di cui considereremo

#Definizione

**Definizione (varianza, deviazione standard).**

Sia $X$ una *variabile aleatoria discreta avente valor medio finito* $E[X] < +\infty$. Se $E[(X - E[X])^2] < +\infty$ (ovvero la quantità appena considerata ha valor medio finito), possiamo definire la sua varianza come

$$\boxed{\operatorname{var} X := E\left[(X - E[X])^2\right]}$$

Notiamo subito che la *varianza* è sempre positiva, dal momento che lo è pure il suo argomento (in azzurro).

Per non scomodarci con questo problema (infatti avvolte vogliamo sapere *"in che modo si discosta dalla media"*), prendiamo la sua radice e la chiamiamo *deviazione standard* (o *scarto quadratico medio*), definito come

$$\boxed{\sigma_X := \sqrt{\operatorname{var} X}}$$

#Osservazione

**Osservazione (l'esistenza di varianza implica l'esistenza della media).**

Notiamo che l'esistenza della varianza $\operatorname{var} X$ implica che esiste la media.

Infatti, $(X-E[X])^2 = X^2 - 2(XE[X]) + E[X]^2$, ci interessa sapere come si comporta $X$. Per il *teorema sulla composizione delle medie di variabili aleatorie* (1) deve seguire che

$$\sum_{x \in \mathbb{R}} x^2 q(x) < +\infty$$

Poiché $|x| \leq 1 + x^2$, si verifica subito che

$$\sum_{x \in \mathbb{R}} |x| q(x) \leq \sum_{x \in \mathbb{R}} (1+x^2) q(x) \leq 1 + \sum_{x \in \mathbb{R}} x^2 q(x) < +\infty$$

## 2. Valore medio avente momento secondo finito

Notiamo che non vale il viceversa: se una variabile aleatoria ha valor medio, non è detto che la media del suo quadrato sia finito. Definiamo una classe di variabili aleatorie che hanno sia *media finita* che il *valor medio del suo quadrato* finito.

#Definizione

Definizione (valor medio avente momento secondo finito).

Si dice che una variabile aleatoria $X$ ha *valor medio finito* se valgono che

$$E[X], E[X^2] < +\infty$$

In altre parole è una variabile aleatoria su cui *possiamo definire* la *varianza*.