Tutorato V
Probabilità e Statistica
a.a. 2022/2023

Argomenti: momenti; varianza e covarianza; retta di regressione; media condizionale come migliore approssimazione; disuguaglianza di Chebycev e legge dei grandi numeri.

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: lavorare 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. 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 5. 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 6. 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 7. 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 8. 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 9. 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 $Z = 1\!\!1_{\{X \geq 0\}} - 1\!\!1_{\{X<0\}}$ 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 10. 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 $3U$.

Esercizio 11. 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 11 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 12. 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 13. 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 14. Una moneta dà testa con probabilità p incognita. Per avere informazioni su p, la moneta viene tirata n volte e si osserva la percentuale di teste uscite su n lanci. Volendo approssimare p con tale percentuale,

a) quanto deve essere grande n affinché con probabilità maggiore di 0.99 l'errore commesso sia al più 0.1?

b) E quanto deve essere grande n affinché con probabilità maggiore di 0.9 l'errore commesso sia al più 0.01?

Qualcuno sostiene che occorrano più lanci per vedere soddisatta la prima condizione, perché la probabilità richiesta è maggiore. Cosa ne pensate?

[Sugg.: si usi la disuguaglianza di Chebycev.]

Esercizio 15. Un dado viene lanciato n volte. Sia S_n in numero di volte, su n lanci, in cui si osserva che è uscito 6. Calcolare $\lim_{n\to\infty} \mathbb{P}(\frac{1}{n} S_n > 2)$.

[Sugg.: si ricorda la legge dei grandi numeri...]

Esercizio 16. Siano X e Y v.a. indipendenti, $X \sim \mathrm{Geomod}(\frac{1}{2})$ e $Y \sim \mathrm{Po}(\frac{1}{2})$. Sia $\{U_n\}_n$ una successione di v.a. i.i.d. tutte con la stessa legge di $U = X - 2Y$. Stimare n affinché la v.a. $S_n = U_1 + \ldots + U_n$ non superi il livello $3n$ con probabilità almeno 0.9.