8 Esercitazione 8: / /

8.1 Vettori aleatori discreti: densità e funzioni di ripartizione.

Esercizio 1 *Un'urna contiene 3 biglie rosse, due biglie bianche ed una verde. Si estraggono due biglie senza reinserimento. Siano R il numero di biglie rosse estratte e B il numero di biglie bianche estratte.*

1. *Qual è la densità congiunta del vettore (R, B)?*
2. *Calcolare la covarianza di R e B. B e R sono variabili aleatorie non correlate?*

Esercizio 2 *I gestori di una famosa localitá sciistica vogliono studiare le percentuali di guasti di due impianti di risalita (che indicheremo con $p_1 \in (0,1)$ e $p_2 \in (0,1)$) durante una giornata di alta stagione. L'addetto al servizio prende nota del numero T_1 di risalite che avvengono senza problemi fino al primo guasto (compreso), e analogamente per il secondo impianto, segnando il numero T_2 di risalite senza problemi fino al primo guasto (compreso).*

1. *Scrivete la legge di (T_1, T_2).*
2. *Calcolare la probabilitá che sia $T_1 = T_2$ e che $T_1 \geq T_2$.*
3. *Verificare che la v.a. $\min(T_1, T_2)$ segue una legge geometrica e identificarne il parametro.*
4. *Qual è la legge e l'istante atteso del primo guasto comune?*

Esercizio 3 *Siano X ed Y due v.a. con la seguente tabella di densità:*

XY	-1	5	10	
0		*0.12*		*0.4*
5				
	0.3			

1. *Completare la tabella in modo che X ed Y siano indipendenti.*
2. *Calcolare $\mathbb{P}(|XY| \geq 5)$.*

Esercizio 4 *Tre giocatori A, B, C lanciano una moneta equilibrata: il primo due volte, il secondo tre volte e il terzo cinque volte. Con la solita convenzione, consistente nell'identificare testa con "successo" si provi che le due probabilità seguenti sono uguali:*

1. *la probabilità che il numero di successi di A risulti uguale al numero di successi di B;*

2. *la probabilità che il numero di successi di C sia uguale al numero di lanci di A.*

Le due probabilità restano uguali anche se A lancia la moneta m volte ($n \leq m$), B n volte e C $m+n$ volte ?

Esercizio 5 *[Tema d'esame MPSPS 17/07/96] Un'urna contiene 15 palline, di cui 4 Rosse, 5 Bianche e 6 Nere. Quattro palline sono estratte senza reimmissione. Siano X il numero di palline rosse, Y il numero di palline bianche e Z il numero di palline nere estratte.*
1. *Qual è la distribuzione congiunta di (X, Y, Z)?*
2. *Calcolare le distribuzioni marginali di X e Y.*
3. *X ed Y sono indipendenti?*
4. *Come cambiano le risposte ai punti precendenti se le estrazioni sono fatte con reimmissione?*

Esercizio 6 *Siano X ed Y due variabili aleatorie discrete e X sia una v.a. costante (i.e. esiste $c \in \mathbb{R}$ tale che $\mathbb{P}(X = c) = 1$), mostrare che X ed Y sono indipendenti.*

Esercizio 7 *Sia p un primo maggiore di 1 e sia $\Omega = \{1, 2, \dots, p\}$ uno spazio campionario dotato della probabilità uniforme. Siano X ed Y due v.a. discrete definite su questo spazio. Mostrare che condizione necessaria e sufficiente affinché siano indipendenti è che almeno una delle due v.a. sia costante.*

SOLUZIONI

Soluzione esercizio 1.

1. Le combinazioni totali distinte sono $\binom{6}{2}$

$$\begin{aligned} p_{RB}(0,1) &= \frac{2}{\binom{6}{2}} = \frac{2}{15} \\ p_{RB}(0,2) &= \frac{1}{\binom{6}{2}} = \frac{1}{15} \\ p_{RB}(1,0) &= \frac{3}{\binom{6}{2}} = \frac{1}{5} \\ p_{RB}(1,1) &= \frac{3 \cdot 2}{\binom{6}{2}} = \frac{2}{5} \\ p_{RB}(2,0) &= \frac{\binom{3}{2}}{\binom{6}{2}} = \frac{1}{5} \end{aligned}$$

$p_{R\,B}(r,b) = 0$ altrove.

2. Si ricordi che se f è $\mathbb{P}_X$-integrabile,

$$\mathbb{E}(f \circ X) = \int f \mathrm{d}\mathbb{P}_X = \int f(\bar{x})\rho_X(\bar{x})\mathrm{d}\bar{x}$$

dove l'ultima uguaglianza vale solo se la legge $\mathbb{P}_X$ di X ammette densità. Si ha che $RB = 1$ se e solo se $R = B = 1$, negli altri casi $RB = 0$, quindi $\mathbb{E}(RB) = 2/5$. Per le marginali otteniamo:

$$\begin{aligned} p_R(1) &= \frac{1}{5} + \frac{2}{5} = \frac{3}{5} \\ p_R(2) &= \frac{1}{5} \\ p_B(1) &= \frac{2}{15} + \frac{2}{5} = \frac{8}{15} \\ p_B(2) &= \frac{1}{15}, \end{aligned}$$

quindi $\mathbb{E}(R) = 3/5 + 2/5 = 1$ mentre $\mathbb{E}(B) = 8/15 + 2/15 = 2/3$. Ne segue che $\mathrm{Cov}(R,B) = 2/5 - 2/3 = -4/15 \neq 0$. Quindi le variabili sono correlate.

Osservazione: quando si chiede la legge di una variabile (al limite a valori vettoriali) X si chiede di descrivere la probabilità $\mathbb{P}_X(\cdot) := \mathbb{P}(X^{-1}(\cdot))$. Per quanto riguarda il supporto di questa misura, se X è a valori in uno spazio misurabile con la σ-algebra di Borel (derivante da una topologia che soddisfa il secondo assioma di numerabilità) allora esiste un supporto minimo descritto da $\mathrm{essRg}(X) := \{z : \forall U \text{ aperto} : z \in U : \mathbb{P}(X^{-1}(U)) > 0\}$. Questo oggetto è spesso abbastanza facile da trattare: ad esempio se f è una funzione misurabile, continua su $\mathrm{essRg}(X)$, allora $\mathrm{essRg}(f \circ X) = \overline{f(\mathrm{essRg}(X))}$. Se invece X è tale che $\mathbb{P}_X(\text{essRg}(X)^c) > 0$ allora $f : \mathbb{1}_{\text{essRg}(X)}$ e si ha $\text{essRg}(f \circ X) = \{0, 1\}$ mentre $\overline{f(\text{essRg}(X))} = \{0\}$.

Soluzione esercizio 2.

1. In prima approssimazione possiamo supporre che ogni risalita sia indipendente dalle altre e che ciascun impianto abbia la stessa probabilità di guastarsi ad ogni risalita. Poniamo X_i la variabile aleatoria che vale 1 se all'i-esima risalita del primo impianto è avvenuto un guasto e che vale 0 altrimenti, analogamente definiamo Y_i la variabile aleatoria che vale 1 se all'i-esima risalita del secondo impianto è avvenuto un guasto e che vale 0 altrimenti. Abbiamo che $T_1 \sim \text{Geom}(p_1)$ e $T_2 \sim \text{Geom}(p_2)$ essendo gli istanti di primo successo di due processi di Bernoulli, i processi $(X_n)_{n\in\mathbb{N}}$ e $(Y_n)_{n\in\mathbb{N}}$ rispettivamente.
   T_1 é indipendente da T_2 quindi abbiamo che il vettore aleatorio (T_1, T_2) ha la seguente densitá:
   $$p_{(T_1,T_2)}(k, h) = p_1(1 - p_1)^{k-1}p_2(1 - p_2)^{h-1} \quad \forall k = 1, \dots, \forall h = 1, \dots.$$

2. $$\begin{aligned}\mathbb{P}(T_1 = T_2) &= \mathbb{P}(\bigcup_{k=1}^{\infty}\{T_1 = k, T_2 = k\}) = \sum_{k=1}^{\infty}\mathbb{P}(\{T_1 = k, T_2 = k\})\\ &= \sum_{k=1}^{\infty}\mathbb{P}(T_1 = k)\mathbb{P}(T_2 = k) = \sum_{k=1}^{+\infty}p_1(1 - p_1)^{k-1}p_2(1 - p_2)^{k-1}\\ &= \frac{p_1p_2}{-p_1p_2 + p_1 + p_2} = \frac{1}{1/p_1 + 1/p_2 - 1}.\end{aligned}$$
   Notiamo che per $k \in \mathbb{N}$,
   $$\mathbb{P}(T_1 \geq k) = \mathbb{P}(X_1 = 0, \dots, X_{k-1} = 0) = (1 - p_1)^{k-1},$$
   perciò:
   $$\begin{aligned}\mathbb{P}(T_1 \geq T_2) &= \mathbb{P}(\bigcup_{k=1}^{\infty}\{T_1 \geq k, T_2 = k\}) = \sum_{k=1}^{\infty}\mathbb{P}(\{T_1 \geq k, T_2 = k\})\\ &= \sum_{k=1}^{\infty}\mathbb{P}(T_1 \geq k)\mathbb{P}(T_2 = k) = \sum_{k=1}^{+\infty}(1 - p_1)^{k-1}p_2(1 - p_2)^{k-1}\\ &= \frac{p_2}{-p_1p_2 + p_1 + p_2}\end{aligned}$$

3. Sia Z una v.a. geometrica di parametro p, allora è sufficiente calcolare la sua funzione di ripartizione nei naturali positivi per conoscere la sua

funzione di ripartizione su tutto $\mathbb{R}$. Si ha che: $\mathbb{P}(Z \leq k) = 1 - \mathbb{P}(Z > k) = 1 - (1-p)^k$, per $k = 1, \cdots$.

Per quanto riguarda $\min(T_1, T_2)$, è chiaro che assume valori in $\mathbb{N}^*$, inoltre:

$$\begin{aligned}\mathbb{P}(\min(T_1,T_2) \leq k) &= 1 - \mathbb{P}(\min(T_1,T_2) > k) = 1 - \mathbb{P}(T_1 > k)\mathbb{P}(T_2 > k) \\ &= 1 - (1-p_1)^k(1-p_2)^k = 1 - (1 + p_1p_2 - p_1 - p_2)^k.\end{aligned}$$

Quindi $\min(T_1, T_2)$ ammette densitá geometrica di parametro $p = -p_1p_2 + p_1 + p_2$.

4. Poniamo $Z_i = X_iY_i$ per ogni $i \in \mathbb{N}^*$, è facile vedere che $Z = (Z_n)_{n\in\mathbb{N}^*}$ è un processo di Bernoulli di parametro p_1p_2 e che la v.a. $T =$ "istante del primo successo di Z'' è proprio la variabile di cui ci interessa calcolare la media. Quindi $\mathbb{E}(T) = \frac{1}{p_1p_2}$, dal momento che T è geometrica di parametro p_1p_2.

Soluzione esercizio 3.

1. Se troviamo le densità marginali abbiamo finito, dal momento che dobbiamo costruire la densità congiunta in modo che X ed Y siano indipendenti. Visto che X assume solo due valori $\mathbb{P}(X = 5) = 1 - \mathbb{P}(X = 0) = 0.6$. Sappiamo inoltre che $\mathbb{P}(X = 0, Y = 5) = 0.4 \cdot \mathbb{P}(Y = 5) = 0.12$, quindi $\mathbb{P}(Y = 5) = 0.3$. Allora si ha:

XY	-1	5	10	
0		0.12		0.4
5				0.6
	0.3	0.3	1-(0.3+0.3)=0.4	

Quindi otteniamo:

XY	-1	5	10	
0	0.12	0.12	0.16	0.4
5	0.18	0.18	0.24	0.6
	0.3	0.3	0.4	

2. $\mathbb{P}(|XY| \geq 5) = \mathbb{P}(X = 5) = 0.6$.

Soluzione esercizio 4.

Vediamo direttamente il caso generale.

Siano S_A, S_B ed S_C in numero di successi di A, B e C rispettivamente. Quindi: $S_A \sim \text{Binom}(n, 1/2)$, $S_B \sim \text{Binom}(m, 1/2)$ e $S_C \sim \text{Binom}(n+m, 1/2)$. Abbiamo che:

$$\mathbb{P}(S_A = S_B) = \sum_{k=0}^{n} \mathbb{P}(S_A = k, S_B = k) = \sum_{k=0}^{n} \binom{n}{k}\binom{m}{k}\frac{1}{2^{n+m}}$$

D' altra parte $\mathbb{P}(S_C = n) = \binom{n+m}{n}\frac{1}{2^{m+n}}$. Quindi si tratta di dimostrare che $\sum_{k=0}^{n}\binom{n}{k}\binom{m}{k} = \binom{n+m}{n}$ (identità che abbiamo già visto in un esercizio precedente). In ogni caso, notiamo innanzitutto che $\binom{n}{k} = \binom{n}{n-k}$, quindi possiamo riscrivere l'identità che dobbiamo dimostrare così: $\sum_{k=0}^{n}\binom{n}{n-k}\cdot\binom{m}{k} = \binom{n+m}{n}$. Consideriamo la v.a. $Y \sim \text{hyper}(n+m, m, n)$ allora la sua densità è la seguente:

$$p_Y(k) = \frac{\binom{n}{n-k}\cdot\binom{m}{k}}{\binom{n+m}{n}}, \qquad \text{per } k = 0, \dots, n$$

Quindi

$$\sum_{k=0}^{n}\frac{\binom{n}{n-k}\cdot\binom{m}{k}}{\binom{n+m}{n}} = 1$$

da cui l'identità cercata.

Soluzione esercizio 5.

[**Caso delle estrazioni senza reimmissione**] Partiamo dal considerare le distribuzioni marginali di X ed Y. Poichè le estrazioni dall'urna sono effettuate senza reimmissione, X ed Y hanno entrambe distribuzione ipergeometrica:

$$\mathbb{P}(X = x) = \frac{\binom{4}{x}\binom{11}{4-x}}{\binom{15}{4}} \quad x = 0, 1, \dots, 4$$

$$\mathbb{P}(Y = y) = \frac{\binom{5}{y}\binom{10}{4-y}}{\binom{15}{4}} \quad y = 0, 1, \dots, 4$$

La funzione di densità del vettore (X, Y, Z) è data da

$$\begin{aligned}
&\mathbb{P}(X = x, Y = y, Z = z) \\
&\quad = \mathbb{P}(\text{“su 4 estrazioni si ottengono } x \text{ palline rosse, } y \text{ palline bianche e } z \text{ palline nere”}) \\
&\quad = \begin{cases} \frac{\binom{4}{x}\binom{5}{y}\binom{6}{z}}{\binom{15}{4}} & x + y + z = 4,\ x, y, z \in \mathbb{N} \\ 0 & \textit{altrove} \end{cases}
\end{aligned}$$

La distribuzione marginale di X si ottiene nel seguente modo:

$$\begin{aligned}
\mathbb{P}(X = x) &= \sum_{y,z\in\mathbb{N}: y+z=4-x} \mathbb{P}(X = x, Y = y, Z = z) \\
&= \sum_{y=0}^{4-x} \mathbb{P}(X = x, Y = y, Z = 4 - x - y) \\
&= \sum_{y=0}^{4-x} \frac{\binom{4}{x}\binom{5}{y}\binom{6}{4-x-y}}{\binom{15}{4}} = \frac{\binom{4}{x}}{\binom{15}{4}} \sum_{y=0}^{4-x} \binom{5}{y}\binom{6}{4-x-y} \\
&= \frac{\binom{4}{x}\binom{11}{4-x}}{\binom{15}{4}}
\end{aligned}$$

poichè, per ogni n intero positivo, vale la seguente eguaglianza

$$\sum_{y=0}^{n} \binom{k_1}{y}\binom{k_2}{n-y} = \binom{k_1+k_2}{n}.$$

Valendo $\mathbb{P}(X=4)\cdot\mathbb{P}(Y=4) > 0 = \mathbb{P}(X=4, Y=4)$, X ed Y non sono indipendenti.

[**Caso delle estrazioni con reimmissione**] Per quanto concerne la distribuzione congiunta:

$$\mathbb{P}(X=x, Y=y, Z=z) = \begin{cases} \frac{4!}{x!y!z!}\left(\frac{4}{15}\right)^x\left(\frac{5}{15}\right)^y\left(\frac{6}{15}\right)^z & x+y+z=4, x,y,z\in\mathbb{N} \\ 0 & \textit{altrove.} \end{cases}$$

cioè (X,Y,Z) ha legge multinomiale di parametri $(4, 4/15, 5/15, 6/15)$. Segue che marginalmente sia X sia Y sono binomiali: X ha distribuzione binomiale di parametri $n=4, p=4/15 \simeq 0.26667$ ed Y ha distribuzione binomiale di parametri $n=4, p=1/3$. Verifichiamo quanto affermato per X:

$$\begin{aligned}
\mathbb{P}(X=x) &= \sum_{y=0}^{4-x} \frac{4!}{x!y!(4-x-y)!}\left(\frac{4}{15}\right)^x\left(\frac{5}{15}\right)^y\left(\frac{6}{15}\right)^{4-x-y} \\
&= \frac{4!}{x!(4-x)!}\left(\frac{4}{15}\right)^x \sum_{y=0}^{4-x} \frac{(4-x)!}{y!(4-x-y)!}\left(\frac{5}{15}\right)^y\left(\frac{6}{15}\right)^{4-x-y} \\
&= \binom{4}{x}\left(\frac{4}{15}\right)^x \sum_{y=0}^{4-x} \binom{4-x}{y}\left(\frac{5}{15}\right)^y\left(\frac{6}{15}\right)^{4-x-y} \\
&= \binom{4}{x}\left(\frac{4}{15}\right)^x \left(1-\frac{4}{15}\right)^{4-x}
\end{aligned}$$

Poichè $\mathbb{P}(X=4)\cdot\mathbb{P}(Y=4) > 0 = \mathbb{P}(X=4, Y=4)$, X e Y non sono indipendenti!

Soluzione esercizio 6.
Sia c quel numero reale tale che $\mathbb{P}(X=c)=1$, allora:

- se $x \geq c$ si ha $\mathbb{P}(X\leq x)=1$ e $\mathbb{P}(X\leq x, Y\leq y) = \mathbb{P}(\Omega\cap\{Y\leq y\})\cdot 1 = \mathbb{P}(Y\leq y)\cdot\mathbb{P}(X\leq x)$, per ogni $y\in\mathbb{R}$,
- se $x < c$ si ha $\mathbb{P}(X\leq x)=0$ e $\mathbb{P}(X\leq x, Y\leq y) = \mathbb{P}(\emptyset\cap\{Y\leq y\}) = 0 = \mathbb{P}(Y\leq y)\cdot 0 = \mathbb{P}(Y\leq y)\cdot\mathbb{P}(X\leq x)$.

In ogni caso si ha che $\mathbb{P}(X\leq x, Y\leq y) = \mathbb{P}(Y\leq y)\cdot\mathbb{P}(X\leq x)$, per ogni $(x,y)\in\mathbb{R}\times\mathbb{R}$ e quindi X ed Y sono indipendenti.

Soluzione esercizio 7.
Sia (h,k) una coppia di valori per cui $\mathbb{P}(X=h, Y=k)\neq 0$. Dal momento

che Ω è dotato della probabilità uniforme, $\mathbb{P}(X = h, Y = k)$ è della forma $\frac{c}{p}$, $\mathbb{P}(X = h) = \frac{a}{p}$ e $\mathbb{P}(Y = k) = \frac{b}{p}$ con a, b, c interi compresi fra 0 e p. Notiamo che $c \leq \min(a, b)$; evidentemente, dalla condizione di indipendenza si ha che:

$$\mathbb{P}(X = h, Y = k) = \frac{c}{p} = \frac{a}{p} \cdot \frac{b}{p} = \mathbb{P}(X = h) \cdot \mathbb{P}(Y = k)$$

da cui $p \cdot c = a \cdot b$. Poiché $c \neq 0$ se e solo se $a, b \neq 0$, si ha che se $c \neq 0$ allora p é un divisore di ab, ma è primo quindi p è divisore di a o di b. Quindi $a = p$ oppure $b = p$. Poiché esiste h, k tale che $c \neq 0$ si ha che la condizione è necessaria.

Dall'esercizio precedente si ha la sufficienza.