Tutorato V
Probabilità e Statistica
a.a. 2015/2016

Argomenti: momenti; varianza e covarianza; interpretazione della media condizionale; retta di regressione.

Esercizio 1. Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilià.

a) Fissato $k = 1, 2, \ldots$, sia

$$L_d^k(\Omega, \mathcal{F}, \mathbb{P}) = \{X : \Omega \to \mathbb{R} \text{ t.c. } X \text{ è v.a. discreta ed esiste } \mathbb{E}(|X|^k)\}.$$

Dimostrare che $L_d^k(\Omega, \mathcal{F}, \mathbb{P})$ è uno spazio vettoriale e che per ogni $1 \leq r \leq k$ si ha $L_d^k(\Omega, \mathcal{F}, \mathbb{P}) \subset L_d^r(\Omega, \mathcal{F}, \mathbb{P})$.

b) Consideriamo il caso $k = 2$. Definiamo

$$\langle \cdot, \cdot \rangle : (X, Y) \in L_d^2(\Omega, \mathcal{F}, \mathbb{P}) \times L_d^2(\Omega, \mathcal{F}, \mathbb{P}) \mapsto \langle X, Y \rangle = \mathbb{E}(XY) \in \mathbb{R}$$

Dimostrare che $\langle \cdot, \cdot \rangle$ definisce un **prodotto scalare**[1] su $L_d^2(\Omega, \mathcal{F}, \mathbb{P})$ e quindi

$$|| \cdot || : X \in L_d^2 \mapsto ||X|| = \sqrt{\mathbb{E}(X^2)}$$

è una **norma** su $L_d^2(\Omega, \mathcal{F}, \mathbb{P})$.

Esercizio 2. Dimostrare che per ogni scelta di $X, Y, X_1, X_2 \in L_d^2(\Omega, \mathcal{F}, \mathbb{P})$,

1) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;
2) $\mathrm{Cov}(X_1 + X_2, Y) = \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y)$;
3) $\mathrm{Cov}(aX, Y) = a\mathrm{Cov}(X, Y)$, per ogni $a \in \mathbb{R}$;
4) $\mathrm{Cov}(X, X) = \mathrm{Var}(X) \geq 0$ e $\mathrm{Var}(X) = 0$ se e solo se $X = \alpha$ per qualche $\alpha \in \mathbb{R}$.
5) $\mathrm{Cov}(X, \alpha) = 0$, per ogni $\alpha \in \mathbb{R}$.

Dedurre che la covarianza è un'applicazione **simmetrica** e **bilineare**:

$$\begin{aligned}\mathrm{Cov}(aX_1 + bX_2, cY_1 + dY_2) =& ac\mathrm{Cov}(X_1, Y_1) + ad\mathrm{Cov}(X_1, Y_2) + \\ &+ bc\mathrm{Cov}(X_2, Y_1) + bd\mathrm{Cov}(X_2, Y_2).\end{aligned}$$

per ogni $a, b, c, d \in \mathbb{R}$. Ma $\mathrm{Cov}(X, Y)$ non definisce un prodotto scalare su L_d^2: perché?

Esercizio 3. La disuguaglianza di Cauchy-Schwarz dice: date X, Y con momento secondo finito allora $|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$. Dimostrare che vale l'uguaglianza se e solo se $XY = 0$ oppure $X = \lambda Y$, per qualche $\lambda \neq 0$, $\lambda \in \mathbb{R}$ (da determinare).
[Sugg: si lavori con $\mathbb{E}((X + \theta Y)^2)$ al variare di $\theta \in \mathbb{R}$ e ricordare che se $\mathbb{E}(Z^2) = 0$ allora $Z = 0$.]

[1] Ricordiamo che se V è uno spazio vettoriale, $\langle \cdot, \cdot \rangle : (v, w) \in V \times V \mapsto \langle v, w \rangle \in \mathbb{R}$ è un prodotto scalare se valgono le seguenti proprietà: 1. (simmetria) $\langle v, w \rangle = \langle w, v \rangle$; 2. (linearità) $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$; 3. (omogeneità) per $\lambda \in \mathbb{R}$, $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$; 4. (positività e non degenerazione) $\langle v, v \rangle \geq 0$ e $\langle v, v \rangle = 0$ se e solo se $v = 0$.

Esercizio 4. Siano X e Y due v.a. discrete. Indichiamo con E_Y l'insieme dei valori che può assumere Y e con p_Y la sua densità discreta. Definiamo $\varphi : E_Y \to \mathbb{R}$ come:

$$\varphi(y) = \mathbb{E}(X \mid Y = y) \text{ quando } p_Y(y) > 0 \text{ e } \varphi(y) = 0 \text{ altrimenti.}$$

Supponiamo che X abbia speranza matematica finita.

a) Dimostrare che anche $\varphi(Y)$ ha speranza matematica finita e si ha $\mathbb{E}(\varphi(Y)) = \mathbb{E}(X)$.

D'ora in poi, supporemo che X abbia anche momento secondo finito.

b) Dimostrare che $\varphi(Y)$ ha momento secondo finito.

[Sugg.: ricordiamo che, data una qualsiasi v.a. Z, si ha $|\mathbb{E}(Z)|^2 \leq \mathbb{E}(|Z|^2)$, cioè $|\sum_z z p_Z(z)|^2 \leq \sum_z |z|^2 p_Z(z)$.]

c) Sia $g : E_Y \to \mathbb{R}$ una funzione tale che $g(Y)$ abbia momento secondo finito. Dimostrare che $\mathbb{E}(g(Y)\varphi(Y)) = \mathbb{E}(g(Y)X)$.

d) Sia $\psi : E_Y \to \mathbb{R}$ una funzione tale che $\psi(Y)$ abbia momento secondo finito. Dimostrare che $\mathbb{E}(|\psi(Y) - X|^2) \geq \mathbb{E}(|\varphi(Y) - X|^2)$.

[Sugg.: potrebbe essere utile scrivere $\mathbb{E}(|\psi(Y) - \varphi(Y)|^2) = \mathbb{E}(|(\psi(Y) - X) + (X - \varphi(Y))|^2)$, sviluppare il quadrato, usare la linearità ed il punto **c)**...]

[Oss.: l'esercizio 4 dà un sacco di informazioni sulla media condizionale (di carattere generale, validi cioè non solo per v.a. discrete). **a)** dimostra che la media della media condizionata è la media stessa, cioè, in media, la media condizionata di X dato Y si comporta proprio come la media di X. Inoltre, da **c)** segue che la media condizionale di X dato Y rappresenta la funzione di Y che "meglio approssima" la v.a. X nella classe delle v.a. che sono funzioni di Y e che hanno momento secondo finito.]

Esercizio 5. Siano X e Y due v.a. bernoulliane di parametro p e $\hat{p}$ rispettivamente. Dimostrare che X e Y sono indipendenti se e solo se $\mathrm{Cov}(X, Y) = 0$. Generalizzare tale proprietà per X e Y v.a. che possono assumere solo due valori: $E_X = \{x_1, x_2\}$ e $E_Y = \{y_1, y_2\}$, con $x_1 \neq x_2$ e $y_1 \neq y_2$.

[Sugg.: Per la seconda parte, si può tenere conto delle v.a. ausiliarie $U = \frac{X - x_1}{x_2 - x_1}$ e $V = \frac{Y - y_1}{y_2 - y_1}$.]

Esercizio 6. Siano X e Y due v.a. a valori in $E_X = \{-2, 1\}$ e $E_Y = \{-1, 0, 2\}$ rispettivamente, con distribuzione congiunta descritta tramite la seguente tabella:

	$X = -2$	$X = 1$	
$Y = -1$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$
$Y = 0$	$\frac{1}{12}$	$\frac{5}{12}$	$\frac{1}{2}$
$Y = 2$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$
	$\frac{1}{3}$	$\frac{2}{3}$	1

Mostrare che $\mathrm{Cov}(X, Y) = 0$ e che X e Y non sono indipendenti. Dedurre che, in generale, covarianza nulla **non** implica indipendenza.

Esercizio 7. Siano X e Y v.a. correlate ($\mathrm{Cov}(X,Y) \neq 0$) e sia $y = g(x)$ l'equazione della retta di regressione di Y rispetto a X:

$$g(x) = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}\big(x - \mathbb{E}(X)\big) + \mathbb{E}(Y).$$

Indichiamo con ψ la funzione inversa di g: $\psi(y) = g^{-1}(y)$. Si noti che ψ è lineare, quindi l'equazione $x = \psi(y)$ definisce una retta. È vero (sempre, qualche volta, mai) che ψ è la retta di regressione di X rispetto a Y?

Esercizio 8. Siano X e Y due v.a. indipendenti t.c. $X,Y \sim \mathrm{Po}(\lambda)$ e $\alpha, \beta \in \mathbb{R}$.

a) Calcolare la covarianza tra X e $Z^{\alpha,\beta} = \alpha\, X + \beta\, Y$. Stabilire il tipo di dipendenza tra X e $Z^{\alpha,\beta}$ al variare di $\alpha, \beta \in \mathbb{R}$.

[Sugg.: non utilizzare la distribuzione congiunta di X e $Z^{\alpha,\beta}$; piuttosto, si consideri che $Z^{\alpha,\beta} = \alpha\, X + \beta\, Y$...]

b) Disegnare la retta di regressione di $Z^{\alpha,\beta}$ rispetto a X.

Esercizio 9. Un dado equo viene lanciato due volte. Sia X il risultato del primo lancio e Y il massimo risultato ottenuto nei due lanci.

a) Calcolare la distribuzione congiunta e le distribuzioni marginali di X e Y.

[Sugg.: potrebbe essere utile ricordare che:
$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$, $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$, $\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2$.]

b) Calcolare $\mathrm{Cov}(X,Y)$. Che tipo di dipendenza c'è tra X e Y?

c) Disegnare la retta di regressione di Y rispetto a X.

d) Stimare il numero di volte in cui occorre lanciare il dado affinché con probabilità minore di 0.2 la media aritmetica dei risultati disti da 3.5 (cioè dal valore atteso) per più di 0.1.

Esercizio 10. Fissate X_1, X_2, X_3 v.a. i.i.d. con distribuzione

$$p(1) = \frac{1}{2} = 1 - p(-1),$$

siano $U = X_1\, X_2$ e $V = X_1\, X_3$.

a) Calcolare la retta di regressione di V rispetto a U.

b) Le v.a. U e V sono indipendenti?

c) Calcolare la distribuzione congiunta di U e V.

Esercizio 11. Sia X una v.a. con densità discreta

$$p_X(x) = \begin{cases} \frac{1}{2} \cdot (1-p)^{|x|-1} p & \text{se } x = \pm 1, \pm 2, \pm 3, \dots \\ 0 & \text{altrimenti} \end{cases}$$

Definiamo $Z = \mathbb{1}_{\{X \geq 0\}} - \mathbb{1}_{\{X < 0\}}$ e $W = |X|$.

a) Calcolare la legge congiunta di Z e W. Z e W sono indipendenti?

b) Calcolare la legge e la media condizionale di W dato $Z = z$.

c) Studiare la dipendenza tra X e Z; disegnare la retta di regressione di Z rispetto a X.

d) Calcolare la media condizionale di X dato che $Z = -1$.

Esercizio 12. Una moneta che dà testa con probabilità $p \in (0, 1)$ viene lanciata 6 volte. Sia $X_i = 1\!\!1_{A_i}$, dove A_i denota l'evento *esce testa all'i-esimo lancio*, $i = 1, \dots, 6$. Definiamo:

$$Y = \sum_{i=1}^{2} X_i, \quad Z = \sum_{i=3}^{4} X_i, \quad W = \sum_{i=5}^{6} X_i, \quad T = Y + Z, \quad U = Z + W.$$

a) Calcolare la covarianza tra T e U. T e U sono indipendenti?

b) Calcolare la media e la varianza di T.

c) Calcolare la media e la varianza di $2T + 1$.

d) Calcolare la covarianza tra $2T + 1$ e $3U$.

Soluzioni

Esercizio 1. a) Usiamo qui alcuni risultati visti a lezione.
Sappiamo che se X e Y hanno momento $k-$esimo allora anche $X+Y$ ha momento di ordine k, così come αX e βY per ogni $\alpha, \beta \in \mathbb{R}$. Quindi L^2_d è uno spazio vettoriale. Sappiamo anche che se X ha momento k-esimo finito allora ha momento di ordine r per ogni $r \leq k$, quindi $L^k_d \subset L^r_d$.

b) Occorre verificare che

1. (simmetria) per ogni $X, Y \in L^2_d$, $\langle X, Y\rangle = \langle Y, X\rangle$: ovvio;

2. (linearità) per ogni $X_1, X_2, Y \in L^2_d$, $\langle X_1 + X_2, Y\rangle = \langle X_1, Y\rangle + \langle X_2, Y\rangle$: $\langle X_1 + X_2, Y\rangle = \mathbb{E}[(X_1 + X_2)Y] = \mathbb{E}(X_1 Y) + \mathbb{E}(X_2 Y) = \langle X_1, Y\rangle + \langle X_2, Y\rangle$;

3. (omogeneità) per ogni $\lambda \in \mathbb{R}$ e $X \in L^2_d$, $\langle \lambda X, Y\rangle = \lambda\langle X, Y\rangle$: $\langle \lambda X, Y\rangle = \mathbb{E}(\lambda X\, Y) = \lambda\mathbb{E}(XY) = \lambda\langle X, Y\rangle$;

4. (positività e non degenerazione) per ogni $X \in L^2_d$, $\langle X, X\rangle \geq 0$ e $\langle X, X\rangle = 0$ se e solo se $X = 0$: $\langle X, X\rangle = \mathbb{E}(X^2) \geq 0$ e se $X = 0$ ovviamente $\langle X, X\rangle = 0$; inoltre, se $\langle X, X\rangle = \mathbb{E}(X^2) = 0$ allora

$$\sum_{x\in E_X} x^2 p_X(x) = 0.$$

La somma su scritta è a termini non negativi, quindi il risultato è nullo se e solo se tutti i termini della somma sono nulli: l'unico elemento in E_X può essere solo lo 0, cioè $X = 0$.

Le 1.–4. provano che $\langle\cdot,\cdot\rangle$ è un prodotto scalare su L^2_d. È ben noto allora che $||X|| = \sqrt{\langle X, X\rangle} = \sqrt{\mathbb{E}(X^2)}$ definisce una norma su L^2_d.

Esercizio 2. La verifica delle **1)**–**4)** è immediata conseguenza delle proprietà del prodotto scalare definito nell'esercizio 1: basta infatti ricordare che, per ogni $X, Y \in L^2_d$,

$$\mathrm{Cov}(X, Y) = \langle X - \mathbb{E}(X), Y - \mathbb{E}(Y)\rangle = \langle TX, TY\rangle$$

dove $T : X \in L^2_d \mapsto TX = X - \mathbb{E}(X) \in L^2_d$ è un operatore lineare di L^2_d.
Per **5)**, basta ricordare che, per ogni $\alpha \in \mathbb{R}$, le v.a. X e $Y = \alpha$ sono indipendenti, quindi $\mathrm{Cov}(X, \alpha) = 0$. Oppure, per verifica diretta, $\mathrm{Cov}(X, \alpha) = \mathbb{E}(\alpha X) - \mathbb{E}(\alpha)\mathbb{E}(X) = \alpha\mathbb{E}(X) - \alpha\mathbb{E}(X) = 0$.
L'applicazione Cov su $L^2_d \times L^2_d$ non definisce un prodotto scalare perché salta la non degenerazione: se $\mathrm{Cov}(X, X) = 0$ allora $\mathrm{Var}(X) = 0$ e quindi $X = \alpha$, ma non è detto che $\alpha = 0$.

Esercizio 3. Il caso $XY = 0$ è banale (perché significa che o X oppure Y oppure entrambe sono identicamente nulle), quindi lo lasciamo da parte e supponiamo nel seguito che $XY \neq 0$, cioè sia X che Y sono v.a. non identicamente nulle.
Supponiamo che $X = \lambda Y$, per qualche $\lambda \in \mathbb{R}$. Allora $\mathbb{E}(XY) = \lambda\mathbb{E}(Y^2)$ e $\mathbb{E}(X^2)\mathbb{E}(Y^2) = \lambda^2\mathbb{E}(Y^2)^2$, e si ha banalmente $|\mathbb{E}(XY)| = \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$.
Viceversa, supponiamo che $|\mathbb{E}(XY)| = \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$, cioè 1) $\mathbb{E}(XY) = \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$ oppure 2) $\mathbb{E}(XY) = -\sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$. Preso $\theta \in \mathbb{R}$, si ha

$$\mathbb{E}((X + \theta Y)^2) = \mathbb{E}(X^2 + \theta^2 Y^2 + 2\theta XY) = \mathbb{E}(X^2) + \theta^2\mathbb{E}(Y^2) + 2\theta\mathbb{E}(XY).$$

Nel caso 1), si ha

$$\mathbb{E}((X+\theta Y)^2) = \mathbb{E}(X^2) + \theta^2\mathbb{E}(Y^2) + 2\theta\sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)} = \big(\sqrt{\mathbb{E}(X^2)} + \theta\sqrt{\mathbb{E}(Y^2)}\big)^2.$$

Preso $\theta = -\sqrt{\mathbb{E}(X^2)}/\sqrt{\mathbb{E}(Y^2)}$, otteniamo $\mathbb{E}((X+\theta Y)^2)$ e quindi $X+\theta Y = 0$, da cui $X = -\theta Y$. Se invece siamo nel caso 2), cioè $\mathbb{E}(XY) = -\sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$, gli stessi ragionamenti danno ancora $X = \lambda Y$ con $\lambda = \sqrt{\mathbb{E}(X^2)}/\sqrt{\mathbb{E}(Y^2)}$.

Esercizio 4. **a)** Sappiamo che $\varphi(Y)$ ha media se e solo se $\sum_y |\varphi(y)|p_Y(y) < \infty$ e in tal caso, $\mathbb{E}(\varphi(Y)) = \sum_y \varphi(y)p_Y(y)$. Si ha:

$$\begin{aligned}\sum_y |\varphi(y)|p_Y(y) &= \sum_y \Big|\sum_x xp_{X|Y}(x|y)\Big|p_Y(y) \leq \sum_y\sum_x |x|p_{X|Y}(x|y)p_Y(y)\\ &= \sum_y\sum_x |x|p_{X,Y}(x,y) = \sum_x |x|\sum_y p_{X,Y}(x,y)\\ &= \sum_x |x|p_X(x) < \infty\end{aligned}$$

perché X ha speranza matematica finita. Inoltre,

$$\begin{aligned}\mathbb{E}(\varphi(Y)) &= \sum_y \varphi(y)p_Y(y) = \sum_y\sum_x xp_{X|Y}(x|y)p_Y(y)\\ &= \sum_y\sum_x xp_{X,Y}(x,y) = \sum_x x\sum_y p_{X,Y}(x,y)\\ &= \sum_x xp_X(x) = \mathbb{E}(X).\end{aligned}$$

b) Sappiamo che $\varphi(Y)$ ha momento secondo finito se e solo se $\sum_y |\varphi(y)|^2p_Y(y) < \infty$. Si ha:

$$\sum_y |\varphi(y)|^2p_Y(y) = \sum_y \Big|\sum_x xp_{X|Y}(x|y)\Big|^2 p_Y(y).$$

Ora, sappiamo che, per una qualunque v.a. Z, $|\mathbb{E}(Z)|^2 \leq \mathbb{E}(|Z|^2)$, cioè

$$\Big|\sum_z zp_Z(z)\Big|^2 \leq \sum_z |z|^2p_Z(z).$$

Applicando questa disuguaglianza con $p_Z(\cdot)$ sostituita da $p_{X|Y}(\cdot \mid y)$, otteniamo

$$\Big|\sum_x xp_{X|Y}(x \mid y)\Big|^2 \leq \sum_x |x|^2p_{X|Y}(x \mid y).$$

Sostituendo,

$$\begin{aligned}\sum_y |\varphi(y)|^2p_Y(y) &\leq \sum_y\sum_x |x|^2p_{X|Y}(x|y)p_Y(y)\\ &= \sum_y\sum_x |x|^2p_{X,Y}(x,y) = \sum_x |x|^2\sum_y p_{X,Y}(x,y)\\ &= \sum_x |x|^2p_X(x) < \infty\end{aligned}$$

perché X ha momento secondo finito.

c) Poiché X, $\varphi(Y)$ e $g(Y)$ hanno tutte momento secondo, sappiamo che $Xg(Y)$ e $\varphi(Y)g(Y)$ hanno speranza matematica finita. Procediamo quindi al calcolo diretto delle medie:

$$\begin{aligned}\mathbb{E}(g(Y)\varphi(Y)) &= \sum_y g(y)\varphi(y)p_Y(y) = \sum_y g(y)\sum_x xp_{X|Y}(x|y)p_Y(y)\\ &= \sum_y g(y)\sum_x xp_{X,Y}(x,y) = \sum_{x,y} xg(y)p_{X,Y}(x,y)\\ &= \mathbb{E}(Xg(Y)).\end{aligned}$$

d) Poiché X, $\varphi(Y)$ e $\psi(Y)$ hanno tutte momento secondo, sappiamo che $X-\psi(Y)$ e $\varphi(Y)-\psi(Y)$ hanno momento secondo finito. E si ha

$$\begin{aligned}\mathbb{E}(\psi(Y)\varphi(Y)) &= \sum_y \psi(y)\varphi(y)p_Y(y) = \sum_y \psi(y)\sum_x xp_{X|Y}(x|y)p_Y(y)\\ &= \sum_y \psi(y)\sum_x xp_{X,Y}(x,y) = \sum_{x,y} x\psi(y)p_{X,Y}(x,y)\\ &= \mathbb{E}(X\psi(Y)).\end{aligned}$$

c) Si ha

$$\begin{aligned}\mathbb{E}(|\psi(Y)-X|^2) &= \mathbb{E}(|\psi(Y)-\varphi(Y)+\varphi(Y)-X|^2)\\ &= \mathbb{E}((\psi(Y)-\varphi(Y))^2+(\varphi(Y)-X)^2+2(\psi(Y)-\varphi(Y))(\varphi(Y)-X))\\ &= \underbrace{\mathbb{E}((\psi(Y)-\varphi(Y))^2)}_{\geq 0\ \forall\psi}+\mathbb{E}((\varphi(Y)-X)^2)+2\mathbb{E}((\psi(Y)-\varphi(Y))(\varphi(Y)-X))\\ &\geq \mathbb{E}(|\varphi(Y)-X)^2|)+2\mathbb{E}((\psi(Y)-\varphi(Y))(\varphi(Y)-X)).\end{aligned}$$

Basta quindi verificare che l'ultimo termine a destra è nullo. Infatti,

$$\mathbb{E}((\psi(Y)-\varphi(Y))(\varphi(Y)-X)) = \mathbb{E}((\psi(Y)-\varphi(Y))\varphi(Y)) - \mathbb{E}((\psi(Y)-\varphi(Y))X) = 0$$

come segue dall'utilizzo del punto **b)** con ψ sostituita da $\psi-\varphi$.

Esercizio 5. Ovviamente, indipendenza implica covarianza nulla. Supponiamo ora che $\mathrm{Cov}(X,Y)=0$:

$$0 = \mathrm{Cov}(X,Y) = \mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y) = \mathbb{P}(X=1,Y=1)-\mathbb{P}(X=1)\mathbb{P}(Y=1),$$

cioè $\mathbb{P}(X=1,Y=1)=\mathbb{P}(X=1)\mathbb{P}(Y=1)$. Allora

$$\begin{aligned}\mathbb{P}(X=0,Y=1) &= \mathbb{P}(Y=1)-\mathbb{P}(X=1,Y=1) = \mathbb{P}(Y=1)-\mathbb{P}(X=1)\mathbb{P}(Y=1)\\ &= \mathbb{P}(Y=1)\Big(1-\mathbb{P}(X=1)\Big) = \mathbb{P}(X=0)\mathbb{P}(Y=1),\\ \mathbb{P}(X=0,Y=0) &= \mathbb{P}(X=0)-\mathbb{P}(X=0,Y=1) = \mathbb{P}(X=0)-\mathbb{P}(X=0)\mathbb{P}(Y=1)\\ &= \mathbb{P}(X=0)\Big(1-\mathbb{P}(Y=1)\Big) = \mathbb{P}(X=0)\mathbb{P}(Y=0),\\ \mathbb{P}(X=1,Y=0) &= \mathbb{P}(X=1)-\mathbb{P}(X=1,Y=1) = \mathbb{P}(X=1)-\mathbb{P}(X=1)\mathbb{P}(Y=1)\\ &= \mathbb{P}(X=1)\Big(1-\mathbb{P}(Y=1)\Big) = \mathbb{P}(X=1)\mathbb{P}(Y=0),\end{aligned}$$

da cui segue che X e Y sono indipendenti.
Supponiamo ora che X e Y possano assumere solo due valori $x_1 \neq x_2$ e $y_1 \neq y_2$ rispettivamente e che $\mathrm{Cov}(X,Y) = 0$. Consideriamo le v.a.

$$U = \frac{X - x_1}{x_2 - x_1} \quad \text{e} \quad V = \frac{Y - y_1}{y_2 - y_1}.$$

U e V sono bernoulliane di parametri, rispettivamente, $p = p_X(x_2)$ e $\hat{p} = p_Y(y_2)$. Inoltre,

$$\mathrm{Cov}(U,V) = \mathrm{Cov}\Big(\frac{X - x_1}{x_2 - x_1}, \frac{Y - y_1}{y_2 - y_1}\Big) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} \mathrm{Cov}(X,Y) = 0,$$

quindi, per **a)**, U e V sono indipendenti. Infine, poiché

$$X = x_1 + (x_2 - x_1)U \quad \text{e} \quad Y = y_1 + (y_2 - y_1)V,$$

segue che anche X e Y sono indipendenti.

Esercizio 6. Si ha:

$$\mathbb{E}(X) = \sum_{x \in E_X} x p_X(x) = -2 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = 0$$
$$\mathbb{E}(Y) = \sum_{y \in E_Y} y p_Y(x) = -1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} = 0$$
$$\mathbb{E}(XY) = \sum_{x \in E_X} \sum_{y \in E_Y} xy p_{X,Y}(x,y) = 2 \cdot \frac{1}{6} - 1 \cdot \frac{1}{6} - 4 \cdot \frac{1}{12} + 2 \cdot \frac{1}{12} = 0$$

da cui segue che $\mathrm{Cov}(X,Y) = 0$. Inoltre, ad esempio, $\mathbb{P}(X = -2, Y = -1) = \frac{1}{6} \neq \frac{1}{9} = \mathbb{P}(X = -2)\mathbb{P}(Y = -1)$, dunque X e Y non sono indipendenti. Abbiamo quindi trovato un controesempio alla (evidentemente falsa, in generale) affermazione "covarianza nulla implica indipendenza".

Esercizio 7. Posto $y = g(x)$, risolvendo rispetto a x (il che si può fare perché $\mathrm{Cov}(X,Y) \neq 0$) si ottiene:

$$x = \frac{\mathrm{Var}(X)}{\mathrm{Cov}(X,Y)}\Big(y - \mathbb{E}(Y)\Big) + \mathbb{E}(X),$$

e quindi $\psi(y) = \frac{\mathrm{Var}(X)}{\mathrm{Cov}(X,Y)}\Big(y - \mathbb{E}(Y)\Big) + \mathbb{E}(X)$. Ora, l'equazione della retta di regressione di X rispetto a Y è

$$x = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}\Big(y - \mathbb{E}(Y)\Big) + \mathbb{E}(X).$$

Allora $x = \psi(y)$ è la retta di regressione richiesta se e solo se

$$\frac{\mathrm{Var}(X)}{\mathrm{Cov}(X,Y)} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}$$

che equivale a

$$|\,\mathrm{Cov}(X,Y)\,| = \sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}.$$

Quindi $x = \psi(y) \equiv g^{-1}(y)$ è l'equazione della retta di regressione richiesta se e solo se vale il caso limite della disuguaglianza di Cauchy-Schwartz: X e Y sono tali che $|\,\mathrm{Cov}(X,Y)\,| =$

$\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}$. La soluzione è data dall'Esercizio 3: applicandone i risultati alle v.a. $\bar{X} = X - \mathbb{E}(X)$ e $\bar{Y} = Y - \mathbb{E}(Y)$, otteniamo (vd anche le soluzioni) che dev'essere $X \equiv \mathbb{E}(X)$ oppure $Y \equiv \mathbb{E}(Y)$ oppure, se nessuna delle precedenti è verificata,

$$X - \mathbb{E}(X) = \sqrt{\frac{\mathrm{Var}(X)}{\mathrm{Var}(Y)}}(Y - \mathbb{E}(Y)),$$

cioè X e Y sono legate da una relazione lineare che è proprio quella appena scritta.

Esercizio 8. a) Usando le proprietà tipiche della covarianza, si ha

$$\mathrm{Cov}(X, Z^{\alpha,\beta}) = \mathrm{Cov}(X, \alpha\, X + \beta\, Y) = \alpha \mathrm{Cov}(X, X) + \beta \mathrm{Cov}(X, Y) = \alpha \mathrm{Var}(X)$$

perché X e Y sono indipendenti, quindi non correlate: $\mathrm{Cov}(X, Y) = 0$. Ora, basta ricordare che la varianza di una $\mathrm{Po}(\lambda)$ è $\lambda (> 0)$:

$$\mathrm{Cov}(X, Z^{\alpha,\beta}) = \alpha\, \lambda \text{ quindi: } \begin{cases} \text{se } \alpha > 0 \text{ allora c'è dipendenza } \textbf{positiva} \\ \text{se } \alpha < 0 \text{ allora c'è dipendenza } \textbf{negativa} \\ \text{se } \alpha = 0 \text{ allora c'è dipendenza } \textbf{nulla} \end{cases}$$

Si noti, in particolare, che il tipo di dipendenza **non** dipende da β. Inoltre, se $\alpha = 0$ allora $Z^{\alpha,\beta} = Z^{0,\beta} = \beta\, Y$, quindi X e $Z^{0,\beta}$ non sono solo incorrelate ma anche indipendenti.

b) L'equazione della retta di regressione di $Z^{\alpha,\beta}$ rispetto a X nel piano (x, z) è

$$z = \frac{\mathrm{Cov}(X, Z^{\alpha,\beta})}{\mathrm{Var}(X)}\big(x - \mathbb{E}(X)\big) + \mathbb{E}(Z^{\alpha,\beta}).$$

Poiché $\mathrm{Cov}(X, Z^{\alpha,\beta}) = \alpha\, \lambda$, $\mathbb{E}(X) = \lambda = \mathbb{E}(Y)$ e $\mathbb{E}(Z^{\alpha,\beta}) = \alpha \mathbb{E}(X) + \beta \mathbb{E}(Y) = (\alpha + \beta)\lambda$, la retta di regressione è $z = \alpha\, x + \beta\, \lambda$.

Esercizio 9. a) Siano X e Z rispettivamente il risultato del primo e del secondo lancio. Allora $Y = X \vee Z$ e

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x, X \vee Z = y) = \mathbb{P}(X = x, x \vee Z = y).$$

Se $(x, y) \notin \{1, \ldots, 6\}^2$, ovviamente $p_{X,Y}(x, y) = 0$. Studiamo quindi il comportamento di $p_{X,Y}$ per $(x, y) \in \{1, \ldots, 6\}^2$. Si ha,

$$\begin{aligned} p_{X,Y}(x, y) &= \mathbb{P}(X = x, x \vee Z = y, Z \leq x) + \mathbb{P}(X = x, x \vee Z = y, Z > x) \\ &= \mathbb{P}(X = x, x = y, Z \leq x) + \mathbb{P}(X = x, Z = y, Z > x) \\ &= \begin{cases} 0 & \text{se } x > y \\ \mathbb{P}(X = x, Z \leq x) & \text{se } x = y \\ \mathbb{P}(X = x, Z = y) & \text{se } x < y \end{cases} \end{aligned}$$

Ricordando che X e Z rappresentano il risultato di due tiri (indipendenti) di un dado, possiamo scrivere, per $(x, y) \in \{1, \ldots, 6\}^2$,

$$p_{X,Y}(x, y) = \begin{cases} 0 & \text{se } x > y \\ \frac{x}{6^2} & \text{se } x = y \\ \frac{1}{6^2} & \text{se } x < y \end{cases}$$

Calcoliamo le distribuzioni marginali di X e Y: per $x, y \in \{1, \dots, 6\}$,

$$
\begin{aligned}
p_X(x) &= \sum_{y=1}^{6} p_{X,Y}(x,y) = \frac{x}{6^2} + \sum_{y=x+1}^{6} \frac{1}{6^2} = \frac{x}{6^2} + \frac{6-x}{6^2} = \frac{6}{6^2} = \frac{1}{6}, \\
p_Y(y) &= \sum_{x=1}^{6} p_{X,Y}(x,y) = \sum_{x=1}^{y-1} \frac{1}{6^2} + \frac{y}{6^2} = \frac{y-1}{6^2} + \frac{y}{6^2} = \frac{2y-1}{6^2}.
\end{aligned}
$$

b) Per calcolare la covarianza tra X e Y occorrono $\mathbb{E}[X]$, $\mathbb{E}[Y]$ e $\mathbb{E}[X\,Y]$. Nel seguito, useremo le uguaglianze:

$$
\sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}, \quad \sum_{k=1}^{n} k^3 = \Big(\frac{n(n+1)}{2}\Big)^2.
$$

Si ha

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{x=1}^{6} x\, p_X(x) = \sum_{x=1}^{6} x \frac{1}{6} = \frac{1}{6} \cdot \frac{6 \cdot 7}{2} = \frac{7}{2}; \\
\mathbb{E}[Y] &= \sum_{y=1}^{6} y\, p_Y(y) = \sum_{y=1}^{6} y \frac{2y-1}{6^2} = \frac{1}{6^2}\Big[2\sum_{y=1}^{6} y^2 - \sum_{y=1}^{6} y\Big] \\
&= \frac{1}{6^2}\Big[2 \cdot \frac{6 \cdot 7 \cdot 13}{6} - \frac{6 \cdot 7}{2}\Big] = \frac{161}{36} (\simeq 4.72); \\
\mathbb{E}[X\,Y] &= \sum_{x,y=1}^{6} x\,y\, p_{X,Y}(x,y) = \sum_{x=1}^{6}\sum_{y=1}^{6} x\,y\, p_{X,Y}(x,y) = \sum_{x=1}^{6} x \sum_{y=x} y\, p_{X,Y}(x,y) \\
&= \sum_{x=1}^{6} x\Big[x \cdot \frac{x}{6^2} + \sum_{y=x+1}^{6} y \cdot \frac{1}{6^2}\Big] = \frac{1}{6^2} \sum_{x=1}^{6} x\Big[x^2 + \sum_{y=1}^{6} y - \sum_{y=1}^{x} y\Big] \\
&= \frac{1}{6^2} \sum_{x=1}^{6} x\Big[x^2 + \frac{6 \cdot 7}{2} - \frac{x(x+1)}{2}\Big] = \frac{1}{6^2} \sum_{x=1}^{6} x\, \frac{x^2 - x + 42}{2} \\
&= \frac{1}{2 \cdot 6^2}\Big[\sum_{x=1}^{6} x^3 - \sum_{x=1}^{6} x^2 + 42 \sum_{x=1}^{6} x\Big] \\
&= \frac{1}{2 \cdot 6^2}\Big[\Big(\frac{6 \cdot 7}{2}\Big)^2 - \frac{6 \cdot 7 \cdot 13}{6} + 42 \cdot \frac{6 \cdot 7}{2}\Big] = \frac{154}{9} (\simeq 17.11)
\end{aligned}
$$

Quindi:

$$
\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = \frac{35}{24}
$$

e la dipendenza tra le due variabili è positiva: all'aumentare dell'una ci si aspetta un aumento anche dell'altra.

c) L'equazione della retta di regressione è:

$$
y = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}\Big(x - \mathbb{E}[X]\Big) + \mathbb{E}[Y].
$$

Occorre quindi calcolare $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X]$. Si ha

$$\mathbb{E}[X^2] = \sum_{x=1}^{6} x^2 \, p_X(x) = \sum_{x=1}^{6} x^2 \cdot \frac{1}{6} = \frac{1}{6} \cdot \frac{6 \cdot 7 \cdot 13}{6} = \frac{91}{6}$$

da cui segue che

$$\mathrm{Var}(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$$

Quindi, la retta di regressione ha equazione

$$y = \frac{1}{2}\left(x - \frac{7}{2}\right) + \frac{161}{36}.$$

d) Denotiamo con X_i il risultato dell'$i-$esimo lancio e $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Allora $3.5 = \mathbb{E}[\bar{X}_n]$ e $\mathrm{Var}(\bar{X}_n) = \frac{1}{n}\mathrm{Var}(X_1)$, dove $\mathrm{Var}(X_1) = \frac{35}{12}$. La disuguaglianza di Chebycev assicura che

$$\mathbb{P}(|\bar{X}_n - 3.5| > 0.1) \leq \frac{\mathrm{Var}(\bar{X}_n)}{0.1^2} = 100 \cdot \frac{\mathrm{Var}(X_1)}{n}.$$

Perché si abbia $\mathbb{P}(|\bar{X}_n - 3.5| > 0.1) \leq 0.2$ è sufficiente che

$$100 \cdot \frac{\mathrm{Var}(X_1)}{n} \leq 0.2$$

da cui segue che $n \geq 2000 \cdot \mathrm{Var}(X_1) = 5833, \bar{3}$: basterà scegliere $n = 5834$.

Esercizio 10. a) Calcoliamo dapprima le quantità che servono per il calcolo dei coefficienti della retta di regressione:

$$\begin{aligned}
&\mathbb{E}(U) = \mathbb{E}(X_1X_2) = \mathbb{E}(X_1)\mathbb{E}(X_2) = 0 \\
&\mathbb{E}(V) = \mathbb{E}(X_1X_3) = \mathbb{E}(X_1)\mathbb{E}(X_3) = 0 \\
&\mathrm{Var}(U) = \mathbb{E}(U^2) = \mathbb{E}(X_1^2X_2^2) = \mathbb{E}(X_1^2)\mathbb{E}(X_2^2) = 1 \\
&\mathrm{Cov}(U,V) = \mathbb{E}(UV) = \mathbb{E}(X_1^2X_2X_3) = \mathbb{E}(X_1^2)\mathbb{E}(X_2)\mathbb{E}(X_3) = 0
\end{aligned}$$

Dunque, nel piano (u, v) la retta di regressione ha equazione $v = 0$.

b) U e V hanno covarianza nulla e possono assumere due solo valori: dall'Esercizio 5 possiamo immediatamente dedurre l'indipendenza.

c) Si ha, per $u, v \in \{-1, 1\}$,

$$\begin{aligned}
p_{U,V}(u,v) &= \mathbb{P}(U = u, V = v) = \mathbb{P}(X_1X_2 = u, X_1X_3 = v) \\
&= \mathbb{P}(X_2 = u, X_3 = v, X_1 = 1) + \mathbb{P}(X_2 = -u, X_3 = -v, X_1 = -1) \\
&= \mathbb{P}(X_2 = u)\mathbb{P}(X_3 = v)\mathbb{P}(X_1 = 1) + \mathbb{P}(X_2 = -u)\mathbb{P}(X_3 = -v)\mathbb{P}(X_1 = -1) \\
&= 2 \cdot \frac{1}{8} = \frac{1}{4}
\end{aligned}$$

Osserviamo che, di conseguenza, $p_U(-1) = 1/2 = p_U(1)$ e analogamente $p_V(-1) = 1/2 = p_V(1)$. Poiché $p_{U,V} = p_U p_V$, anche da qui segue l'indipendenza.

Esercizio 11. a) Dobbiamo calcolare $p_{Z,W}(z,w) = \mathbb{P}(Z = z, W = w)$, con $z = \pm 1$ e $w = 1, 2, \ldots$ (per ogni altro valore di z e w, si ha evidentemente $p_{Z,W}(z,w) = 0$). Ora, se $z = 1$,

$$\begin{aligned} p_{Z,W}(1,w) &= \mathbb{P}(Z = 1, W = w) = \mathbb{P}(X \geq 0, |X| = w) = \\ &= \mathbb{P}(X \geq 0, X = w) = \mathbb{P}(X = w) = \frac{1}{2}\, p\,(1-p)^{w-1}. \end{aligned}$$

Analogamente, se $z = -1$,

$$\begin{aligned} p_{Z,W}(-1,w) &= \mathbb{P}(Z = -1, W = w) = \mathbb{P}(X < 0, |X| = w) = \\ &= \mathbb{P}(X < 0, -X = w) = \mathbb{P}(X = -w) = \frac{1}{2}\, p\,(1-p)^{w-1}. \end{aligned}$$

Dunque, riassumendo si ha

$$p_{Z,W}(z,w) = \begin{cases} \frac{1}{2}\, p\,(1-p)^{w-1} & \text{se } z = \pm 1 \text{ e } w = 1, 2, \ldots \\ 0 & \text{altrimenti} \end{cases}$$

Z e W sono indipendenti, perché $p_{Z,W}(z,w) = g_1(z) g_2(w)$ dove

$$g_1(z) = \begin{cases} \frac{1}{2} & \text{se } z = \pm 1 \\ 0 & \text{altrimenti} \end{cases} \qquad \text{e} \qquad g_2(w) = \begin{cases} (1-p)^{w-1} p & \text{se } w = 1, 2, \ldots \\ 0 & \text{altrimenti} \end{cases}$$

Essendo g_1 e g_2 due densità (in particolare, g_2 è la densità del primo successo in prove bernoulliane), si ha immediatamente che $p_Z(z) = g_1(z)$ e $p_W(w) = g_2(w)$.

b) Poiché, come abbiamo detto, Z e W sono indipendenti, si ha $p_{W\,|\,Z}(w\,|\,z) = p_W(w)$. Dunque, $\mathbb{E}(W\,|\,Z = z) = \sum_w w p_{W\,|\,Z}(w\,|\,z) = \sum_w w p_W(w) = \mathbb{E}(W) = 1/p$.

c) Occorre $\mathrm{Cov}(X, Z)$. Per evidenti ragioni di simmetria, $\mathbb{E}(X) = \mathbb{E}(Z) = 0$; inoltre,

$$\mathbb{E}(XZ) = \mathbb{E}(X(\mathbb{1}_{\{X \geq 0\}} - \mathbb{1}_{\{X < 0\}})) = \mathbb{E}(|X|) = \sum_x |x| p_X(x).$$

Ora, poiché $x \mapsto |x| p_X(x)$ è una funzione pari[2] si ha

$$\mathbb{E}(XZ) = \sum_x |x| p_X(x) = 2 \sum_{x=1}^{+\infty} x p_X(x) = \frac{1}{p}$$

(perché $2\sum_{x=1}^{+\infty} x p_x(x) = \mathbb{E}(T)$, con T istante di primo successo in prove bernoulliane con probabilità di successo $= p$). Dunque, $\mathrm{Cov}(X, Z) = \mathbb{E}(XZ) = 1/p > 0$, da cui segue che tra X e Z c'è dipendenza positiva. Per scrivere l'equzione dlla retta di regressione di Z rispetto a X, occorre calcolare $\mathrm{Var}(X)$: ricordando che $\mathbb{E}(X) = 0$, si ha

$$\mathrm{Var}(X) = \mathbb{E}(X^2) = \sum_x x^2 p_X(x) = 2 \sum_{x=1}^{+\infty} x^2 p_X(x),$$

dove abbiamo usato il fatto che $x \mapsto x^2 p_X(x)$ è pari. Ora, detta T la v.a. istante di primo successo in prove bernoulliane con prob. successo $= p$, si ha

[2] Ricordiamo che $g(x)$ è pari se $g(x) = g(-x)$.

$$2\sum_{x=1}^{+\infty} x^2 p_X(x) = \mathbb{E}(T^2) = \mathrm{Var}(T) + \mathbb{E}^2(T) = \frac{1-p}{p^2} + \frac{1}{p^2} = \frac{2-p}{p^2}.$$

Dunque, nel piano (x, z) la retta di regressione ha equazione

$$z = \frac{p}{2-p}\, x.$$

d) $\mathbb{E}(X \mid Z = -1) = \sum_x x p_{X \mid Z}(x \mid -1)$, essendo $p_{X \mid Z}(x \mid -1) = p_{X,Z}(x, -1)/p_Z(-1) = 2\, p_{X,Z}(x, -1)$. Ora,

$$\begin{aligned} p_{X,Z}(x, -1) &= |P(X = x, Z = -1) = \\ = \mathbb{P}(X = x, X < 0) &= \begin{cases} \mathbb{P}(X = x) = p_X(x) & \text{se } x = -1, -2, -3, \dots \\ 0 & \text{altrimenti} \end{cases} \end{aligned}$$

Quindi,

$$\mathbb{E}(X \mid Z = -1) = \sum_{x=-\infty}^{-1} x p(1-p)^{-x-1} = -\sum_{k=1}^{+\infty} k p(1-p)^{k-1} = -\mathbb{E}(T) = -\frac{1}{p}$$

dove T denota la v.a. istante di primo successo come sopra.

Esercizio 12. a) Poiché $X_1, \dots, X_6$ sono indipendenti, segue che anche le tre v.a. Y, Z, W sono indipendenti. Quindi,

$$\begin{aligned} \mathrm{Cov}(T, U) &= \mathrm{Cov}(Y + Z, Z + W) = \mathrm{Cov}(Y, Z) + \mathrm{Cov}(Y, W) + \mathrm{Cov}(Z, Z) + \mathrm{Cov}(Z, W) = \\ &= 0 + 0 + \mathrm{Var}(Z) + 0 = \mathrm{Var}(X_3 + X_4) = \mathrm{Var}(X_3) + \mathrm{Var}(X_4) = 2p(1-p), \end{aligned}$$

b) $\mathbb{E}(T) = \mathbb{E}(X_1 + \cdots + X_4) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_4) = 4p$. Inoltre, poiché le X_i sono indipendenti, $\mathrm{Var}(T) = \mathrm{Var}(X_1 + \cdots + X_4) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_4) = 4p(1-p)$.

c) Si ha: $\mathbb{E}(2T + 1) = 2\mathbb{E}(T) + 1 = 8p + 1$ e $\mathrm{Var}(2T + 1) = \mathrm{Var}(2T) = 4\mathrm{Var}(T) = 16p(1-p)$.

d) Si ha: $\mathrm{Cov}(2T + 1, 3U) = 6\mathrm{Cov}(T, U) + 3\mathrm{Cov}(1, U) = 4p(1-p)$ perché si ha sempre $\mathrm{Cov}(c, X) = 0$ quando c è costante (e X è qui una v.a. qualsiasi).