Esercizi - Fascicolo IV

Esercizio 1 Una compagnia di assicurazioni emette una polizza che pagherà n euro se l'evento E si verificherà entro un anno. Se la compagnia stima che l'evento E si verificherà entro un anno con probabilità p , quanto dovrebbe essere il costo della polizza per il cliente in modo che il profitto atteso per la compagnia sia $n/10$?

Soluzione. Sia X il profitto della compagnia e c il costo della polizza. Allora

$$P(X = c) = 1 - p, \quad P(X = c - n) = p.$$

Allora

$$E(X) = c(1 - p) + (c - n)p = c - np$$

Dunque

$$E(X) = n/10 \iff c = n \left(p + \frac{1}{10} \right).$$

Esercizio 2 Una persona lancia una moneta equilibrata finchè non ottiene croce la prima volta. Se appare croce all' n -mo lancio, la persona vince 2^n euro. Denotiamo con X il guadagno del giocatore. Mostrare che $E(X) = +\infty$.

Soluzione. Basta osservare che $P(X = 2^n) = \frac{1}{2^n}$.

Esercizio 3 Si estraggono 3 biglie, senza reinserimento, da un'urna contenente 5 biglie rosse, 4 biglie bianche e 2 biglie verdi. Sia X il numero di biglie rosse estratte e Y il numero di biglie bianche estratte.

- (a) Scrivere la densità (discreta) di (X, Y) .
- (b) Calcolare le densità marginali di X e Y e i valori attesi $E(X)$ ed $E(Y)$.
- (c) Calcolare la covarianza di X e Y .

Soluzione.

- (a) Sappiamo che i casi possibili, cioè le terne distinte senza tener conto dell'ordine di estrazione, sono $\binom{11}{3}$. Osservando r biglie rosse, b bianche e v verdi, i casi favorevoli sono:

$$\binom{5}{r} \binom{4}{b} \binom{2}{v}$$

$\Pr(X = i, Y = j) \neq 0$ se e solo se $(i, j) \in \{(i, j) : 1 \leq i \leq 3, 1 \leq j \leq 3, 1 \leq i + j \leq 3\}$.

$$\Pr(X = 0, Y = 1) = \frac{\binom{5}{0} \binom{4}{1} \binom{2}{2}}{\binom{11}{3}} = \frac{4}{165} \simeq 0.024;$$

$$\Pr(X = 0, Y = 2) = \frac{\binom{5}{0}\binom{4}{2}\binom{2}{1}}{\binom{11}{3}} = \frac{4}{55} \simeq 0.073;$$

$$\Pr(X = 0, Y = 3) = \frac{\binom{5}{0}\binom{4}{3}\binom{2}{0}}{\binom{11}{3}} = \frac{4}{165} \simeq 0.024;$$

$$\Pr(X = 1, Y = 0) = \frac{\binom{5}{1}\binom{4}{0}\binom{2}{2}}{\binom{11}{3}} = \frac{1}{33} \simeq 0.03;$$

$$\Pr(X = 1, Y = 1) = \frac{\binom{5}{1}\binom{4}{1}\binom{2}{1}}{\binom{11}{3}} = \frac{8}{33} \simeq 0.242$$

$$\Pr(X = 1, Y = 2) = \frac{\binom{5}{1}\binom{4}{2}\binom{2}{0}}{\binom{11}{3}} = \frac{2}{11} \simeq 0.182;$$

$$\Pr(X = 2, Y = 0) = \frac{\binom{5}{2}\binom{4}{0}\binom{2}{1}}{\binom{11}{3}} = \frac{4}{33} \simeq 0.121;$$

$$\Pr(X = 2, Y = 1) = \frac{\binom{5}{2}\binom{4}{1}\binom{2}{0}}{\binom{11}{3}} = \frac{8}{33} \simeq 0.242;$$

$$\Pr(X = 3, Y = 0) = \frac{\binom{5}{3}\binom{4}{0}\binom{2}{0}}{\binom{11}{3}} = \frac{2}{33} \simeq 0.061.$$

(b) Si ha:

$$\Pr(X = 0) = \Pr(X = 0, Y = 1) + \Pr(X = 0, Y = 2) + \Pr(X = 0, Y = 3) = \frac{4}{33} \simeq 0.121;$$

$$\Pr(X = 1) = \Pr(X = 1, Y = 0) + \Pr(X = 1, Y = 1) + \Pr(X = 1, Y = 2) = \frac{5}{11} \simeq 0.455;$$

$$\Pr(X = 2) = \Pr(X = 2, Y = 0) + \Pr(X = 2, Y = 1) = \frac{4}{11} \simeq 0.364;$$

$$\Pr(X = 3) = \Pr(X = 3, Y = 0) = \frac{2}{33} \simeq 0.061.$$

Per gli altri valori i si ha $\Pr(X = i) = 0$.

$$\Pr(Y = 0) = \Pr(X = 1, Y = 0) + \Pr(X = 2, Y = 0) = \Pr(X = 3, Y = 0) = \frac{7}{33} \simeq 0.212;$$

$$\Pr(Y = 1) = \Pr(X = 0, Y = 1) + \Pr(X = 1, Y = 1) + \Pr(X = 2, Y = 1) = \frac{28}{55} \simeq 0.509;$$

$$\Pr(Y = 2) = \Pr(X = 0, Y = 2) + \Pr(X = 1, Y = 2) = \frac{14}{55} \simeq 0.255;$$

$$\Pr(Y = 3) = \Pr(X = 0, Y = 3) = \frac{4}{165} \simeq 0.024.$$

Per gli altri valori j si ha $\Pr(Y = j) = 0$.

Per i valori attesi si ottiene:

$$E(X) = 1 \cdot \Pr(X = 1) + 2 \cdot \Pr(X = 2) + 3 \cdot \Pr(X = 3) = \frac{15}{11} \simeq 1.364;$$

$$E(Y) = 1 \cdot \Pr(Y = 1) + 2 \cdot \Pr(Y = 2) + 3 \cdot \Pr(Y = 3) = \frac{12}{11} \simeq 1.091.$$

(c) XY può assumere solo i valori 0, 1 e 2.

$$\begin{aligned} E(XY) &= \\ &= \Pr(X = 1, Y = 1) + 2[\Pr(X = 1, Y = 2) + \Pr(X = 2, Y = 1)] = \\ &= \frac{12}{11} \simeq 1.091, \end{aligned}$$

quindi:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = -\frac{48}{121} \simeq -0.397.$$

Ne segue che X e Y sono correlate e quindi dipendenti.

Esercizio 4 Il piatto di una roulette contiene i numeri da 0 a 36. Consideriamo due possibili puntate: A) quella sui numeri pari: si vince se esce un numero pari tra 2 e 36, e in questo caso si riceve il doppio del capitale puntato (quindi si resta in attivo di una quantità uguale al capitale puntato); B) quella sulla prima dozzina: si vince se esce un numero tra 1 e 12, e in questo caso si riceve il triplo del capitale puntato (quindi si resta in attivo del doppio del capitale puntato).

Supponiamo di puntare contemporaneamente due Euro sui numeri pari, e un Euro sulla prima dozzina. Sia X il bilancio della giocata (capitale finale - capitale iniziale).

- Determinare la densità di X e $E(X)$.
- Si considerino le seguenti variabili casuali:

$$Y = \begin{cases} 1 & \text{se esce un numero pari} \\ 0 & \text{altrimenti} \end{cases}$$

$$Z = \begin{cases} 1 & \text{se esce un numero della prima dozzina} \\ 0 & \text{altrimenti} \end{cases}$$

Calcolare $\text{Cov}(Y, Z)$.

- Dopo aver espresso X come funzione di Y e Z , calcolare $E(X)$ e $\text{Var}(X)$ senza usare la densità calcolata in a).

Soluzione.

- Introduciamo gli eventi

$$A = \{\text{esce un numero pari (tra 2 e 36)}\}, \quad B = \{\text{esce un numero tra 1 e 12}\}$$

È chiaro che l'esito degli eventi A e B determina X : più precisamente,

$$X = -3 \cdot 1_{(A \cup B)^c} + 1 \cdot 1_{A \setminus B} + 0 \cdot 1_{B \setminus A} + 4 \cdot 1_{A \cap B}.$$

Notando che gli eventi $\{(A \cup B)^c, A \setminus B, B \setminus A, A \cap B\}$ sono a due a due disgiunti, si ha che $X(\Omega) = \{-3, 0, 1, 4\}$ e

$$\begin{aligned} p_X(-3) &= P((A \cup B)^c) = \frac{13}{37}, & p_X(0) &= P(B \setminus A) = \frac{6}{37} \\ p_X(1) &= P(A \setminus B) = \frac{12}{37}, & p_X(4) &= P(A \cap B) = \frac{6}{37}, \end{aligned}$$

da cui $E(X) = -3p_X(-3) + p_X(1) + 4p_X(4) = -\frac{3}{37}$.

- Dato che Y e Z sono variabili di Bernoulli, si ha che

$$\begin{aligned} \text{Cov}(Y, Z) &= E(YZ) - E(Y)E(Z) = P(Y = 1, Z = 1) - P(Y = 1)P(Z = 1) \\ &= \frac{6}{37} - \frac{18}{37} \frac{12}{37} = \frac{6}{1369}. \end{aligned}$$

c) Si ha che

$$X = (4Y - 2) + (3Z - 1) = 4Y + 3Z - 3.$$

Dato che $Y \sim Be(\frac{18}{37})$, $Z \sim Be(\frac{12}{37})$, si ha che

$$E(X) = 4E(Y) + 3E(Z) - 3 = 4\frac{18}{37} + 3\frac{12}{37} - 3 = -\frac{3}{37}.$$

Analogamente si ha che

$$\begin{aligned} Var(X) &= 16 Var(Y) + 9 Var(Z) + 2 \cdot 12 \cdot Cov(Y, Z) \\ &= 16\frac{18}{37}\frac{19}{37} + 9\frac{12}{37}\frac{25}{37} + 24\frac{6}{1369} = \frac{8316}{1369}. \end{aligned}$$

Esercizio 5 Siano $X, Y \sim Be(p)$ indipendenti, con $p \in (0, 1)$. Mostrare che le variabili casuali $X + Y$ e $X - Y$ sono scorrelate ma non indipendenti.

Soluzione.

$$Cov(X + Y, X - Y) = Var(X) - Var(Y) = 0,$$

ma, ad esempio

$$P(X + Y = 0, X - Y = 1) = 0 \neq P(X + Y = 0)P(X - Y = 1).$$

Esercizio 6 In un canale di trasmissione vengono trasmessi simboli binari. I disturbi sul canale fanno sì che ogni simbolo trasmesso ha la probabilità del 2% di essere ricevuto errato, indipendentemente dagli altri simboli. I messaggi vengono trasmessi in "parole" composte da 50 simboli. Qual è la probabilità che una parola venga ricevuta con almeno due simboli errati? (Usare l'approssimazione di Poisson)

Soluzione. Sia, per $i = 1, \dots, 50$,

$$X_i = \begin{cases} 0 & \text{se l}'i\text{-mo simbolo é ricevuto correttamente} \\ 1 & \text{altrimenti} \end{cases}$$

Posto $Y = \sum_{i=1}^{50} X_i$, dobbiamo calcolare $P(Y \geq 2)$. Essendo $X_i \sim Be(0.02)$, allora $Y \sim B(50, 0.02)$. Usando l'approssimazione di Poisson, $Y \approx Po(1)$. Dunque

$$P(Y \geq 2) \simeq 1 - 2e^{-1}.$$

Esercizio 7 Sia X una variabile casuale a valori in $\mathbb{N}$, tale che

$$P(X = k) = \frac{e^{-2}2^k}{2k!}(1 + \lambda k),$$

con $\lambda > 0$.

- Determinare il valore di λ
- Sia $Y \sim Po(2)$. Mostrare che

$$P(X = k) = \frac{1}{2}[P(Y = k) + P(Y + 1 = k)].$$

- Usando il risultato in b., calcolare media e varianza di X

Soluzione. a. Dev'essere

$$\sum_{k=0}^{+\infty} P(X = k) = 1.$$

Ma

$$\sum_{k=0}^{+\infty} \frac{e^{-2} 2^k}{2k!} (1 + \lambda k) = \frac{1}{2} \sum_k e^{-2} \frac{2^k}{k!} + \frac{\lambda}{2} \sum_k k e^{-2} \frac{2^k}{k!} = \frac{1}{2} + \lambda,$$

ove si è usato il fatto che $\sum_k k e^{-2} \frac{2^k}{k!}$ è la media di una variabile di Poisson di parametro 2. Dunque $\lambda = 1/2$.

b. Dal punto a., per $k \geq 1$,

$$\begin{aligned} P(X = k) &= \frac{1}{2} e^{-2} \frac{2^k}{k!} + \frac{1}{4} e^{-2} \frac{2^k}{(k-1)!} = \frac{1}{2} e^{-2} \frac{2^k}{k!} + \frac{1}{2} e^{-2} \frac{2^{k-1}}{(k-1)!} \\ &= \frac{1}{2} P(Y = k) + \frac{1}{2} P(Y = k-1). \end{aligned}$$

Per $k = 0$ l'identità da dimostrare è banale.

c. Dal punto b.

$$E(X) = \frac{1}{2} \sum_k k P(Y = k) + \frac{1}{2} \sum_k k P(Y + 1 = k) = \frac{1}{2} E(Y) + \frac{1}{2} E(Y + 1) = 1 + 3/2 = 5/2.$$

Analogamente

$$E(X^2) = \frac{1}{2} E(Y^2) + \frac{1}{2} E((Y + 1)^2) = E(Y^2) + E(Y) + 1/2 = 6 + 2 + 1/2 = 17/2.$$

Infine

$$\text{Var}(X) = 17/2 - 25/4 = 9/4.$$

Esercizio 8 Sia $\lambda > 0$ e $c : \mathbb{N} \rightarrow [0, +\infty)$ una funzione tale che $c(0) = 0$ e $\inf_{n>0} c(n) > \lambda$. Sia X una variabile casuale a valori in $\mathbb{N}$ con densità

$$p_X(n) = \begin{cases} \frac{1}{Z} & \text{se } n = 0 \\ \frac{1}{Z} \frac{\lambda^n}{c(1)c(2)\dots c(n)} & \text{se } n \geq 1, \end{cases}$$

dove $Z = 1 + \sum_{n=1}^{+\infty} \frac{\lambda^n}{c(1)c(2)\dots c(n)}$.

a. Mostrare che $E(c(X)) = \lambda$.

b. Mostrare che per ogni funzione $f : \mathbb{N} \rightarrow \mathbb{R}$ tale che la v.c. $f(X + 1)$ ammette valor medio, si ha

$$E(c(X)f(X)) = \lambda E(f(X + 1)).$$

c.* Si assuma che, per ogni $n \geq 0$, $|c(n + 1) - c(n)| \leq 1$. Sia $Y \sim Po(\lambda)$. Mostrare, per induzione su k , che per ogni $k \geq 1$

$$E[c(X)^k] \leq E(Y^k).$$

(Sugg.: usare, dopo averla verificata, l'uguaglianza $E(Y^{k+1}) = \lambda \sum_{i=0}^k \binom{k}{i} E(Y^i)$)

Soluzione. a.

$$E(c(X)) = \frac{1}{Z} \sum_{n=1}^{+\infty} c(n) \frac{\lambda^n}{c(1)c(2)\dots c(n)} = \frac{1}{Z} \left[\lambda + \sum_{n=2}^{+\infty} \frac{\lambda^n}{c(1)c(2)\dots c(n-1)} \right]$$

$$= \frac{\lambda}{Z} \left[1 + \sum_{n=1}^{+\infty} \frac{\lambda^n}{c(1)c(2) \cdots c(n)} \right] = a.$$

b. Si procede in modo analogo al punto a.

$$\begin{aligned} E(f(X)c(X)) &= \frac{1}{Z} \sum_{n=1}^{+\infty} f(n)c(n) \frac{\lambda^n}{c(1)c(2) \cdots c(n)} = \frac{1}{Z} \left[\lambda f(1) + \sum_{n=2}^{+\infty} f(n) \frac{\lambda^n}{c(1)c(2) \cdots c(n-1)} \right] \\ &= \frac{\lambda}{Z} \left[f(1) + \sum_{n=1}^{+\infty} f(n+1) \frac{\lambda^n}{c(1)c(2) \cdots c(n)} \right] = \lambda E(f(X+1)). \end{aligned}$$

c. Cominciamo coll'osservare che (vedi esercizio 2.3.4).

$$E[(Y+1)^k] = \lambda \sum_{i=0}^k \binom{k}{i} E(Y^i).$$

Per il punto b. e il fatto che $c(X+1) \leq c(X) + 1$, si ha

$$\begin{aligned} E[c(X)^{k+1}] &= aE[c(X+1)^k] \leq aE\{[c(X)+1]^k\} = a \sum_{i=0}^k \binom{k}{i} E[c(X)^i] \leq \text{ipotesi induttiva} \\ &\leq a \sum_{i=0}^k \binom{k}{i} E(Y^i) = E(Y^{k+1}). \end{aligned}$$

Resta da verificare il passo base dell'induzione ($k=1$). Ma $E(c(X)) = a = E(Y)$.

Esercizio 9 a. Siano $X_1, X_2 \sim Po(\lambda)$ indipendenti. Fissati $n \in \mathbb{N}$ e $0 \leq k \leq n$, calcolare

$$(1) \quad P(X_1 = k | X_1 + X_2 = n).$$

b. Supponiamo che n sia fissato, e chiamiamo $q(k)$ il valore dell'espressione in (1). Mostrare che $q(\cdot)$ è il valore della densità di una variabile casuale binomiale, e determinarne i parametri.

c. Siano ora $X_1, X_2, \dots, X_m \sim Po(\lambda)$ indipendenti, con $m > 2$. Calcolare

$$q_m(k) = P(X_1 = k | X_1 + X_2 + \cdots + X_m = n),$$

e determinare i parametri della variabile casuale binomiale la cui densità è $q_m(\cdot)$.

Soluzione. a. e b.

$$\begin{aligned} P(X_1 = k | X_1 + X_2 = n) &= \frac{P(X_1 = k, X_1 + X_2 = n)}{P(X_1 + X_2 = n)} = \frac{P(X_1 = k, X_2 = n - k)}{P(X_1 + X_2 = n)} \\ &= \frac{P(X_1 = k)P(X_2 = n - k)}{P(X_1 + X_2 = n)} = \frac{e^{-\lambda} \frac{\lambda^k}{k!} e^{-\lambda} \frac{\lambda^{n-k}}{(n-k)!}}{e^{-2\lambda} \frac{(2\lambda)^n}{n!}} = \binom{n}{k} 2^{-n}, \end{aligned}$$

dove abbiamo usato l'indipendenza di X_1 e X_2 e il fatto che $X_1 + X_2 \sim Po(2\lambda)$. Si noti che il valore ottenuto è la densità di una v.c. $B(n, 1/2)$.

c. In modo analogo al punto precedente, usando il fatto che $X_2 + \dots + X_m \sim Po((m-1)\lambda)$,

$$\begin{aligned} P(X_1 = k | X_1 + X_2 + \dots + X_m = n) &= \frac{P(X_1 = k)P(X_2 + \dots + X_m = n - k)}{P(X_1 + \dots + X_m = n)} \\ &= \binom{n}{k} \frac{1}{m^k} \left(1 - \frac{1}{m}\right)^{n-k}, \end{aligned}$$

che è la densità di una $B(n, 1/m)$.

Esercizio 10 Siano X, Z e W variabili casuali indipendenti con $X \sim Be(p)$, $Z, W \sim Po(\lambda)$. Definiamo

$$Y = XZ + W.$$

- Determinare le densità $p_{X,Y}$ e p_Y .
- Utilizzando la densità calcolata al punto a., determinare $E(Y)$ e $Var(Y)$.
- Calcolare $E(Y)$ e $Var(Y)$ *senza* utilizzare p_Y .

Soluzione. a. Si ha

$$p_{X,Y}(0, y) = P(X = 0, Y = y) = P(X = 0, W = y) = P(X = 0)P(W = y) = (1 - p)e^{-\lambda} \frac{\lambda^y}{y!},$$

$$p_{X,Y}(1, y) = P(X = 1, Y = y) = P(X = 1, Z + W = y) = pe^{-2\lambda} \frac{(2\lambda)^y}{y!},$$

ove si è usato il fatto che X, Z, W sono indipendenti e che $Z + W \sim Po(2\lambda)$. Segue che

$$p_Y(y) = p_{X,Y}(0, y) + p_{X,Y}(1, y) = (1 - p)e^{-\lambda} \frac{\lambda^y}{y!} + pe^{-2\lambda} \frac{(2\lambda)^y}{y!}.$$

b.

$$\begin{aligned} E(Y) &= (1 - p) \sum_{y=0}^{+\infty} ye^{-\lambda} \frac{\lambda^y}{y!} + p \sum_{y=0}^{+\infty} ye^{-2\lambda} \frac{(2\lambda)^y}{y!} \\ &= (1 - p)\lambda + 2\lambda p = \lambda(p + 1), \end{aligned}$$

dove si è usato il fatto che le due somme nella formula precedente coincidono con le medie di v.c. di Poisson di parametri λ e 2λ rispettivamente. Un analogo ragionamento con i momenti secondi conduce a

$$E(Y^2) = (1 - p)(\lambda^2 + \lambda) + p(4\lambda^2 + 2\lambda) = \lambda^2(1 + 3p) + \lambda(1 + p).$$

Perciò

$$Var(Y) = \lambda^2(p - p^2) - \lambda(1 + p).$$

c. Usando il fatto che la media del prodotto di due v.c. indipendenti è uguale al prodotto dei valori medi, si ha

$$E(Y) = E(XZ) + E(W) = E(X)E(Z) + E(W) = p\lambda + \lambda = \lambda(p + 1)$$

e

$$\begin{aligned} E(Y^2) &= E(X^2 Z^2) + E(W^2) + 2E(XZW) = E(X^2)E(Z^2) + E(W^2) + 2E(X)E(Z)E(W) \\ &= p(\lambda^2 + \lambda) + \lambda^2 + \lambda + 2p\lambda^2 = \lambda^2(1 + 3p) + \lambda(1 + p), \end{aligned}$$

e si ritrova quanto trovato sopra.

Esercizio 11 Siano X e Y due variabili casuali a valori in $\mathbb{N}$ aventi la seguente densita' congiunta:

$$P(X = k, Y = n) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda} \frac{\lambda^n}{n!} & \text{se } 0 \leq k \leq n \\ 0 & \text{altrimenti,} \end{cases}$$

dove $p \in (0, 1)$ e $\lambda > 0$ sono due parametri fissati.

a. Determinare le densita' marginali di X e Y . Mostrare, in particolare, che $X \sim Po(p\lambda)$ e $Y \sim Po(\lambda)$.

b. Calcolare il coefficiente di correlazione $\rho_{X,Y}$.

Soluzione. a.

$$\begin{aligned} P(X = k) &= \sum_{n=k}^{+\infty} \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda} \frac{\lambda^n}{n!} = \frac{\lambda^k p^k}{k!} e^{-\lambda} \sum_{n=k}^{+\infty} \frac{\lambda^{n-k} (1-p)^{n-k}}{(n-k)!} \\ &= \frac{\lambda^k p^k}{k!} e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n (1-p)^n}{n!} = \frac{\lambda^k p^k}{k!} e^{-\lambda} e^{(1-p)\lambda} = e^{-p\lambda} \frac{(p\lambda)^k}{k!}. \end{aligned}$$

$$P(Y = n) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \frac{\lambda^n}{n!} \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^n}{n!}.$$

b. Dal punto a. abbiamo che $Var(X) = p\lambda$ e $Var(Y) = \lambda$. Inoltre

$$E(XY) = \sum_{k,n} kn P(X = k, Y = n) = \sum_{n=0}^{+\infty} n e^{-\lambda} \frac{\lambda^n}{n!} \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k}.$$

Osservando che $\sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k}$ è la media di una $B(n, p)$, abbiamo

$$E(XY) = p \sum_{n=0}^{+\infty} n^2 e^{-\lambda} \frac{\lambda^n}{n!} = p(\lambda + \lambda^2),$$

dove abbiamo usato il fatto che $\sum_{n=0}^{+\infty} n^2 e^{-\lambda} \frac{\lambda^n}{n!} = p(\lambda + \lambda^2)$ è il momento secondo di una $Po(\lambda)$. Ne segue che

$$Cov(X, Y) = E(XY) - E(X)E(Y) = p(\lambda + \lambda^2) - p\lambda^2 = p\lambda.$$

Infine

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = \frac{p\lambda}{\sqrt{p\lambda^2}} = \sqrt{p}.$$

Esercizio 12 Un'urna contiene 100 palline numerate da 0 a 99. Si estrae una pallina, e si denotano con X e Y le due cifre del numero estratto (la cifra delle decine, X , si considera uguale a zero per numeri minori di 10). Mostrare che X e Y sono indipendenti, e determinarne la distribuzione.

Soluzione. Sia Z il valore del numero estratto. Chiaramente $P(Z = i) = \frac{1}{100}$ per $i = 0, 1, \dots, 99$. X e Y assumono valori interi tra 0 e 9. Per $n, m \in \{0, 1, \dots, 9\}$ si ha

$$P(X = n, Y = m) = P(Z = 10n + m) = \frac{1}{100},$$

$$P(X = n) = P(10n \leq Z \leq 10n + 9) = \frac{1}{10},$$

$$P(Y = m) = P(Z \in \{m, m + 10, m + 20, \dots, m + 90\}) = \frac{1}{10},$$

da cui segue l'indipendenza.

Esercizio 13 Dimostrare che le seguenti affermazioni sono equivalenti:

- i) Gli eventi A e B sono indipendenti.
- ii) Le variabili casuali 1_A e 1_B sono indipendenti.
- iii) Le variabili casuali 1_A e 1_B sono scorrelate (cioè $Cov(1_A, 1_B) = 0$).

Soluzione.

i) $\Rightarrow$ ii) Dimostriamo che la densità congiunta è il prodotto delle densità marginali.

$$\begin{aligned} P(1_A = 1, 1_B = 1) &= P(A \cap B) = (\text{per l'indipendenza}) \\ &= P(A)P(B) = P(1_A = 1)P(1_B = 1). \end{aligned}$$

Analogamente:

$$P(1_A = 0, 1_B = 1) = P(A^c \cap B) = P(A^c)P(B) = P(1_A = 0)P(1_B = 1),$$

e allo stesso modo negli altri casi.

ii) $\Rightarrow$ iii) È una proprietà generale.

iii) $\Rightarrow$ i) Si noti che

$$E(1_A 1_B) = E(1_{A \cap B}) = P(A \cap B),$$

e

$$E(1_A)E(1_B) = P(A)P(B).$$

Sicché:

$$0 = Cov(1_A, 1_B) = E(1_A 1_B) - E(1_A)E(1_B) = P(A \cap B) - P(A)P(B),$$

da cui segue la conclusione.

Esercizio 14 Siano $X \sim Ge(p)$ e $Y \sim Ge(q)$, Determinare la densità di $Z = \min(X, Y)$.

Soluzione. Si procede esattamente come nell'Esempio 2.8.2:

$$F_Z(k) = 1 - [(1-p)(1-q)]^k,$$

da cui $Z \sim Ge(1 - [(1-p)(1-q)])$.

Esercizio 15 Siano $X, Y \sim Ge(p)$ indipendenti, $0 < p < 1$, e $Z := \max(X, Y)$.

- a. Determinare la densità congiunta di (X, Z) .
- b. Posto $W := Z - X$, determinare la densità p_W di W .
- c. Calcolare $E(W)$.
- d. È vero o falso che X e W sono indipendenti?

Soluzione.

a.

$$P(X = n, Z = m) = \begin{cases} P(X = n, Y = m) = p^2(1-p)^{n+m} & \text{se } m > n \geq 0 \\ P(X = n, Y \leq n) = p(1-p)^n[1 - (1-p)^{n+1}] & \text{se } m = n \geq 0 \\ 0 & \text{altrimenti.} \end{cases}$$

b. Per $k \geq 1$

$$P(W = k) = \sum_{n=0}^{+\infty} P(X = n, Z = n + k) = \sum_{n=0}^{+\infty} p^2(1-p)^{2n}(1-p)^k = \frac{p}{2-p}(1-p)^k.$$

Perciò

$$P(W = 0) = 1 - \sum_{k=1}^{+\infty} P(W = k) = 1 - \frac{p}{2-p} \left(\frac{1}{p} - 1 \right) = 1 - \frac{1-p}{2-p} = \frac{1}{2-p}.$$

c.

$$E(W) = \sum_{k=1}^{+\infty} kP(W = k) = \frac{1}{2-p} p \sum_{k=1}^{+\infty} k(1-p)^k = \frac{1-p}{p(2-p)},$$

dove si è usato il fatto che $p \sum_{k=1}^{+\infty} k(1-p)^k$ è la media di una $Ge(p)$.

d. Falso. Infatti, ad esempio

$$P(X = 0, W = 0) = P(X = 0, Y = 0) = p^2 \neq P(X = 0)P(W = 0) = p \frac{1}{2-p}$$

per ogni $p \in (0, 1)$.

Esercizio 16 Siano $X, Y \sim Ge(p)$ indipendenti, e si definisca $Z := \max(X, Y) - Y$.

a. Determinare la densità discreta di Z .

b. Determinare $E(Z)$ (sugg.: ricordare che $E(Z) = \sum_{n=1}^{+\infty} P(Z \geq n)$, oppure usare la funzione generatrice dei momenti)

Soluzione. a. Se $n \geq 1$,

$$\begin{aligned} P(Z = n) &= \sum_{k=0}^{+\infty} P(X = k + n, Y = k) = \sum_{k=0}^{+\infty} P(X = k + n)P(Y = k) \\ &= \sum_{k=0}^{+\infty} p^2(1-p)^{k+n}(1-p)^k = p^2(1-p)^n \sum_{k=0}^{+\infty} (1-p)^{2k} = \frac{p(1-p)^n}{2-p}. \end{aligned}$$

Inoltre

$$\begin{aligned} P(Z = 0) &= 1 - \sum_{n=1}^{+\infty} P(Z = n) = 1 - \frac{p}{2-p} \sum_{n=1}^{+\infty} (1-p)^n \\ &= 1 - \frac{p}{2-p} \frac{(1-p)}{p} = \frac{1}{2-p}. \end{aligned}$$

b. Notare che, per $n \geq 1$,

$$P(Z \geq n) = \frac{p}{2-p} \sum_{k=n}^{+\infty} (1-p)^k = \frac{(1-p)^n}{2-p}.$$

Perciò

$$E(Z) = \sum_{n=1}^{+\infty} P(Z \geq n) = \frac{1-p}{p(2-p)}.$$

Esercizio 17 Sia X una variabile casuale che assume i valori 0 e 1 e Y una variabile casuale a valori in $\mathbb{N}$, la cui densità congiunta è determinata dalle identità: per $n \in \mathbb{N}$

$$\begin{aligned} P(X=0, Y=n) &= p \frac{e^{-1}}{n!} \\ P(X=1, Y=n) &= (1-p)e^{-2} \frac{2^n}{n!}. \end{aligned}$$

- Determinare le densità marginali p_X e p_Y .
- Calcolare la probabilità condizionata $P(X=0|Y>0)$.
- Calcolare $E(Y)$ e $Var(Y)$.
- Calcolare $Cov(X, Y)$ e il coefficiente di correlazione $\rho(X, Y)$.

Soluzione.

a.

$$p_X(0) = \sum_n P(X=0, Y=n) = p = 1 - p_X(1).$$

$$p_Y(n) = P(X=0, Y=n) + P(X=1, Y=n) = p \frac{e^{-1}}{n!} + (1-p)e^{-2} \frac{2^n}{n!}.$$

b.

$$\begin{aligned} P(X=0|Y>0) &= \frac{P(X=0, Y>0)}{P(Y>0)} = \frac{P(X=0) - P(X=0, Y=0)}{1 - P(Y=0)} \\ &= \frac{p - pe^{-1}}{1 - pe^{-1} - (1-p)e^{-2}}. \end{aligned}$$

c.

$$\begin{aligned} \gamma_Y(t) &= \sum_{n=0}^{+\infty} e^{tn} p_Y(n) = pe^{-1} \sum_{n=0}^{+\infty} \frac{e^{tn}}{n!} + (1-p)e^{-2} \sum_{n=0}^{+\infty} \frac{2^n e^{tn}}{n!} \\ &= p \exp(e^t - 1) + (1-p) \exp(2e^t - 2). \end{aligned}$$

Non è difficile calcolare

$$E(Y) = \gamma'_Y(0) = p + 2(1-p) = 2-p, \quad E(Y^2) = \gamma''_Y(0) = 2p + 6(1-p) = 6-4p,$$

da cui si calcola $Var(Y) = E(Y^2) - [E(Y)]^2$.

d. Abbiamo appena calcolato $E(Y)$ e $Var(Y)$. Facilmente, dal punto a., si ha $E(X) = 1-p$ e $Var(X) = p(1-p)$. Inoltre

$$E(XY) = \sum_{x,y} xyp_{X,Y}(x,y) = \sum_{n=0}^{+\infty} np_{X,Y}(1,n) = (1-p)e^{-2} \sum_{n=0}^{+\infty} n \frac{2^n}{n!} = 2(1-p),$$

dove abbiamo usato il fatto che $e^{-2} \sum_{n=0}^{+\infty} n \frac{2^n}{n!}$ è la media di una $Po(2)$. A questo punto è sufficiente sostituire i valori ottenuti per calcolare

$$Cov(X, Y) = E(XY) - E(X)E(Y),$$

e

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.$$

Esercizio 18 Sia X una variabile casuale che assume i valori 0 e 1 e Y una variabile casuale a valori in $\mathbb{N}$, la cui densità congiunta è determinata dalle identità: per $n \in \mathbb{N}$

$$\begin{aligned} P(X = 0, Y = n) &= \frac{p}{3} \left(\frac{2}{3}\right)^n \\ P(X = 1, Y = n) &= \frac{1-p}{4} \left(\frac{3}{4}\right)^n. \end{aligned}$$

- Determinare le densità marginali p_X e p_Y .
- Calcolare la probabilità condizionata $P(X = 0|Y > 0)$.
- Calcolare $E(Y)$ e $Var(Y)$.
- Calcolare $Cov(X, Y)$ e il coefficiente di correlazione $\rho(X, Y)$.

Soluzione.

a.

$$p_X(0) = \sum_n P(X = 0, Y = n) = p = 1 - p_X(1).$$

$$p_Y(n) = P(X = 0, Y = n) + P(X = 1, Y = n) = \frac{p}{3} \left(\frac{2}{3}\right)^n + \frac{1-p}{4} \left(\frac{3}{4}\right)^n.$$

b.

$$P(X = 0|Y > 0) = \frac{P(X = 0, Y > 0)}{P(Y > 0)} = \frac{P(X = 0) - P(X = 0, Y = 0)}{1 - P(Y = 0)} = \frac{p - \frac{p}{3}}{1 - \frac{p}{3} - \frac{1-p}{4}}.$$

c.

$$\gamma_Y(t) = \sum_{n=0}^{+\infty} e^{tn} p_Y(n) = \frac{p}{3} \sum_{n=0}^{+\infty} \left(\frac{2}{3} e^t\right)^n + \frac{1-p}{4} \sum_{n=0}^{+\infty} \left(\frac{3}{4} e^t\right)^n = \frac{p}{3} \frac{1}{1 - \frac{2}{3} e^t} + \frac{1-p}{4} \frac{1}{1 - \frac{3}{4} e^t},$$

per t tale che $\frac{3}{4} e^t < 1$, mentre $\gamma_X(t) = +\infty$ altrimenti. Non è difficile calcolare

$$E(Y) = \gamma_Y'(0) = 3 - p, \quad E(Y^2) = \gamma_Y''(0) = 21 - 13p,$$

da cui si calcola $Var(Y) = E(Y^2) - [E(Y)]^2$.

d. Abbiamo appena calcolato $E(Y)$ e $Var(Y)$. Facilmente, dal punto a., si ha $E(X) = 1 - p$ e $Var(X) = p(1 - p)$. Inoltre

$$E(XY) = \sum_{x,y} xyp_{X,Y}(x, y) = \sum_{n=0}^{+\infty} np_{X,Y}(1, n) = \frac{(1-p)}{4} \sum_{n=0}^{+\infty} n \left(\frac{3}{4}\right)^n = 3(1-p),$$

dove abbiamo usato il fatto che $\frac{1}{4} \sum_{n=0}^{+\infty} n \left(\frac{3}{4}\right)^n$ è la media di una $Ge(1/4)$. A questo punto è sufficiente sostituire i valori ottenuti per calcolare

$$Cov(X, Y) = E(XY) - E(X)E(Y),$$

e

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Esercizio 19 Sia $X_n \sim \text{Po}(n\lambda)$, e $Y_n = \frac{X_n - n\lambda}{n^{2/3}}$. Usando la Disuguaglianza di Chebyshev mostrare che per ogni $\epsilon > 0$

$$\lim_{n \rightarrow +\infty} P(|Y_n| > \epsilon) = 0.$$

Soluzione. Si noti che $E(Y_n) = 0$ e $\text{Var}(Y_n) = \frac{n\lambda}{n^{4/3}} = \frac{\lambda}{n^{1/3}}$. Pertanto

$$P(|Y_n| > \epsilon) \leq \frac{\text{Var}(Y_n)}{\epsilon^2} = \frac{\lambda}{n^{1/3}\epsilon^2},$$

da cui si conclude facilmente.

Esercizio 20 Siano $\{X_1, X_2, Y_1, Y_2, S\}$ variabili aleatorie indipendenti con le seguenti distribuzioni:

$$X_1 \sim X_2 \sim \text{Po}(\lambda), \quad Y_1 \sim Y_2 \sim \text{Po}(\mu), \quad S \sim \text{Be}\left(\frac{1}{2}\right) = \text{B}\left(1, \frac{1}{2}\right),$$

dove $\lambda, \mu \in \mathbb{R}^+$ sono parametri fissati, con $\lambda \neq \mu$. Introduciamo le variabili Z_1, Z_2 definite da

$$Z_1 := X_1 \mathbf{1}_{\{S=0\}} + Y_1 \mathbf{1}_{\{S=1\}}, \quad Z_2 := X_2 \mathbf{1}_{\{S=0\}} + Y_2 \mathbf{1}_{\{S=1\}}.$$

In altre parole, lancio una moneta equilibrata (descritta dalla variabile S) e pongo $(Z_1, Z_2) := (X_1, X_2)$ se esce testa (cioè se $S = 0$) mentre pongo $(Z_1, Z_2) := (Y_1, Y_2)$ se esce croce ($S = 1$).

a) Calcolare $E(Z_1)$ e $E(Z_2)$.

b) Mostrare che Z_1 e Z_2 **non** sono indipendenti (sugg.: calcolare $E(Z_1 Z_2)$).

Soluzione.

a) Si noti che Z_1 e Z_2 hanno la stessa legge perché sono la stessa funzione dei vettori (X_1, Y_1, S) e (X_2, Y_2, S) , che hanno la stessa legge. In particolare $E(Z_1) = E(Z_2)$ e per l'indipendenza delle variabili in gioco si ha

$$E(Z_1) = E(X_1 \mathbf{1}_{\{S=0\}} + Y_1 \mathbf{1}_{\{S=1\}}) = E(X_1)P(S=0) + E(Y_1)P(S=1) = \frac{\lambda + \mu}{2}.$$

b) Si ha che

$$\begin{aligned} E(Z_1 Z_2) &= E(X_1 X_2 \mathbf{1}_{\{S=0\}} + Y_1 Y_2 \mathbf{1}_{\{S=1\}}) \\ &= E(X_1)E(X_2)P(S=0) + E(Y_1)E(Y_2)P(S=1) = \frac{\lambda^2 + \mu^2}{2} \neq E(Z_1)E(Z_2), \end{aligned}$$

per cui Z_1 e Z_2 non possono essere indipendenti.

Esercizio 21 Ad un bambino vengono regalati N palloncini gonfiabili. Ognuno di questi palloncini rimane “integro”, cioè non scoppia, un numero di giorni la cui distribuzione è geometrica di parametro $p \in (0, 1)$. Si considerino indipendenti le durate dei diversi palloncini. Per $n = 1, 2, \dots$ denotiamo con X_n il numero di palloncini ancora integri dopo n giorni.

a. Qual è la distribuzione di X_n ?

b. Sia T la variabile casuale a valori naturali tale che l'ultimo palloncino scoppia dopo il $T + 1$ -mo giorno. Calcolate $P(T > n)$ per ogni $n \geq 0$, e la densità discreta p_T di T .

c. Siano $1 \leq m \leq n$ e $0 \leq h \leq k \leq N$. Calcolare la probabilità condizionata

$$P(X_n = h | X_m = k).$$

d. Per ogni fissato $n \geq 1$, calcolare

$$P(X_n = N | X_{2n} = 0).$$

Soluzione.

a. Per $i = 1, 2, \dots, N$, sia ξ_i il numero di giorni in cui rimane integro il palloncino i -mo. Dunque $\xi_1, \dots, \xi_N \sim Ge(p)$ indipendenti. Posto $Y_{n,i} := 1_{\{\xi_i \geq n\}}$, abbiamo che $Y_{n,1}, \dots, Y_{n,N} \sim Be((1-p)^n)$ indipendenti. Inoltre

$$X_n = \sum_{i=1}^N Y_{n,i} \sim B(N, (1-p)^n).$$

b. Si ha $T = \max\{\xi_1, \dots, \xi_N\}$. Perciò

$$P(T \leq n) = P(\xi_1 \leq n, \xi_2 \leq n, \dots, \xi_N \leq n) = \prod_{i=1}^N P(\xi_i \leq n) = [1 - (1-p)^{n+1}]^N$$

da cui

$$P(T > n) = 1 - P(T \leq n) = 1 - [1 - (1-p)^{n+1}]^N.$$

Inoltre

$$p_T(n) = P(T = n) = P(T > n-1) - P(T > n) = [1 - (1-p)^{n+1}]^N - [1 - (1-p)^n]^N.$$

c. Cominciamo col calcolare $P(X_n = h, X_m = k)$.

$$\begin{aligned} P(X_n = h, X_m = k) &= \sum_{\substack{I \subseteq J \subseteq \{1, 2, \dots, N\} \\ |I|=h, |J|=k}} P(\xi_i \geq n \forall i \in I; m \leq \xi_i < n \forall i \in J \setminus I; \xi_i < m \forall i \notin J) \\ &= \sum_{\substack{I \subseteq J \subseteq \{1, 2, \dots, N\} \\ |I|=h, |J|=k}} [P(\xi_1 \geq n)]^h [P(m \leq \xi_1 < n)]^{k-h} [P(\xi_1 < m)]^{N-k} \\ &= \binom{N}{k} \binom{k}{h} [P(\xi_1 \geq n)]^h [P(m \leq \xi_1 < n)]^{k-h} [P(\xi_1 < m)]^{N-k} \end{aligned}$$

In modo del tutto simile:

$$P(X_m = k) = \binom{N}{k} [P(\xi_1 \geq m)]^k [P(\xi_1 < m)]^{N-k}.$$

Pertanto

$$\begin{aligned} P(X_n = h | X_m = k) &= \frac{P(X_n = h, X_m = k)}{P(X_m = k)} = \binom{k}{h} \left[\frac{P(m \leq \xi_1 < n)}{P(\xi_1 \geq m)} \right]^{k-h} \\ &= \binom{k}{h} \left[\frac{P(\xi_1 \geq m) - P(\xi_1 \geq n)}{P(\xi_1 \geq m)} \right]^{k-h} = \binom{k}{h} [1 - (1-p)^{n-m}]^{k-h}. \end{aligned}$$

d. Per la Formula di Bayes e per quanto visto al punto c.

$$P(X_n = N | X_{2n} = 0) = \frac{P(X_{2n} = 0 | X_n = N)P(X_n = N)}{P(X_{2n} = 0)} = \frac{[1 - (1-p)^n]^N (1-p)^{nN}}{(1 - (1-p)^{2n})^N}.$$

Esercizio 22 Siano $X_1, \dots, X_n, T$ variabili casuali scalari indipendenti. Le variabili $X_1, \dots, X_n$ hanno la stessa distribuzione, con $E(X_i) = \mu$ e $\text{Var}(X_i) = \sigma^2 < \infty$, mentre la variabile T prende valori in $\{1, \dots, n\}$.

a) Si spieghi perché, per $i, j \in \{1, \dots, n\}$, le variabili casuali X_i e $1_{\{T \geq j\}}$ sono indipendenti.

Si introduca la variabile S definita da

$$S := \sum_{i=1}^n X_i 1_{\{i \leq T\}}.$$

b) Si mostri che $E(S) = \mu E(T)$.

c) Si mostri che $\text{Cov}(S, X_j) = \sigma^2 P(T \geq j)$, per ogni $j \in \{1, \dots, n\}$.

Soluzione.

a) Dato che X_i e T sono indipendenti, per un risultato visto a lezione anche $f(X)$ e $g(Y)$ sono indipendenti, qualunque siano le funzioni (misurabili) f e g . Dato che $1_{\{T \geq j\}} = 1_{[j, \infty)}(T)$, la conclusione segue prendendo $f =$ funzione identica e $g = 1_{[j, \infty)}(\cdot)$.

b) Dato che $|S| \leq \sum_{i=1}^n |X_i|$ e le variabili X_i ammettono valor medio, segue che anche S ammette valor medio, e per linearità e indipendenza

$$E(S) = \sum_{i=1}^n E(X_i 1_{\{i \leq T\}}) = \sum_{i=1}^n E(X_i) P(T \geq i) = \mu \sum_{i=1}^n P(T \geq i) = \mu E(T).$$

c) Dato che $|S| \leq \sum_{i=1}^n |X_i|$ e le variabili X_i sono per ipotesi in L^2 , segue che anche $S \in L^2$, per cui $\text{Cov}(S, X_j)$ è ben definita. Per linearità e indipendenza si ha che

$$\begin{aligned} E(SX_j) &= \sum_{i=1}^n E(X_i X_j 1_{\{i \leq T\}}) = \sum_{i=1}^n E(X_i X_j) P(T \geq i) \\ &= \sum_{i \neq j} \mu^2 P(T \geq i) + E(X_j^2) P(T \geq j). \end{aligned}$$

D'altro canto, $E(S)E(X_j) = \mu^2 E(T) = \mu^2 \sum_{i=1}^n P(T \geq i)$, da cui

$$\text{Cov}(S, X_j) = E(SX_j) - E(S)E(X_j) = E(X_j^2)P(T \geq j) - \mu^2 P(T \geq j) = \sigma^2 P(T \geq j).$$

Esercizio 23 Sia $X \sim B(n, p)$. Mostrare che

$$E \left[\frac{1}{X+1} \right] = \frac{1 - (1-p)^{n+1}}{(n+1)p}.$$

Soluzione. Usando il fatto che

$$\frac{1}{k+1} \binom{n}{k} = \frac{1}{n+1} \binom{n+1}{k+1},$$

abbiamo:

$$\begin{aligned} E \left[\frac{1}{X+1} \right] &= \sum_{k=0}^n \frac{1}{k+1} \binom{n}{k} p^k (1-p)^{n-k} = \frac{1}{n+1} \sum_{k=0}^n \binom{n+1}{k+1} p^k (1-p)^{n-k} \\ &= \frac{1}{(n+1)p} \sum_{k=0}^n p^{k+1} (1-p)^{n-k} = \frac{1}{(n+1)p} \sum_{k=1}^{n+1} p^k (1-p)^{n+1-k} \\ &= \frac{1}{(n+1)p} \left[\sum_{k=0}^{n+1} p^k (1-p)^{n+1-k} - (1-p)^{n+1} \right] = \frac{1 - (1-p)^{n+1}}{(n+1)p}. \end{aligned}$$

Esercizio 24 Sia $X \sim Po(\lambda)$. Quale valore k massimizza $P(X = k)$?

Soluzione. Si noti che

$$\frac{P(X = k+1)}{P(X = k)} = \frac{\lambda}{k+1},$$

e perciò

$$p(X = k+1) \geq P(X = k) \iff \lambda \geq k+1.$$

Ne segue che $P(X = k)$ è massimizzata per $k = \lfloor \lambda \rfloor$.

Esercizio 25 Sia $X \sim Po(\lambda)$. Mostrare che

$$E[X^n] = \lambda E[(X+1)^{n-1}].$$

Si utilizzi questo risultato per calcolare $E(X^3)$.

Soluzione.

$$E[X^n] = e^{-\lambda} \sum_{k=0}^{+\infty} k^n \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{+\infty} k^{n-1} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=0}^{+\infty} (k+1)^{n-1} \frac{\lambda^k}{k!} = \lambda E[(X+1)^{n-1}].$$

In particolare, ricordando che $E(X) = \lambda$ e $E(X^2) = \lambda^2 + \lambda$,

$$E(X^3) = \lambda E[(X+1)^2] = \lambda[E(X^2) + 2E(X) + 1] = \lambda[\lambda^2 + 3\lambda + 1].$$

Esercizio 26 * Sia S un insieme di n elementi, $\Omega = \mathcal{P}(S) \setminus \{\emptyset\}$, e P la probabilità uniforme su Ω . Consideriamo la variabile casuale X definita da $X(\omega) = |\omega|$. Mostrare che

$$E(X) = \frac{n}{2 - \left(\frac{1}{2}\right)^{n-1}}$$

$$\text{Var}(X) = \frac{n2^{2n-2} - n(n+1)2^{n-2}}{(2^n - 1)^2}.$$

Si mostri inoltre che, per n grande, $\text{Var}(X) \sim n/4$.

Soluzione. Si noti che, per $k = 1, 2, \dots, n$

$$P(X = k) = \frac{\binom{n}{k}}{2^n - 1}.$$

Perciò

$$E(X) = \frac{1}{2^n - 1} \sum_{k=1}^n k \binom{n}{k} = \frac{\frac{n2^n}{2}}{2^n - 1} = \frac{n}{2 - (\frac{1}{2})^{n-1}},$$

dove si è usato il fatto che $\frac{1}{2^n} \sum_{k=1}^n k \binom{n}{k}$ è la media di una $B(n, 1/2)$. In modo analogo, $\frac{1}{2^n} \sum_{k=1}^n k^2 \binom{n}{k} = \frac{n(n+1)}{4}$, essendo il momento secondo di una $B(n, 1/2)$. Quindi

$$E(X^2) = \frac{1}{2^n - 1} \sum_{k=1}^n k^2 \binom{n}{k} = \frac{2^n \frac{n(n+1)}{4}}{2^n - 1},$$

da cui il valore della varianza si deduce facilmente. Per determinare il comportamento asintotico, è sufficiente osservare che, nella frazione che esprime la varianza, il termine dominante a numeratore è $n2^{2n-2}$, mentre il termine dominante a denominatore è 2^{2n} .

Esercizio 27 Siano $X_1, X_2, \dots, X_n$ variabili casuali i.i.d., la cui funzione generatrice dei momenti $\gamma(t)$ è finita per ogni $t \in R$, e sia N una variabile casuale a valori in $\{1, 2, \dots, n\}$, indipendente da $X_1, X_2, \dots, X_n$, la cui densità discreta indichiamo con $p_N(\cdot)$. Definiamo la variabile casuale S come segue

$$S(\omega) = \sum_{k=1}^{N(\omega)} X_k(\omega).$$

- Calcolare, in termini di $\gamma(t)$ e $\{p_N(j)\}_{1 \leq j \leq n}$, la funzione generatrice $\gamma_S(t)$ di S . (Sugg: osservare che $e^{tS} = \sum_{j=1}^n e^{t(X_1 + \dots + X_j)} \mathbf{1}_{N=j}$.)
- Calcolare media e varianza di S in termini di media e varianza di X_1 e di N .

Soluzione.

- Per l'indipendenza di $S, X_1, \dots, X_n$ si ha

$$\begin{aligned} \gamma_S(t) &= E(e^{tS}) = \sum_{j=1}^n E(e^{t(X_1 + \dots + X_j)} \mathbf{1}_{N=j}) = \sum_{j=1}^n E(e^{t(X_1 + \dots + X_j)}) P(N = j) \\ &= \sum_{j=1}^n \gamma(t)^j p_N(j). \end{aligned}$$

b) Dall'espressione ottenuta in a) è chiaro che $\gamma_S(t)$ è finita per ogni $t \in \mathbb{R}$, per cui $E(S) = \gamma'_S(0)$ e $E(S^2) = \gamma''_S(0)$. Calcoliamo dunque le derivate di $\gamma_S(t)$:

$$\begin{aligned}\gamma'_S(t) &= \sum_{j=1}^n j \gamma(t)^{j-1} \gamma'(t) p_N(j), \\ \gamma''_S(t) &= \sum_{j=1}^n j(j-1) \gamma(t)^{j-2} \gamma'(t)^2 p_N(j) + \sum_{j=1}^n j \gamma(t)^{j-1} \gamma''(t) p_N(j).\end{aligned}$$

Ponendo $t = 0$ si ottiene

$$\begin{aligned}E(S) &= \gamma'_S(0) = \gamma'(0) \sum_{j=1}^n j p_N(j) = E(X_1) E(N), \\ E(S^2) &= \gamma''_S(0) = \gamma'(0)^2 \sum_{j=1}^n j(j-1) p_N(j) + \gamma''(0) \sum_{j=1}^n j p_N(j) \\ &= E(X_1)^2 (E(N^2) - E(N)) + E(X_1^2) E(N),\end{aligned}$$

da cui

$$\text{Var}(S) = E(S^2) - E(S)^2 = E(N) \text{Var}(X_1) + E(X_1)^2 \text{Var}(N).$$

Esercizio 28 Sia $(X_n)_{n \geq 1}$ una successione di variabili casuali indipendenti, tali che $X_n \sim Be(1/2^n)$. Definiamo la variabile casuale T , a valori in $\mathbb{N} \cup \{+\infty\} = \{1, 2, \dots, n, \dots, +\infty\}$:

$$T = \begin{cases} \min\{k \in \mathbb{N} \text{ tali che } X_k = 1\} & \text{se } \{k \in \mathbb{N} \text{ tali che } X_k = 1\} \neq \emptyset \\ +\infty & \text{altrimenti.} \end{cases}$$

- Determinare un'espressione per $P(T > n)$. [Sugg.: si noti che l'evento $\{T > n\}$ può essere espresso in modo semplice in funzione delle prime n variabili $X_1, \dots, X_n$.]
- Scrivere $\log P(T > n)$ in forma di una somma $\sum_{k=1}^n a_k$, e mostrare che la serie $\sum_k a_k$ converge.
- Dedurre che $P(T = +\infty) > 0$.

Soluzione.

a) Dalla definizione di T e dall'indipendenza delle X_i si ha che

$$\begin{aligned}P(T > n) &= P(X_1 = 0, X_2 = 0, \dots, X_n = 0) \\ &= P(X_1 = 0) \cdot P(X_2 = 0) \cdots P(X_n = 0) = \prod_{i=1}^n \left(1 - \frac{1}{2^i}\right).\end{aligned}$$

b) Dal risultato del punto a) possiamo scrivere

$$\log P(T > n) = \sum_{i=1}^n \log \left(1 - \frac{1}{2^i}\right).$$

Si noti che $-cx \leq \log(1-x) \leq 0$ per ogni $x \in [\frac{1}{2}, 1]$, per un'opportuna costante positiva c (si può prendere per es. $c = 2 \log 2$). Si ha dunque, per ogni $n \in \mathbb{N}$,

$$\left| \sum_{i=1}^n \log \left(1 - \frac{1}{2^i} \right) \right| \leq \sum_{i=1}^n \left| \log \left(1 - \frac{1}{2^i} \right) \right| \leq c \sum_{i=1}^n \frac{1}{2^i} \leq c,$$

il che mostra che la serie $\sum_i \log \left(1 - \frac{1}{2^i} \right)$ converge assolutamente.

c) Si noti che $\{T = +\infty\} = \bigcap_{n=1}^{\infty} \{T > n\}$ e che $\{T > n\} \supseteq \{T > n+1\}$. La continuità della probabilità dà dunque che

$$P(T = +\infty) = \lim_{n \rightarrow \infty} P(T > n) = e^{\lim_{n \rightarrow \infty} \log P(T > n)} > 0,$$

dove si è usato il fatto che il limite di $\log P(T > n)$ esiste finito, come dimostrato nel punto b).

Esercizio 29 Siano $X \sim Be(p)$, $Y \sim Po(\lambda)$ e $Z \sim Po(\mu)$ variabili casuali indipendenti. Definiamo

$$W := XY + (1 - X)Z.$$

- Senza fare calcoli con densità, calcolare media e varianza di W , e la covarianza $Cov(W, Y)$.
- Determinare la densità congiunta di (W, Y) , e la densità marginale di W .

Soluzione.

a)

$$E(W) = E(XY) + E((1 - X)Z) = E(X)E(Y) + (1 - E(X))E(Z) = p\lambda + (1 - p)\mu.$$

Osservando che $X(1 - X) \equiv 0$ e che $X^2 = X$, $(1 - X)^2 = 1 - X$, si ha

$$E(W^2) = E(X)E(Y^2) + (1 - E(X))E(Z^2) = p(\lambda + \lambda^2) + (1 - p)(\mu + \mu^2),$$

da cui si ottiene

$$Var(W) = p\lambda + (1 - p)\mu + p(1 - p)(\lambda - \mu)^2.$$

Infine

$$E(WY) = E(XY^2) + E((1 - X)ZY) = p(\lambda + \lambda^2) + (1 - p)\lambda\mu,$$

da cui

$$Cov(W, Y) = E(WY) - E(W)E(Y) = p\lambda.$$

b)

$$\begin{aligned} P(W = k, Y = h) &= P(W = k, Y = h, X = 0) + P(W = k, Y = h, X = 1) \\ &= (1 - p)P(Z = k, Y = h) + pP(Y = k, Y = h) \\ &= (1 - p)e^{-\lambda - \mu} \frac{\lambda^k}{k!} \frac{\mu^h}{h!} + p\mathbf{1}_{\{0\}}(k - h)e^{-\lambda} \frac{\lambda^k}{k!}. \end{aligned}$$

Da cui si trova

$$\begin{aligned} P(W = k) &= \sum_{h \geq 0} P(W = k, Y = h) = \sum_{h \neq k} P(W = k, Y = h) + P(W = k, Y = k) \\ &= \sum_{h \geq 0} (1-p)e^{-\lambda-\mu} \frac{\lambda^k \mu^h}{k! h!} + pe^{-\lambda} \frac{\lambda^k}{k!} = (1-p)e^{-\mu} \frac{\mu^k}{k!} + pe^{-\lambda} \frac{\lambda^k}{k!}. \end{aligned}$$

Esercizio 30 Sia X una variabile aleatoria a valori in $\mathbb{N}$, la cui funzione di ripartizione F_X è data, per $n \in \mathbb{N}$, dalla formula

$$F_X(n) := 1 - \frac{1}{n+1}.$$

(a) Si determini la densità p_X di X , e si mostri che $X \notin L^1$.

Per $\alpha > 0$, diciamo che una variabile Y ha distribuzione $S(\alpha)$ ($Y \sim S(\alpha)$) se per ogni $x \in \mathbb{R}$

$$F_Y(x) := 1 - (1 - F_X(x))^\alpha.$$

(b) Siano $Y \sim S(\alpha)$ e $Z \sim S(\beta)$ variabili aleatorie indipendenti. Si determini la distribuzione di $W := \min(Y, Z)$.

(c) Determinare i valori di α tali che se $Y \sim S(\alpha)$ allora $Y \in L^1$. (Può essere utile ricordare che $(1+x)^\alpha = 1 + \alpha x + o(x)$ per $x \rightarrow 0$)

Soluzione.

(a)

$$p_X(n) = F_X(n) - F_X(n-1) = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}.$$

Inoltre

$$E(X) = \sum_n np_X(n) = \sum_n \frac{1}{n+1} = +\infty.$$

Quindi $X \notin L^1$

(b) Notare che se $Y \sim S(\alpha)$

$$1 - F_Y(n) = (1 - F_X(n))^\alpha = \frac{1}{(n+1)^\alpha}.$$

Pertanto

$$\begin{aligned} 1 - F_W(n) &= P(W > n) = P(Y > n)P(Z > n) = (1 - F_Y(n))(1 - F_Z(n)) \\ &= (1 - F_X(n))^\alpha (1 - F_X(n))^\beta = (1 - F_X(n))^{\alpha+\beta}, \end{aligned}$$

da cui segue che $W \sim S(\alpha + \beta)$.

(c) Anzitutto

$$\begin{aligned} p_Y(n) = F_Y(n) - F_Y(n-1) &= \frac{1}{n^\alpha} - \frac{1}{(n+1)^\alpha} = \frac{1}{n^\alpha} \left[1 - \frac{1}{\left(1 + \frac{1}{n}\right)^\alpha} \right] \\ &= \frac{1}{n^\alpha} \left[1 - \frac{1}{1 + \frac{\alpha}{n} + o(1/n)} \right] = \frac{1}{n^\alpha} \frac{\frac{\alpha}{n} + o(1/n)}{1 + \frac{\alpha}{n} + o(1/n)} \end{aligned}$$

da cui segue che $p_Y(n) \sim \frac{\alpha}{n^{\alpha+1}}$, nel senso che

$$\lim_{n \rightarrow +\infty} \frac{p_Y(n)}{\frac{\alpha}{n^{\alpha+1}}} = 1.$$

Perciò

$$E(Y) = \sum_n n p_Y(n) < +\infty$$

se e solo se

$$\sum_n \frac{1}{n^\alpha} < +\infty$$

cioè per $\alpha > 1$.

Un modo alternativo, e assai più rapido, di ottenere il risultato, consiste nel ricordare che, pr variabili aleatorie a valori in $\mathbb{N}$, si ha

$$E(Y) = \sum_{n \geq 0} P(Y > n) = \sum_{n \geq 0} [1 - F_Y(n)].$$

In questo caso

$$E(Y) = \sum_{n \geq 0} \frac{1}{(n+1)^\alpha}$$

che è finito se e solo se $\alpha > 1$.

Esercizio 31 Sia $(X_i)_{i \geq 1}$ una successione di variabili aleatorie i.i.d. con distribuzione di Bernoulli di parametro $p \in (0, 1)$. Sia $q \in (p, 1)$.

(a) Utilizzando la disuguaglianza di Chebychev, si dimostri che per ogni $N \in \mathbb{N}$, ogni $t \geq 0$,

$$\mathbf{P} \left(\frac{1}{N} \sum_{i=1}^N X_i \geq q \right) \leq e^{-tNq} (\mathbf{E} [e^{tX_1}])^N = \psi_N(t),$$

ove $\psi_N(t) \doteq e^{-tNq} (e^t p + (1-p))^N$.

(b) Poniamo $h(q, p) \doteq q \log(q/p) + (1-q) \log((1-q)/(1-p))$. Utilizzando la differenziabilità di $\psi_N(\cdot)$, si calcoli il parametro $t = t_* \in [0, \infty)$ per il quale $\psi_N(t)$ sia minima. Si deduca che

$$\mathbf{P} \left(\frac{1}{N} \sum_{i=1}^N X_i \geq q \right) \leq e^{-Nh(q,p)}.$$

Soluzione. Abbiamo per ogni $N \in \mathbb{N}$, ogni $t \geq 0$,

$$\begin{aligned} \mathbf{P}\left(\frac{1}{N}\sum_{i=1}^N X_i \geq q\right) &= \mathbf{P}\left(t\sum_{i=1}^N X_i \geq tNq\right) \\ &= \mathbf{P}\left(e^{t\sum_{i=1}^N X_i} \geq e^{tNq}\right) \\ &\leq e^{-tNq} \mathbf{E}\left[\prod_{i=1}^N e^{tX_i}\right] \\ &= e^{-tNq} (\mathbf{E}[e^{tX_1}])^N. \end{aligned}$$

Siccome $X_1 \sim Be(p)$, si ha $\mathbf{E}[e^{tX_1}] = e^tp + (1-p)$.

Visto che $\psi_N(\cdot)$ possiede derivate continue e $t = 0$ non è la posizione di un minimo, otteniamo t_* come soluzione di $\frac{d}{dt}\psi_N(t) = 0$. Infatti

$$\frac{d}{dt}\psi_N(t) = -Nqe^{-tNq} (e^tp + (1-p))^N + Ne^tpe^{-tNq} (e^tp + (1-p))^{N-1},$$

da cui $t_* = \log(q/p) - \log((1-q)/(1-p))$. Infine si vede che $\psi_N(t_*) = Nh(q, p)$.

Esercizio 32 In quanto segue, per $A \subseteq \mathbb{R}^2$ e $y \in \mathbb{R}$, sia $A_y := \{x \in \mathbb{R} : (x, y) \in A\}$.

(a) Siano X e Y due variabili aleatorie a valori in $\mathbb{R}$, discrete e indipendenti. Si mostri che, per ogni $A \subseteq \mathbb{R}^2$ e $y \in \mathbb{R}$ tale che $P(Y = y) > 0$, si ha

$$(2) \quad P((X, Y) \in A | \{Y = y\}) = P(X \in A_y).$$

(b) Siano $W, Z \sim Po(1)$ indipendenti. Si ponga $Y := Z$, $X := W + Z$, $A := \{(0, 0)\}$, $y = 0$. Mostrare che l'identità (2) non vale, e si spieghi perché ciò non contraddice quanto visto al punto precedente.

Soluzione.

(a)

$$\begin{aligned} P((X, Y) \in A | \{Y = y\}) &= \frac{P(\{(X, Y) \in A\} \cap \{Y = y\})}{P(Y = y)} = \frac{P(\{(X, y) \in A\} \cap \{Y = y\})}{P(Y = y)} \\ &= \frac{P(\{X \in A_y\} \cap \{Y = y\})}{P(Y = y)} = P(X \in A_y). \end{aligned}$$

(b)

$$P((X, Y) \in \{(0, 0)\} | \{Y = 0\}) = \frac{P(W = 0, Z = 0)}{P(Z = 0)} = e^{-1},$$

mentre

$$P(X \in A_0) = P(X = 0) = e^{-2}.$$

Non vi è contraddizione con il punto precedente, in quanto X e Y non sono indipendenti.

Esercizio 33 Siano X e Y variabili aleatorie discrete, definite nello stesso spazio di probabilità. La funzione

$$\varphi_{X,Y}(x, z) := \begin{cases} P(X = x|X + Y = z) & \text{se } P(X + Y = z) \neq 0 \\ 0 & \text{altrimenti} \end{cases}$$

viene detta *distribuzione di X condizionata a $X + Y$* .

- Determinare $\varphi(x, y)$ nel caso in cui $X \sim B(n, p)$ e $Y \sim B(m, p)$ sono indipendenti. In particolare, mostrare che $\varphi(x, y)$ non dipende da p
- Determinare $\varphi(x, y)$ nel caso in cui $X \sim Po(\lambda)$ e $Y \sim Po(\mu)$ sono indipendenti. In particolare, mostrare che $\varphi(x, y)$ dipende da λ e μ solo attraverso la quantità $\frac{\lambda}{\lambda + \mu}$.
- Siano X e Y due variabili aleatorie discrete indipendenti, con densità rispettivamente p_X e p_Y , e sia $f : \mathbb{R} \rightarrow (0, +\infty)$ tale che

$$\sum_{x,y} f(x+y)p_X(x)p_Y(y) =: c \in (0, +\infty).$$

Siano inoltre X', Y' due variabili aleatorie discrete la cui densità congiunta sia data da

$$p_{X',Y'}(x, y) := \frac{1}{c} f(x+y)p_X(x)p_Y(y).$$

Mostrare che $\varphi_{X',Y'} = \varphi_{X,Y}$.

Soluzione.

- Si noti che $X + Y \sim B(n + m, p)$. Pertanto, se $x = 0, 1, \dots, n$, $z = 0, 1, \dots, n + m$, e $x \leq z$,

$$\varphi_{X,Y}(x, z) = \frac{P(X = x)P(Y = z - x)}{P(X + Y = z)} = \frac{\binom{n}{x} p^x (1-p)^{n-x} \binom{m}{z-x} p^{z-x} (1-p)^{m-z+x}}{\binom{n+m}{z} p^z (1-p)^{n+m-z}} = \frac{\binom{n}{x} \binom{m}{z-x}}{\binom{n+m}{z}}$$

- Analogamente, essendo $X + Y \sim Po(\lambda + \mu)$, se $0 \leq x \leq z$,

$$\varphi_{X,Y}(x, z) = \frac{P(X = x)P(Y = z - x)}{P(X + Y = z)} = \frac{e^{-\lambda} \frac{\lambda^x}{x!} e^{-\mu} \frac{\mu^{z-x}}{(z-x)!}}{e^{-\lambda-\mu} \frac{(\lambda+\mu)^z}{z!}} = \binom{z}{x} \left(\frac{\lambda}{\lambda + \mu} \right)^x \left(\frac{\mu}{\lambda + \mu} \right)^{z-x}$$

- Anzitutto, notiamo che

$$P(X' + Y' = z) = \sum_x P(X' = x, Y' = z - x) = f(z) \sum_x P(X = x, Y = z - x) = f(z) P(X + Y = z).$$

Quindi $P(X' + Y' = z) > 0$ se e solo se $P(X + Y = z) > 0$. In tal caso

$$\varphi_{X',Y'}(x, z) = \frac{P(X' = x, Y' = z - x)}{P(X' + Y' = z)} = \frac{f(z) p_X(x) p_Y(z - x)}{f(z) P(X + Y = z)} = \varphi_{X,Y}(x, z).$$