Indipendenza e covarianza (incorrelazione):

1) Se $X, Y \in \mathcal{L}^1(\Omega, P)$ sono indipendenti, allora

$$\begin{aligned} cov(X,Y) &= E[(X-E[X])\cdot(Y-E[Y])] \\ &\overset{\text{indip.}}{=} E[X-E[X]]\cdot E[Y-E[Y]] \\ &= 0. \end{aligned}$$

2) Se $X_1, \ldots, X_n \in \mathcal{L}^2(\Omega, P)$ sono indipendenti a due a due, allora

$$var\Big(\sum_{i=1}^n X_i\Big) = \sum_{i=1}^n var(X_i).$$

In particolare: X, Y indipendenti (e in $\mathcal{L}^1$) implica X, Y incorrelate.

Attenzione: In generale, non vale l'implicazione inversa.

V.a. incorrelate non necessariamente indipendenti.

(Contro-) Esempio:

Sia Z una v.a. su (Ω, P) a valori in $\{0, \frac{\pi}{2}, \pi\}$

con $P_Z = Unif(\{0, \frac{\pi}{2}, \pi\})$.

Poniamo $X \doteq \cos(Z)$, $\quad Y \doteq \sin(Z)$.

Allora $E[X] = \frac{1}{3} \cdot (1 + 0 - 1) = 0$,

$E[Y] = \frac{1}{3} \cdot (0 + 1 + 0) = \frac{1}{3}$,

$\leadsto$ $\operatorname{cov}(X, Y) = E[X \cdot (Y - \frac{1}{3})]$

| X, Y funzioni di Z

$= \frac{1}{3} \cdot \left(1 \cdot (0 - \frac{1}{3}) + 0 \cdot (1 - \frac{1}{3}) + (-1) \cdot (0 - \frac{1}{3}) \right)$

$= \frac{1}{3} \cdot \left(-\frac{1}{3} + \frac{1}{3} \right) = 0$.

$\leadsto$ X, Y incorrelate.

Ma X, Y <u>non</u> indipendenti. Ad esempio,

$P(X = 1) = \frac{1}{3} > 0$, $\quad P(Y = 1) = \frac{1}{3} > 0$,

mentre $P(X = 1, Y = 1) = 0 \neq P(X = 1) \cdot P(Y = 1)$.

//

Applicazione alla media empirica:

Siano $X_1, \dots, X_n$ v.a. in $L^2(\Omega, P)$ identicamente distribuite (cioè $P_{X_1} = \dots = P_{X_n}$) e indipendenti a due a due.

Poniamo $\bar{S}_n \doteq \frac{1}{n} \sum_{i=1}^{n} X_i$

$\rightsquigarrow$ $\bar{S}_n$ media empirica di $X_1, \dots, X_n$

Valor medio e varianza di $\bar{S}_n$:

- $$E[\bar{S}_n] \overset{\text{linearità}}{=} \frac{1}{n} \sum_{i=1}^{n} E[X_i] \overset{\text{identi. distr.}}{=} \frac{1}{n} \sum_{i=1}^{n} E[X_1] = \frac{n}{n} \cdot E[X_1]$$

$\rightsquigarrow$ $E[\bar{S}_n] = E[X_1]$.

implica incorrelate a due a due

- $$\text{var}(\bar{S}_n) = \frac{1}{n^2} \cdot \text{var}\left(\sum_{i=1}^{n} X_i\right) \overset{\substack{\text{indip.} \\ \text{a 2 a 2}}}{=} \frac{1}{n^2}\left(\sum_{i=1}^{n} \text{var}(X_i)\right)$$

$$\overset{\text{identi. distr.}}{=} \frac{1}{n^2}\left(\sum_{i=1}^{n} \text{var}(X_1)\right) = \frac{n}{n^2} \text{var}(X_1)$$

$\rightsquigarrow$ $\text{var}(\bar{S}_n) = \frac{1}{n} \cdot \text{var}(X_1)$.

Nota: $\text{var}(\bar{S}_n) \xrightarrow{n \to \infty} 0$.

Teorema (disuguaglianza di Markov-Chebyshev):

Sia X v.a. reale su (Ω, P).

1) Se $X \geq 0$, allora per ogni $\varepsilon > 0$

Markov
$$P(X \geq \varepsilon) \leq \frac{E[X]}{\varepsilon}.$$

2) Sia $f : [0,\infty) \to [0,\infty)$ una funzione crescente
con $f(x) > 0$ per $x > 0$.
Se $X \geq 0$, allora per ogni $\varepsilon > 0$,

$$P(X \geq \varepsilon) \leq \frac{E[f(X)]}{f(\varepsilon)}.$$

3) Se $X \in L^2(\Omega, P)$. allora per ogni $\varepsilon > 0$,

$$P(|X - E[X]| \geq \varepsilon) \leq \frac{var(X)}{\varepsilon^2}.$$

Dim.: 3) segue da 2) con $f(x) = x^2$.

2) e 1) sono equivalenti:

$$\{X \geq \varepsilon\} = \{f(X) \geq f(\varepsilon)\}$$ poiché f crescente

1): Sia $\varepsilon > 0$. $P(X \geq \varepsilon) = E[\mathbb{1}_{[\varepsilon,\infty)}(X)]$

$$\leq E\left[\frac{X}{\varepsilon}\right] = \frac{E[X]}{\varepsilon}.$$

//

Osservazione:

La disuguaglianza di Chebyshev (parte 3) del Teorema) dà un significato alla deviazione standard:

Sia $X \in L^2(\Omega, P)$. Poniamo $\sigma^2 \doteq var(X)$

$\rightsquigarrow$ σ deviazione standard.

Supponiamo $\sigma^2 > 0$. Allora per la disuguaglianza di Chebyshev con $\varepsilon = c \cdot \sigma$, dove $c > 0$:

$$P(|X - E[X]| \geq c \cdot \sigma) \leq \frac{var(X)}{c^2 \cdot \sigma^2} = \frac{1}{c^2}.$$

In particolare,

deviazioni dal valor medio $\leq \sigma$ sono "da aspettarsi";

deviazioni più grandi di $5 \cdot \sigma$ hanno probabilità $\leq 4\%$.