**Tutorato di Probabilità 1, VII**
**a.a. 2003/2004**

**Esercizio 1**

Siano $X, Y, Z$ tre v.a. indipendenti, geometriche di parametro $p \in (0,1)$.

**a)** Calcolare $\mathrm{P}(X = Y)$ e $\mathrm{P}(X \geq 2Y)$.

**b)** Calcolare la distribuzione e la media condizionata di $X$ dato $X+Y$ e di $Z$ dato $X+Y$.

**Esercizio 2**

In uno schema di Bernoulli, con $p =$probabilità di ottenere 1, calcolare[1] il numero medio di esperimenti che occorre effettuare perché per la prima volta si osservi la sequenza 01.

**Esercizio 3**

**a)** Fissato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathrm{P})$, sia

$$L^2_d = \{X : \Omega \to \mathbb{R} \text{ t.c. } X \text{ è discreta ed esiste } \mathbb{E}(X^2)\}.$$

Dimostrare che $L^2_d$ è uno spazio vettoriale.

**b)** Dimostrare la **disuguaglianza di Cauchy-Schwartz**: per $X, Y \in L^2_d$,

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

e dedurre che se $X, Y \in L^2_d$ allora esiste $\mathrm{Cov}(X,Y)$ e

$$|\mathrm{Cov}(X,Y)| \leq \sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}$$

[Sugg.: si consideri la funzione $g(\lambda) = \mathbb{E}[(\lambda X + Y)^2]$: $g(\lambda)$ è un trinomio di secondo grado e $g(\lambda) \geq 0$, per ogni $\lambda \in \mathbb{R}$. Ciò significa che il discriminante dev'essere necessariamente...]

**c)** Sia

$$\langle \cdot, \cdot \rangle : (X,Y) \in L^2_d \times L^2_d \mapsto \langle X, Y \rangle = \mathbb{E}(XY) \in \mathbb{R}$$

Dimostrare che $\langle \cdot, \cdot \rangle$ definisce un **prodotto scalare**[2] su $L^2_d$ e quindi

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

è una **norma** su $L^2_d$. Dedurre che $L^2_d$ è uno spazio euclideo.

**Esercizio 4**

Dimostrare che per ogni scelta di $X, Y, X_1, X_2 \in L^2_d$,

**a1)** $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$;

**a2)** $\mathrm{Cov}(X_1 + X_2, Y) = \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y)$;

**a3)** $\mathrm{Cov}(aX, Y) = a\mathrm{Cov}(X,Y)$, per ogni $a \in \mathbb{R}$;

**a4)** $\mathrm{Cov}(X,X) = \mathrm{Var}(X) \geq 0$ e $\mathrm{Var}(X) = 0$ se e solo se $X = \alpha$ per qualche $\alpha \in \mathbb{R}$.

**a5)** $\mathrm{Cov}(X, \alpha) = 0$, per ogni $\alpha \in \mathbb{R}$.

Dedurre che la covarianza è un'applicazione **simmetrica** e **bilineare**: per ogni $a, b, c, d \in \mathbb{R}$,

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

Ma $\mathrm{Cov}(X,Y)$ non definisce un prodotto scalare su $L^2_d$: perché?

---

[1] Potrebbe essere utile ricordare che, per $\alpha \in (-1,1)$, $\sum_{n=1}^{\infty} n\alpha^{n-1} = \frac{1}{(1-\alpha)^2}$.

[2] 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 $v = 0$.

## Soluzioni

**1) a1)**

$$\mathrm{P}(X=Y)=\sum_{k=0}^{\infty}\mathrm{P}(X=k,Y=k)=\sum_{k=0}^{\infty}p^2(1-p)^{2k}=p^2\frac{1}{1-(1-p)^2}=\frac{p}{2-p};$$

**a2)**

$$\mathrm{P}(X\geq 2Y)=\sum_{k=0}^{\infty}\mathrm{P}(Y=k,X\geq 2k)=\sum_{k=0}^{\infty}\mathrm{P}(Y=k)\mathrm{P}(X\geq 2k)$$

dove

$$\mathrm{P}(X\geq 2k)=\sum_{j=2k}^{\infty}p(1-p)^j=p(1-p)^{2k}\sum_{j=2k}^{\infty}(1-p)^{j-2k}=p(1-p)^{2k}\sum_{\ell=0}^{\infty}(1-p)^{\ell}=(1-p)^{2k},$$

quindi

$$\mathrm{P}(X\geq 2Y)=\sum_{k=0}^{\infty}\mathrm{P}(Y=k)\mathrm{P}(X\geq 2k)=p\sum_{k=0}^{\infty}(1-p)^{3k}=p\frac{1}{1-(1-p)^3}=\frac{1}{1+(1-p)+(1-p)^2};$$

**b1)** Cerchiamo la distribuzione congiunta di $X$ e $X+Y$:

$$p_{X,X+Y}(i,j)=\mathrm{P}(X=i,Y=j-i)=\begin{cases}p^2(1-p)^j & \text{se } j\geq i\geq 0\\ 0 & \text{altrimenti.}\end{cases}$$

La marginale di $X+Y$ vale

$$p_{X+Y}(j)=\sum_i p_{X,X+Y}(i,j)=\sum_{i=0}^{j}p^2(1-p)^j=(j+1)p^2(1-p)^j$$

con $j\geq 0$. Fissato $j\geq 0$, la distribuzione condizionata è

$$p_{X\,|X+Y}(i\,|j)=\frac{p_{X,X+Y}(i,j)}{p_{X+Y}(j)}=\begin{cases}\frac{1}{j+1} & \text{se } 0\leq i\leq j\\ 0 & \text{altrimenti}\end{cases}$$

cioè $X\,|X+Y=j$ si distribuisce uniformemente nell'insieme $\{0,1,\ldots,j\}$. Allora, ancora fissato $j\geq 0$,

$$\mathbb{E}(X\,|X+Y=j)=\sum_i i p_{X\,|X+Y}(i\,|j)=\frac{1}{j+1}\sum_{i=0}^{j}i=\frac{1}{j+1}\frac{j(j+1)}{2}=\frac{j}{2}.$$

La media condizionata è dunque la v.a.

$$\mathbb{E}(X\,|X+Y)=\frac{X+Y}{2}.$$

**b2)** Poiché $Z$ è indipendente sia da $X$ che da $Y$, $Z$ è indipendente da $X+Y$, quindi

$$p_{Z\,|X+Y}(i\,|j)=p_Z(i)$$

per ogni $i,j\geq 0$, quindi $\mathbb{E}(Z\,|X+Y=j)=\mathbb{E}(Z)=\frac{1-p}{p}$ e la media condizionata non è aleatoria ma deterministica:

$$\mathbb{E}(Z\,|X+Y)=\mathbb{E}(Z)=\frac{1-p}{p}.$$

**2)** Indicando con $X_1, X_2, \dots$ lo schema di Bernoulli, possiamo scrivere

$$T = \inf\{n \geq 2 \,:\, X_{n-1} = 0 \text{ e } X_n = 1\}.$$

Si richiede è $\mathbb{E}(T)$, per il cui calcolo occorre avere la distribuzione di $T$.

Il punto è che $\{T = k\} = \cup_{j=0}^{n-2} A_j$, dove

$$A_j = \begin{cases} \{X_1 = \cdots = X_{n-2} = 0, X_{n-1} = 0, X_n = 1\} & \text{se } j = 0, \\ \{X_1 = \cdots = X_j = 1, X_{j+1} = \cdots = X_{n-2} = 0, X_{n-1} = 0, X_n = 1\} & \text{se } j \geq 1. \end{cases}$$

Poiché gli $A_j$ sono disgiunti e $\mathrm{P}(A_j) = p^j \cdot (1-p)^{n-2-j} \cdot (1-p) \cdot p = p^{j+1}(1-p)^{n-1-j}$, si ha

$$\begin{aligned} p_T(n) &= \sum_{j=0}^{n-2} \mathrm{P}(A_j) = \sum_{j=0}^{n-2} p^{j+1}(1-p)^{n-1-j} = p(1-p)^{n-1} \sum_{j=0}^{n-2} \Big(\frac{p}{1-p}\Big)^j \\ &= p(1-p)^{n-1} \frac{1 - \Big(\frac{p}{1-p}\Big)^{n-1}}{1 - \frac{p}{1-p}} = p(1-p) \frac{(1-p)^{n-1} - p^{n-1}}{1-2p} \end{aligned}$$

(si noti che in effetti $p_T(n) > 0$ per ogni $k \geq n$). Possiamo ora calcolare la media di $T$:

$$\mathbb{E}(T) = \sum_{n=2}^{\infty} n\, p_T(n) = \sum_{n=2}^{\infty} n\, p(1-p) \frac{(1-p)^{n-1} - p^{n-1}}{1-2p} = \frac{p(1-p)}{1-2p} \Big( \sum_{n=2}^{\infty} n\,(1-p)^{n-1} - \sum_{n=2}^{\infty} n\, p^{n-1} \Big).$$

Sappiamo che, per $\alpha \in (-1,1)$, $\sum_{n=1}^{\infty} n\,\alpha^{n-1} = \frac{1}{(1-\alpha)^2}$, quindi $\sum_{n=2}^{\infty} n\,\alpha^{n-1} = \frac{1}{(1-\alpha)^2} - 1 = \frac{\alpha(2-\alpha)}{(1-\alpha)^2}$. Allora,

$$\mathbb{E}(T) = \frac{p(1-p)}{1-2p} \Big( \frac{(1-p)(2-(1-p))}{(1-(1-p))^2} - \frac{p(2-p)}{(1-p)^2} \Big) = \frac{1}{p(1-p)}.$$

**3) a)** Sappiamo[3] che se $X$ e $Y$ hanno momento $k-$esimo allora anche $X+Y$ ha momento di ordine $k$, così come $\alpha X$ e $\beta Y$ per ogni $\alpha, \beta \in \mathbb{R}$. Quindi $L^2_d$ è uno spazio vettoriale.

**b)** Per $x, y \in \mathbb{R}$, è ben noto che $|xy| \leq \frac{1}{2}(|x|^2 + |y|^2)$, da cui segue che se $X, Y \in L^2_d$ allora $XY$ ha media. Inoltre, posto $g(\lambda) = \mathbb{E}[(\lambda X + Y)^2]$, per $\lambda \in \mathbb{R}$ e $X, Y \in L^2_d$, si ha

$$0 \leq g(\lambda) = \lambda^2 \mathbb{E}(X^2) + 2\lambda \mathbb{E}(XY) + \mathbb{E}(Y^2)$$

per ogni $\lambda \in \mathbb{R}$. Ma allora il discriminante deve essere non positivo:

$$\mathbb{E}(XY)^2 - \mathbb{E}(X^2)\,\mathbb{E}(Y^2) \leq 0$$

da cui segue che $|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)\,\mathbb{E}(Y^2)}$. In particolare, la disuguaglianza vale con $X - \mathbb{E}(X)$ e $Y - \mathbb{E}(Y)$:

$$|\mathrm{Cov}(X,Y)| = |\mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]| \leq \sqrt{\mathbb{E}[(X - \mathbb{E}(X))^2]\,\mathbb{E}[(Y - \mathbb{E}(Y))^2]} = \sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}.$$

**c)** 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 $X \in L^2_d$, $\langle \lambda X, Y \rangle = \lambda \langle X, Y \rangle$: $\langle \lambda X, Y \rangle = \mathbb{E}(\lambda X\, Y) = \lambda \mathbb{E}(XY) = \lambda \langle X, Y \rangle$;

[3] La verifica segue dalla validità della disuguaglianza: $|x+y|^k \leq 2^{k-1}(|x|^k + |y|^k)$, per ogni intero $k \geq 1$ e $x, y \in \mathbb{R}$.

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 $L^2_d$.

**4)** La verifica delle **a1)–a4)** è immediata conseguenza delle proprietà **1.–4.** valide per il prodotto scalare definito nell'esercizio 3: 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 **a5)**, 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 $X = \alpha$ e non è detto che $\alpha = 0$.