• v.a. binomiale $X \sim Bin(n,p)$

$Im(X) = \{0, \dots, n\}$

$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k \in \{0, \dots, n\} = Im(X)$

$\mathbb{E}(X) = np \quad , \quad \mathbb{V}ar(X) = ?$

In generale:

$$\mathbb{V}(X) = \mathbb{E}\left([X - \mathbb{E}(X)]^2 \right)$$

$$= \mathbb{E}\left(X^2 - 2\,\mathbb{E}(X)\cdot X + \mathbb{E}(X)^2 \right)$$

linearità di $\mathbb{E}$

$$= \mathbb{E}(X^2) - 2\,\mathbb{E}(X)\cdot\mathbb{E}(X) + \mathbb{E}(X)^2$$

$$= \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Calcoliamo $\mathbb{E}(X^2)$ per $X \sim Bin(n,p)$

$$\mathbb{E}(X^2) = \sum_{k=0}^{n} k^2 \, \mathbb{P}(X=k) = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{\cancel{k=0}\; 1}^{n} k^{\cancel{2}} \frac{n!}{\cancel{k!}\,(k-1)!\,(n-k)!} p^k (1-p)^{n-k}$$

$h = k-1$

$$= \sum_{h=0}^{n-1} (h+1) \, n \frac{(n-1)!}{h!\,(n-1-h)!} \, p \, p^h (1-p)^{n-1-h}$$

distribuisco

$$= \underbrace{\sum_{h=0}^{n-1} h \, np \binom{n-1}{h} p^h (1-p)^{n-1-h}}_{= (n-1)p} + \underbrace{\sum_{h=0}^{n-1} np \binom{n-1}{h} p^h (1-p)^{n-1-h}}_{= 1}$$

$$= np\,(n-1)p + np = n(n-1)p^2 + np$$

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = n(n-1)p^2 + np - n^2p^2 = n\,p(1-p)$$

è giustamente invariante per $p \to 1-p$

- È capitato un fatto peculiare

$$X_i(\omega) = \begin{cases} 1 & \text{se } \omega_i = 1 \\ 0 & \omega_i = 0 \end{cases} \qquad i = 1, \dots, n$$

$X_i \sim Bern(p)$ e $X = \sum_{i=1}^{n} X_i$

giustamente, per linearità di $\mathbb{E}(\cdot)$,

$$np = \mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i) = \sum_{i=1}^{n} p = np$$

non si capisce come possa può succedere

$$np(1-p) = \mathbb{V}(X) \qquad \sum_{i=1}^{n} \mathbb{V}(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

la varianza non ha proprio modo di essere lineare c'è un quadrato.

Evidentemente, $\mathbb{V}(\cdot)$ è ~~omog~~ omogenea di grado due

$$X \to \alpha X \qquad \alpha \in \mathbb{R}$$

$$\mathbb{V}(X) \to \alpha^2 \, \mathbb{V}(X)$$

com'è possibile che ci sia capitato

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) \qquad ?$$

... questo capita quando X e Y sono v.a. <u>indipendenti</u>

- Seguiamo l'algebra. X, Y v.a. limitate (per non avere difficoltà di sommabilità)

$$\mathbb{V}(X+Y) = \mathbb{E}[(X+Y-\underbrace{\mathbb{E}(X+Y)}_{\mathbb{E}(X)+\mathbb{E}(Y)})^2]$$

$$= \mathbb{E}[(X-\mathbb{E}(X) + Y-\mathbb{E}(Y))^2]$$

$$= \mathbb{E}[(X-\mathbb{E}(X))^2 + (Y-\mathbb{E}(Y))^2 + 2(X-\mathbb{E}(X))\cdot(Y-\mathbb{E}(Y))]$$

$$= \mathbb{V}(X) + \mathbb{V}(Y) + 2\,\mathrm{cov}(X,Y)$$

ove

$$\mathrm{cov}(X,Y) := \mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))$$

<u>OSS</u>: $\mathbb{V}(X) = \mathrm{cov}(X,X)$

e da $\mathbb{V}(\cdot)$ recuperiamo $\mathrm{cov}(\cdot,\cdot)$ per polarizzazione, ovvero

$$\mathrm{cov}(X,Y) = \frac{1}{4}[\mathbb{V}(X+Y) - \mathbb{V}(X-Y)]$$

<u>DEF</u> <u>Indipendenza di v.a.</u>

X, Y v.a. discrete sono <u>indipendenti</u> sse

$$\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y)$$

$$\forall x \in \mathrm{Im}(X), \quad \forall y \in \mathrm{Im}(Y)$$

ovvero gli eventi

$X^{-1}(\{x\})$ e $Y^{-1}(\{y\})$ sono indipendenti $\forall x$ e $\forall y$

- Nel caso dello schema di Bernoulli

$$X_i = \begin{cases} 1 & \text{se } \omega_i = 1 \\ 0 & \text{se } \omega_i = 0 \end{cases}$$

e se $i \neq j$ $\quad X_i$ è indip da X_j

(risultati di lanci diversi)

<u>Prop</u> se X, Y v.a. indipendenti allora

$$\mathrm{cov}(X, Y) = 0$$

come si dice,

indipendenti $\Rightarrow$ scorrelate $\quad (\mathrm{cov}(X,Y) = 0)$

<u>OSS 1</u> Per l'algebra di prima

$$\mathrm{cov}(X,Y) = 0 \iff \mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$$

~~Dimm~~

<u>OSS 2</u>

$$\mathrm{cov}(X,Y) = \mathbb{E}\big((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\big)$$

linearità $\mathbb{E}(\cdot)$

$$= \mathbb{E}(XY) - 2\,\mathbb{E}(X)\cdot\mathbb{E}(Y) + \mathbb{E}(X)\,\mathbb{E}(Y)$$

$$= \mathbb{E}(XY) - \mathbb{E}(X)\cdot\mathbb{E}(Y)$$

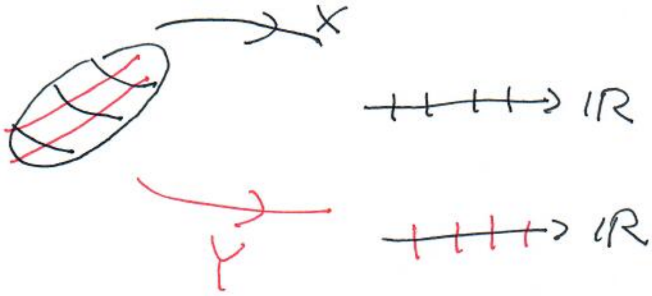
quindi

$$\mathrm{cov}(X,Y) = 0 \iff \mathbb{E}(X\cdot Y) = \mathbb{E}(X)\,\mathbb{E}(Y)$$

<u>dim</u> verifichiamo che X, Y indip $\Rightarrow \mathbb{E}(XY) = \mathbb{E}(X)\cdot\mathbb{E}(Y)$

Calcoliamo $\mathbb{E}(\cdot)$ sommando su $\omega \in \Omega$

$$\mathbb{E}(XY) = \sum_{\omega\in\Omega} X(\omega)\,Y(\omega)\;\mathbb{P}(\{\omega\})$$



$X^{-1}(\{x\}),\ x \in Im(X)$

$Y^{-1}(\{y\}),\ y \in Im(Y)$ sono partizioni di Ω

$$\sum_{\omega\in\Omega} \dots = \sum_{\substack{x \in Im(X)\\ y \in Im(Y)}} \ \sum_{\omega \in X^{-1}(\{x\}) \cap Y^{-1}(\{y\})} \dots$$

trovo

$$\mathbb{E}(X\cdot Y) = \sum_{\substack{x \in Im(X)\\ y \in Im(Y)}} \ \sum_{\omega \in X^{-1}(\{x\}) \cap Y^{-1}(\{y\})} \overbrace{X(\omega)}^{x}\ \underbrace{Y(\omega)}_{y}\ \mathbb{P}(\{\omega\})$$

$$= \sum_{\substack{x \in Im(X)\\ y \in Im(Y)}} x y \underbrace{\sum_{\omega \in X^{-1}(\{x\}) \cap Y^{-1}(\{y\})} \mathbb{P}(\{\omega\})}_{}$$

$$\parallel$$

$$\mathbb{P}(X^{-1}(\{x\}) \cap Y^{-1}(\{y\}))$$

$$\parallel$$

$$\mathbb{P}(X=x,\ Y=y)$$

(indipendenza) $\longrightarrow$ $\parallel$

$$\cancel{\mathbb{P}((X}\ \mathbb{P}(X=x)\cdot\mathbb{P}(Y=y)$$

$$= \sum_{\substack{x \in Im(X)\\ y \in Im(Y)}} x y\ \mathbb{P}(X=x)\ \mathbb{P}(Y=y)$$

$$= \mathbb{E}(X)\cdot\mathbb{E}(Y)$$

- Sunto proprietà di $\mathbb{E}(\cdot)$ e $\mathbb{V}(\cdot)$

$A \subset \Omega$

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \in A^c \end{cases}$$ è una v.a.

$(\mathbb{1}_A \sim \text{Bern}(\mathbb{P}(A)))$

- $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A) \qquad A \subset \Omega$

- $\mathbb{E}(\alpha X + \beta Y) = \alpha\,\mathbb{E}(X) + \beta\,\mathbb{E}(Y) \qquad \alpha, \beta \in \mathbb{R}$

  $\mathbb{E}(\cdot)$ è lineare

- $X \leq Y$ (ovvero $X(\omega) \leq Y(\omega) \quad \forall \omega \in \Omega$)

  ordine di $\mathbb{R}$

  $\Rightarrow \quad \mathbb{E}(X) \leq \mathbb{E}(Y)$

- $\mathbb{V}(X) \geq 0 \qquad = 0$ sse X è v.a. certa

- $\mathbb{V}(\alpha X + \beta) = \alpha^2\,\mathbb{V}(X) \qquad \alpha, \beta \in \mathbb{R}$

  v.a. certa

- X, Y indip $\Rightarrow \mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$

  $\Uparrow$ $\text{cov}(X,Y) = 0$

- $|\text{cov}(X,Y)| \leq \sqrt{\mathbb{V}(X)} \cdot \sqrt{\mathbb{V}(Y)}$

  può essere > 0 o < 0

l'unica da dimostrare è l'ultima. In effetti è Cauchy-Schwartz, dipende da una struttura euclidea.

se $(\cdot,\cdot)$ è un prodotto scalare

$$|(x,y)| \leq \sqrt{(x,x)}\,\sqrt{(y,y)}$$

- Variabili aleatorie come spazio euclideo

$(\Omega, \mathcal{F}, \mathbb{P})$ $\quad |\Omega| < +\infty \quad |\Omega| = n$

$\mathcal{F} = \{A \subset \Omega\}$

$\mathcal{V} = \{X : \Omega \to \mathbb{R}\} = \{\text{v.a su } \Omega\}$ è uno spazio vettoriale su $\mathbb{R}$

lo dotiamo di una struttura euclidea, ovvero di un prodotto scalare

Non è il prodotto scalare canonico su $\mathbb{R}^n$, ma dipende da $\mathbb{P}$.

$L^2(\Omega, \mathbb{P}) = \{X : \Omega \to \mathbb{R}\}$ con il prodotto scalare def. da

$$(X, Y) = \sum_{\omega \in \Omega} X(\omega) Y(\omega) \mathbb{P}(\{\omega\})$$

bilineare e $(X, X) \geq 0$ $= 0$ sse $X = 0$
$(X(\omega) = 0 \ \forall \omega \in \Omega)$

OSS Se $\mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}$ Prob. uniforme
è il prodotto scalare canonico su $\mathbb{R}^n$ (a meno di una costante)

Diseg. di Cauchy-Schwartz

$$|(X, Y)| \leq |X| |Y|$$

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)} \sqrt{\mathbb{E}(Y^2)}$$

se $X \to X - \mathbb{E}(X)$
$Y \to Y - \mathbb{E}(Y)$
è l'ultima proprietà di prima.

- Indipendenza di tante v.a.

Per gli eventi era un po' complicato

$A_1, \dots, A_n$ eventi sono indipendenti

sse $\forall k \leq n$ $\quad i_1, i_2, \dots, i_k$

$$\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_k})$$

per v.a. è più facile

[da scrivere, in realtà è la stessa cosa]

<u>DEF</u> <u>equivalente</u>

X, Y v.a. (sia esse discrete o continue)

sono indipendenti sse $\forall$ $A \subset \mathbb{R}$, $B \subset \mathbb{R}$

$$\mathbb{P}(\underbrace{X \in A, Y \in B}_{\{\omega:\, X(\omega) \in A\} \cap \{\omega:\, X(\omega) \in B\}}) = \mathbb{P}(X \in A) \cdot \mathbb{P}(X \in B)$$

infatti dalla definizione precedente [X, Y discrete]

(additività)

$$\mathbb{P}(X \in A, Y \in B) = \sum_{\substack{x \in \mathrm{Im}(X) \cap A \\ y \in \mathrm{Im}(Y) \cap B}} \mathbb{P}(X = x, Y = y)$$

(definizione precedente)

$$= \sum_{\substack{x \in \dots \\ y \in \dots}} \mathbb{P}(X = x)\, \mathbb{P}(Y = y) = \mathbb{P}(X \in A) \cdot \mathbb{P}(X \in B)$$

<u>Def</u> $X_1, \dots, X_n$ v.a.

sono <u>indipendenti</u> sse

$$\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1) \cdots \mathbb{P}(X_n \in A_n)$$

$$\forall\, A_1, \dots, A_n \subset \mathbb{R}$$

<u>OSS</u> Se prendo $A_i = \mathbb{R}$ per qualche i

ricavo che

$X_{i_1}, \dots, X_{i_k}$ $\qquad \{i_1, \dots, i_k\} \subset \{1, \dots, n\}$

sono anche indipendenti

[per gli eventi lo dovevo chiedere in più]

Nel caso di schema di Bernoulli

[o in genere per probabilità prodotto]

$\Omega = \{0,1\}^n = \{(\omega_1, \dots, \omega_n) \quad \omega_i \in \{0,1\}\}$ $\qquad \mathbb{P}$ prodotto

se X_i si riferisce ~~alla~~ solo al lancio i

$$(X_i(\omega) = \omega_i)$$

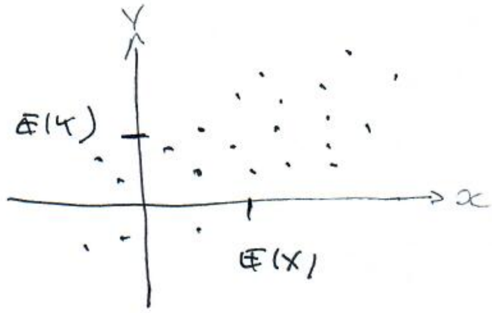
$X_1, \dots, X_n$ sono v.a. indipendenti ~~con~~

- Significato intuitivo di $cov(\cdot,\cdot)$

X, Y v.a. Fai un po' di esperimenti e misura X e Y

Riporta sul piano cartesiano

(I)



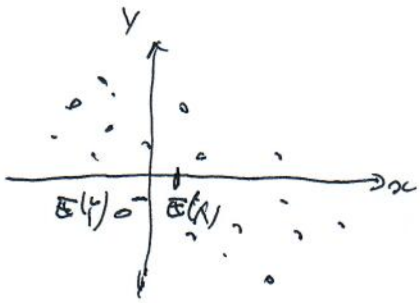
se così vuol dire che se $X > \mathbb{E}(X)$ anche $Y > \mathbb{E}(Y)$ (di solito)

quindi $cov(X,Y) > 0$

X e Y sono positivamente correlate

se ti è capitata una fluttuation $X > \mathbb{E}(X)$ è più probabile che anche Y ~~#~~ $> \mathbb{E}(Y)$

(II)



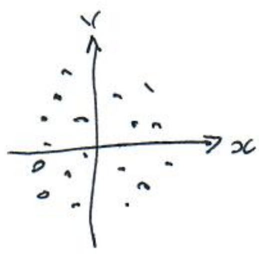
qui è il contrario

$X > \mathbb{E}(X)$ è facile $Y < \mathbb{E}(Y)$

$cov(X,Y) < 0$

X e Y sono negativamente correlati

(III)



in questo caso non sembra esserci correlazione tra X, Y

$cov(X,Y) \approx 0$