Esercizi - Fascicolo III

Esercizio 1 In una procedura di controllo di produzione, n processori prodotti da un processo industriale vengono sottoposti a controllo. Si assuma che ogni pezzo, indipendentemente dagli altri, abbia probabilità $p \in (0,1)$ di essere difettoso. Se un processore è funzionante supera sicuramente il test di controllo, se il processore è difettoso fallisce il test con probabilità $q \in (0,1)$, indipendentemente dagli altri. Sia X = numero di processori che hanno fallito il test. Determinare la distribuzione di X.
Soluzione. Basta osservare che la probabilità che un pezzo fallisca il test è pq, e usare lo schema delle prove ripetute indipendenti, per ottenere

$$p_X(k) = \binom{n}{k}(pq)^k(1-pq)^{n-k}.$$

Esercizio 2 Due dadi truccati sono tali che la probabilità di ottenere un sei è il doppio della probabilità di ottenere ogni altro punteggio. Qual è la media del punteggio ottenuto lanciando i due dadi?
Soluzione. Sia X_i il punteggio del dado i-mo, sicché il punteggio totale X è dato da $X = X_1 + X_2$. X_i assume i valori $1,2,3,4,5$ con probabilità $1/7$, e 6 con probabilità $2/7$, e perciò $E(X_i) = 27/7$, da cui $E(X) = 54/7$.

Esercizio 3 Da un'urna contenente r palline rosse e v palline verdi, si estraggono successivamente, senza reintroduzione, k palline, con $k \leq \min(r,v)$. Per $i = 1,2,\dots,k$, sia

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-ma pallina estratta è rossa} \\ 0 & \text{altrimenti,} \end{cases}$$

e $X = X_1 + \cdots + X_k$.

a. Determinare la distribuzione di X.
b. Determinare le distribuzioni delle X_i.
c. Calcolare $E(X)$.
d.* Mostrare, per induzione su k, che la densità congiunta delle X_i è data da

$$p_{X_1,\dots,X_k}(x_1,\dots,x_k) = \frac{r(r-1)\cdots(r-\sum_{i=1}^k x_i + 1)v(v-1)\cdots(v-k+\sum_{i=1}^k x_i + 1)}{(r+v)(r+v-1)\cdots(r+v-k+1)}$$

Soluzione. a. X è il numero di palline rosse estratte in k estrazioni senza reintroduzione, e quindi

$$p_X(n) = \binom{r}{n}\binom{v}{k-n}/\binom{r+v}{n}.$$

b. Identifichiamo l'insieme delle palline con $\{1,2,\dots,r+v\}$, convenendo che le palline rosse siano le prime r. Lo schema di estrazioni senza reitroduzione si può modellare scegliendo, come spazio campionario, Ω = insieme delle permutazioni di $\{1,2,\dots,r+v\}$, con la probabilità uniforme. Allora $\{X_i = 1\} = \{\sigma \in \Omega : \sigma(i) \in \{1, 2, \dots, r\}\}$, che ha cardinalità $r(r + v - 1)!$, e quindi probabilità $r/r + v$. In conclusione

$$p_{X_i}(1) = 1 - p_{X_i}(0) = \frac{r}{r+v}.$$

c. Essendo $E(X_i) = r/r + v$, si ha $E(X) = kr/(r + v)$.

d. Per induzione su k. Per $k = 1$ la risposta è già in b. Altrimenti, se $x = (x_1, \dots, x_{k-1}, x_k) \in \{0, 1\}^k$,

$$P(X = x) = P(X_k = x_k | X_{k-1} = x_{k-1}, \dots, X_1 = x_1) P(X_{k-1} = x_{k-1}, \dots, X_1 = x_1).$$

Per $P(X_{k-1} = x_{k-1}, \dots, X_1 = x_1)$ si usa l'ipotesi induttiva, mentre

$$P(X_k = 1 | X_{k-1} = x_{k-1}, \dots, X_1 = x_1) = \frac{r - \sum_{i=1}^{k-1} x_i}{r + v - k + 1},$$

da cui si giunge alla conclusione.

Esercizio 4 Un'urna contiene $n \geq 1$ palline bianche e 2 palline rosse. Si eseguono estrazioni ripetute *senza reimmissione*. Introduciamo la variabile casuale

$$X = \text{numero di palline bianche estratte prima di estrarre una pallina rossa},$$

la cui densità discreta verrà indicata con $p_X(k) = P(X = k)$.

a) Mostrare che, per $k = 0, 1, \dots, n$,

$$p_X(k) = \frac{2}{(n+2)(n+1)}(n - k + 1).$$

b) Calcolare $E(X)$ (ricordare le formule $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ e $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$, valide per ogni $n \geq 1$).

Soluzione.

a) Consideriamo gli eventi $A =$ "la $k+1$-ma pallina estratta è rossa", $B =$ "le prime k palline estratte sono tutte bianche". Si ha

$$P(X = k) = P(A \cap B) = P(A|B)P(B) = \frac{2}{n - k + 2} \frac{\binom{n}{k}}{\binom{n+2}{k}} = \frac{2}{(n+2)(n+1)}(n - k + 1),$$

dove nell'ultimo passaggio si sono eseguite le dovute semplificazioni.

b)

$$E(X) = \frac{2}{(n+2)(n+1)} \sum_{k=0}^{n} k(n - k + 1) = \frac{2}{n+2} \sum_{k=1}^{n} k - \frac{2}{(n+2)(n+1)} \sum_{k=1}^{n} k^2$$
$$= \frac{n(n+1)}{n+2} - \frac{n(2n+1)}{3(n+2)} = \frac{n}{3}.$$

Esercizio 5 Per $n \geq 1$, sia X_n una variabile casuale che assume, con la stessa probabilità, i valori $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1$. Se f è una funzione continua, sia

$$m_n = E(f(X_n)).$$

Mostrare che

$$\lim_{n \to +\infty} m_n = \int_0^1 f(x)dx.$$

Soluzione. Si osserva che

$$E(f(X_n)) = \frac{1}{n} \sum_{k=1}^{n} f(k/n),$$

che è una somma di Riemann per l'integrale $\int_0^1 f(x)dx$.

Esercizio 6 Siano X_1, X_2 variabili uniformi discrete sull'insieme $\{1, \ldots, n\}$ (cioè con densità costante su tale insieme), dove $n \in \mathbb{N}$, tra loro indipendenti. Definiamo la variabile $Y := \min\{X_1, X_2\}$.

a) Si calcoli $P(Y = k)$ per ogni $k \in \mathbb{N}$.

b) Si mostri che, per ogni $t \in (0, 1)$, si ha che

$$\lim_{n \to \infty} P(Y \leq tn) \;=\; 2t - t^2 \,.$$

Soluzione.

a) Chiaramente $P(Y = k) = 0$ per $k > n$. Conviene innanzitutto calcolare, per $k \in \{1, \ldots, n\}$, la probabilità $P(Y \geq k)$ che è data da

$$P(Y \geq k) \;=\; P(X_1 \geq k,\, X_2 \geq k) \;=\; P(X_1 \geq k)\, P(X_2 \geq k) \;=\; \left(\frac{n-k+1}{n}\right)^2 \;=\; \left(1 - \frac{k-1}{n}\right)^2 .$$

Si ha dunque

$$P(Y = k) \;=\; P(Y \geq k) - P(Y \geq k+1) \;=\; \left(\frac{n-k+1}{n}\right)^2 - \left(\frac{n-k}{n}\right)^2 \;=\; \frac{1}{n^2} + \frac{2}{n}\left(1 - \frac{k}{n}\right) .$$

b) Dalla formula per $P(Y \geq k)$ calcolata al punto a) si ottiene

$$P(Y \leq tn) \;=\; 1 - P(Y > tn) \;=\; 1 - P(Y > \lfloor tn \rfloor) \;=\; 1 - \left(1 - \frac{\lfloor tn \rfloor - 1}{n}\right)^2 ,$$

e dato che $(\lfloor tn \rfloor - 1)/n \to t$ per $n \to \infty$, per ogni $t \in (0, 1)$, si ha che

$$\lim_{n \to \infty} P(Y \leq tn) \;=\; 1 - (1 - t)^2 \;=\; 2t - t^2 \,.$$

Esercizio 7 Siano X, Y variabili casuali a valori in $\mathbb{N}$, definite sullo stesso spazio di probabilità. Giustificare l'identità

$$P(X+Y=n)=\sum_{k=0}^{+\infty} P(X=k, Y=n-k).$$

Soluzione. Basta osservare che

$$\{X+Y=n\}=\bigcup_{k=0}^{+\infty}\{X=k, Y=n-k\},$$

e usare la σ-addidività.

Esercizio 8 * Sia X una variabile casuale a valori in $\mathbb{N}$. Allora

$$E(X)=\sum_{n=1}^{+\infty} P(X \geq n).$$

Soluzione.

$$\begin{aligned}\sum_{n=1}^{+\infty} P(X \geq n) &= \sum_{n=1}^{+\infty}\sum_{k=n}^{+\infty} P(X=k)=\sum_{k=1}^{+\infty}\sum_{n=1}^{k} P(X=k)\\ &=\sum_{k=1}^{+\infty} kP(X=k)=E(X).\end{aligned}$$

Esercizio 9 Si consideri la seguente classica strategia per il gioco della roulette. Gioco sempre sul rosso. Alla prima giocata punto un dollaro. Se perdo raddoppio la giocata, se vinco smetto. In ogni caso, dato che il mio capitale iniziale è 1023 dollari, se perdo 10 volte di seguito devo smettere. Sia X la differenza tra il mio capitale alla fine del gioco e il capitale iniziale. Calcolare $E(X)$.
Soluzione. Notare che $X=-1023$ se perdo 10 volte di seguito, il che avviene con probabilità $\left(\frac{19}{37}\right)^{10}$. Altrimenti, se vinco in uno dei primi 10 tentativi, un facile calcolo mostra che $X=1$. Il calcolo della media è allora immediato:

$$E(X)=-1023\left(\frac{19}{37}\right)^{10}+\left[1-\left(\frac{19}{37}\right)^{10}\right] \simeq -0.3.$$

Esercizio 10 Un gioco a premi ha un montepremi di 512\$. Vengono poste ad un concorrente 10 domande. Ad ogni risposta errata il montepremi viene dimezzato. Alla prima risposta esatta il concorrente vince il montepremi rimasto. Se non si da alcuna risposta esatta non si vince nulla. Un certo concorrente risponde esattamente ad una domanda con probabilità $p \in (0,1)$, indipendentemente dalle risposte alle altre domande. Sia X la vincita di questo concorrente. Determinare la densità p_X di X.

Soluzione. La variabile casuale X assume i valori $0,1,2,2^2,2^3,\ldots,2^9$. Per $0 \leq k \leq 9$, il valore 2^{9-k} viene assunto se le prime k risposte sono errate, e la $k+1$-ma è corretta. Ciò avviene con probabilità $(1-p)^k p$. Infine il valore 0 viene assunto se tutte le 10 risposte sono sbagliate, il che avviene con probabilità $(1-p)^{10}$. Riassumendo

$$p_X(2^{9-k}) = p(1-p)^k$$

per $0 \leq k \leq 9$, e $p_X(0) = (1-p)^{10}$.

Esercizio 11 In un concorso vengono assegnate le idoneità per un dato servizio. Si assuma che ogni partecipante, indipendentemente dagli altri, abbia probabilità $p = \frac{3}{4}$ di ottenere l'idoneità. Al termine del concorso, a 10 tra gli idonei viene assegnato un posto di lavoro (se gli idonei sono meno di 10 vengono assegnati tanti posti di lavoro quanti sono gli idonei). Supponiamo che al concorso partecipino 15 persone, e sia X il numero dei partecipanti che ottengono l'idoneità ma non il posto di lavoro.

a. Determinare la distribuzione di X.

b. Calcolare $E(X)$.

Soluzione. a. Notare che, se $X > 0$, allora X è il numero di idonei meno 10. Allora, se Y è il numero di idonei:

$$P(X=1) = P(Y=11) = \binom{15}{11}\left(\frac{3}{4}\right)^{11}\left(\frac{1}{4}\right)^{4}.$$

Analogamente:

$$P(X=2) = \binom{15}{12}\left(\frac{3}{4}\right)^{12}\left(\frac{1}{4}\right)^{3}$$

$$P(X=3) = \binom{15}{13}\left(\frac{3}{4}\right)^{13}\left(\frac{1}{4}\right)^{2}$$

$$P(X=4) = \binom{15}{14}\left(\frac{3}{4}\right)^{14}\left(\frac{1}{4}\right)$$

$$P(X=5) = \left(\frac{3}{4}\right)^{15}$$

e, infine, $P(X=0) = 1 - \sum_{i=1}^{5} P(X=i)$.

b.

$$E(X) = \sum_{i=0}^{5} iP(X=i) \simeq 1.4774.$$

Esercizio 12 Si considerino 5 urne identiche, ognuna contenente una pallina rossa e quattro palline verdi. Ogni urna viene assegnata ad uno di cinque giocatori, e ogni giocatore estrae una pallina dalla propria urna. Un montepremi di 3000 Euro viene diviso tra i giocatori che estraggono la pallina rossa.

a. Sia X il numero di Euro vinti da ogni giocatore vincente ($X = 0$ se nessun giocatore estrae la pallina rossa). Determinare la densità e la media di X.

b. Si supponga di considerare uno dei cinque giocatori, chiamiamolo Tizio, e sia Y il numero di Euro vinti da Tizio. Si determinino la densità e la media di Y.

Soluzione. a. Sia $Z =$ numero di giocatori vincenti. Chiaramente $Z \sim B(5, 0.2)$. Abbiamo:

$$\begin{aligned}
P(X=0) &= P(Z=0) = (0.8)^5 \\
P(X=3000) &= P(Z=1) = 5(0.8)^4(0.2) \\
P(X=1500) &= P(Z=2) = ... \\
P(X=1000) &= P(Z=3) = ... \\
P(X=750) &= P(Z=4) = ... \\
P(X=600) &= P(Z=5) = ...
\end{aligned}$$

da cui si calcola

$$E(X) = 3000P(X=3000) + 1500P(X=1500) + 1000P(X=1000) +$$
$$750P(X=750) + 600P(X=600) = ...$$

b. Si noti che $Y = 0$ se Tizio non estrae la pallina rossa, quindi $P(Y=0) = 4/5$. Inoltre, per $k = 0, 1, 2, 3, 4$, si ha che $Y = \frac{3000}{k+1}$ se Tizio ha estratto la pallina rossa, e altri k giocatori hanno estratto la pallina rossa. Si ha, perciò

$$P\left(Y = \frac{3000}{k+1}\right) = \frac{1}{5}\binom{4}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{4-k}.$$

Allora

$$E(Y) = \sum_{k=0}^{4} \frac{3000}{k+1}\frac{1}{5}\binom{4}{k}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{4-k} = ...$$

Esercizio 13 Si sceglie "a caso" un campione di 5 oggetti da un lotto di 100 di cui 10 sono difettosi per effettuare un controllo di qualità. Sia X il numero di oggetti difettosi contenuti nel campione. Determinare la densità discreta di X.

Soluzione. La variabile X assume solo i valori $0, 1 \ldots, 5$. $X = k$ significa che nel lotto di 5 oggetti k sono difettosi. Possiamo calcolare $\Pr(X = k)$ riconoscendo che X ha una distribuzione ipergeometrica oppure calcolare questa probabilità come casi favorevoli su casi possibili. *Casi possibili*: ci sono $\binom{100}{5}$ modi di scegliere 5 oggetti tra 100. *Casi favorevoli*: devo scegliere 5 oggetti di cui k difettosi. Scelgo prima i k difettosi tra i 10 difettosi in $\binom{10}{k}$ modi, poi scelgo i rimanenti $5-k$ oggetti tra i rimanenti $100-10 = 90$ non difettosi in $\binom{90}{5-k}$ modi. In definitiva i casi favorevoli sono: $\binom{10}{k}\binom{90}{5-k}$ e quindi:

$$\Pr(X = k) = \frac{\binom{10}{k}\binom{90}{5-k}}{\binom{100}{5}}, \qquad k = 0, 1, \ldots, 5$$

Facendo un po' di conti si ottiene: $\Pr(X = 0) \simeq 0.583$, $\Pr(X = 1) \simeq 0.340$, $\Pr(X = 2) \simeq 0.070$, $\Pr(X = 3) \simeq 0.007$, $\Pr(X = 4) \simeq \Pr(X = 5) \simeq 0$. Per essere sicuri di non aver sbagliato i conti su può fare la verifica $\sum_{k=0}^{5} \Pr(X = k) = 1$.

Esercizio 14 Sia $N \geq 2$ e sia Ω l'insieme dei sottoinsiemi non vuoti di $\{1, 2, \ldots, N\}$. In altre parole

$$\Omega := \{\omega \subseteq \{1, 2, \ldots, N\} : \omega \neq \emptyset\}.$$

Se $\omega \in \Omega$ sia $X(\omega) := \max(\omega)$ il massimo elemento di ω e $Y(\omega) := \min(\omega)$ il minimo elemento di ω. Infine, sia P la probabilità uniforme su Ω.

i) Mostrare che, per $n \in \{1, 2, \ldots, N\}$,

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

ii) Calcolare la funzione generatrice dei momenti di X.

iii) Determinare la densità congiunta di (X, Y).

iv)* Determinare la densità di $X - Y$.

Soluzione.

i) Anzitutto, è noto che $|\Omega| = 2^N - 1$, visto che un insieme di N elementi ha 2^N sottoinsiemi, di cui uno è l'insieme vuoto. Inoltre, $X(\omega) = n$ se e solo se $n \in \omega$, e $\omega \setminus \{n\} \subseteq \{1, 2, \ldots, n-1\}$. Perciò

$$|\{\omega : X(\omega) = n\}| = |\{\omega' : \omega' \subseteq \{1, 2, \ldots, n-1\}\} = 2^{n-1},$$

da cui si ottiene la formula richiesta.

ii)

$$\gamma_X(t) = E\left[e^{tX}\right] = \sum_{n=1}^{N} e^{tn} \frac{2^{n-1}}{2^N - 1} = \frac{e^t}{2^N - 1} \sum_{m=0}^{N-1} \left(2e^t\right)^m = \frac{e^t}{2^N - 1} \frac{\left(2e^t\right)^N - 1}{2e^t - 1}.$$

iii) Ovviamente, per ogni $\omega \in \Omega$, $X(\omega) \geq Y(\omega)$. Pertanto, per $N \geq n \geq m \geq 1$ dobbiamo calcolare

$$P(X = n, Y = m).$$

Si noti che $X(\omega) = n, Y(\omega) = m$ se e solo se $n, m \in \omega$ e $\omega \setminus \{n, m\} \subseteq \{m+1, \ldots, n-1\}$. Perciò, il numero di ω per cui $X(\omega) = n, Y(\omega) = m$ è uguale al numero di sottoinsiemi di $\{m+1, \ldots, n-1\}$. Notare che quest'ultimo insieme è vuoto se $n = m$, altrimenti ha $n - m - 1$ elementi. Perciò

$$P(X = n, Y = m) = \begin{cases} \frac{1}{2^N - 1} & \text{se } n = m \\ \frac{2^{n-m-1}}{2^N - 1} & \text{se } n > m. \end{cases}$$

iv) L'evento $\{X - Y = k\}$, che è non vuoto per $k = 0, 1, \ldots, N-1$, si può scrivere come unione disgiunta come segue:

$$\{X - Y = k\} = \bigcup_{m=1}^{N-k} \{Y = m, X = m + k\}.$$

Usando allora quanto mostrato al punto iii),

$$P(X - Y = k) = \begin{cases} \sum_{m=1}^{N} \frac{1}{2^N - 1} = \frac{N}{2^N - 1} \text{se } k = 0 \\ \sum_{m=1}^{N-k} \frac{2^{k-1}}{2^N - 1} = (N - k) \frac{2^{k-1}}{2^N - 1} & \text{se } k > 0. \end{cases}$$

Esercizio 15 Sia $n \in \mathbb{N}$, $n \geq 3$ e indichiamo con S_n il gruppo delle permutazioni di $\{1, \ldots, n\}$, munito della probabilità P uniforme. Gli elementi di S_n saranno indicati con $\sigma = (\sigma(1), \ldots, \sigma(n))$. Introduciamo le variabili casuali scalari X, Y definite su S_n:

$$X(\sigma) \;:=\; \sigma(1)\,, \qquad Y(\sigma) \;:=\; \sigma(2)\,.$$

(a) Si mostri che, per ogni $i, j \in \{1, 2, \ldots, n\}$, la densità congiunta di (X, Y) è data da

$$p_{X,Y}(i,j) \;=\; \begin{cases} \dfrac{1}{c_n} & \text{se } i \neq j \\ 0 & \text{se } i = j \end{cases}\,,$$

dove c_n è un'opportuna costante che è richiesto di determinare.

(b) (*) Si determini la densità della variabile $D := Y - X$.
[*Sugg:* basta calcolare $p_D(m)$ per $m > 0$, poiché per simmetria $p_D(-m) = p_D(m)$.]

Indichiamo ora con Z, W due variabili casuali scalari indipendenti, definite su un altro spazio di probabilità $(\Omega, \widetilde{P})$, ciascuna con distribuzione uniforme nell'insieme $\{1, \ldots, n\}$: in altri termini, $\widetilde{P}(Z = i) = \frac{1}{n}$, per ogni $i \in \{1, \ldots, n\}$, e analogamente per W.

(c) Si calcoli $\widetilde{P}(Z \neq W)$.

(d) Si mostri che, per ogni $i, j \in \{1, 2, \ldots, n\}$, si ha che $\widetilde{P}(Z = i, W = j \,|\, Z \neq W) = p_{X,Y}(i,j)$.

Soluzione.

(a) Per definizione di spazio di probabilità uniforme

$$p_{X,Y}(i,j) \;=\; P(X = i, Y = j) \;=\; \frac{|\{\sigma \in S_n :\ \sigma(1) = j,\, \sigma(2) = j\}|}{|S_n|}\,.$$

Chiaramente $p_{X,Y}(i,j) = 0$ se $i = j$, perché le permutazioni sono biunivoche e dunque non si può avere $\sigma(1) = \sigma(2)$. Per $i \neq j$, le permutazioni $\sigma \in S_n$ tali che $\sigma(1) = j$, $\sigma(2) = j$ sono in corrispondenza con le applicazioni biunivoche da $\{3, \ldots, n\}$ a valori in $\{1, \ldots, n\} \setminus \{i, j\}$, di conseguenza $|\{\sigma \in S_n :\ \sigma(1) = j,\, \sigma(2) = j\}| = (n-2)!$ e si ottiene

$$p_{X,Y}(i,j) \;=\; \frac{(n-2)!}{n!} \;=\; \frac{1}{c_n}\,, \qquad \text{dove } \; c_n \;:=\; n(n-1)\,.$$

(Si noti che $c_n = |\{(i,j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} : \, i \neq j\}| = |(X,Y)(S_n)|$.)

(b) Chiaramente i valori possibili per la variabile D sono $D(S_n) = \{-(n-1), \ldots, (n-1)\} \setminus \{0\}$. Per $m \in D(S_n)$, con $m > 0$, si ha

$$\{D = m\} \;=\; \bigcup_{k=1}^{n-m} \{X = k, Y = k + m\}\,,$$

e dato che gli eventi che appaiono nell'unione sono disgiunti segue che per ogni $m \in \{1, \ldots, n-1\}$

$$P(D = m) \;=\; \sum_{k=1}^{n-m} P(X = k, Y = k + m) \;=\; \frac{1}{c_n}\,(n - m) \;=\; \frac{n - m}{n(n-1)}\,.$$

Con analoghi argomenti (oppure per simmetria) si ha $P(D = m) = P(D = -m)$ per $m < 0$, per cui la formula generale è

$$P(D = m) \;=\; \frac{n - |m|}{n(n-1)}\,,$$

per ogni $m \in \{-(n-1), \ldots, (n-1)\} \setminus \{0\}$.

(c)

$$\widetilde{P}(Z \neq W) \;=\; 1 - \widetilde{P}(Z = W) \;=\; 1 - \sum_{i=1}^{n} \widetilde{P}(Z = i, W = i) \;=\; 1 - n \cdot \frac{1}{n^2} \;=\; 1 - \frac{1}{n}\,.$$

(d) Chiaramente $\widetilde{P}(Z = i, W = j \,|\, Z \neq W) = 0$ se $i = j$. Per $i \neq j$, con $i, j \in \{1, \ldots, n\}$, si ha

$$\begin{aligned}\widetilde{P}(Z = i, W = j \,|\, Z \neq W) \;&=\; \frac{\widetilde{P}(Z = i, W = j)}{\widetilde{P}(Z \neq W)} \;=\; \frac{\widetilde{P}(Z = i)\widetilde{P}(W = j)}{\widetilde{P}(Z \neq W)} \;=\; \frac{\frac{1}{n}\frac{1}{n}}{1 - \frac{1}{n}} \\ &=\; \frac{1}{n(n-1)} \;=\; \frac{1}{c_n} \;=\; p_{X,Y}(i,j)\,.\end{aligned}$$

Esercizio 16 Siano W, T variabili casuali indipendenti, con la seguente distribuzione:

$$P(T = 0) \;=\; P(T = 1) \;=\; \frac{1}{2}\,, \qquad P(W = n) \;=\; p\,(1-p)^{n-1}\,,\ \forall n \in \mathbb{N} := \{1, 2, \ldots\}\,,$$

dove $p \in (0,1)$ è un parametro fissato. In altri termini, $T \sim Be(\frac{1}{2})$ mentre $W - 1 \sim Ge(p)$.

Definiamo la variabile

$$X \;:=\; W\,\mathbf{1}_{\{T=0\}} \;+\; \frac{1}{W}\,\mathbf{1}_{\{T=1\}}\,,$$

che può dunque assumere come valori i numeri naturali e i reciproci dei numeri naturali, ossia $X(\Omega) = \mathbb{N} \cup \{\frac{1}{n}\}_{n \in \mathbb{N}}$, dove $\mathbb{N} = \{1, 2, \ldots\}$ (0 escluso).

(a) Si determini la densità discreta di X.

(b) Si mostri che la variabile $Y := 1/X$ ha la stessa distribuzione di X

(c) Si calcoli $E(X)$.
[Ricordiamo le relazioni $\sum_{n=1}^{\infty} \frac{1}{n} x^n \;=\; -\log(1-x)$ e $\sum_{n=1}^{\infty} n\, x^{n-1} \;=\; (1-x)^{-2}$, valide per $|x| < 1$.]

Soluzione.

(a) La densità discreta di X vale

$$\begin{aligned} p_X(n) &= P(W=n,\, T=0) = \frac{1}{2}p(1-p)^{n-1}\,, \quad \text{per } n\in\mathbb{N},\, n\geq 2\,,\\ p_X(1) &= P(W=1) = p\,,\\ p_X\left(\frac{1}{n}\right) &= P(W=n,\, T=1) = \frac{1}{2}p(1-p)^{n-1}\,, \quad \text{per } n\in\mathbb{N},\, n\geq 2\,. \end{aligned}$$

(b) Si ha $Y(\Omega)=\mathbb{N}\cup\{\frac{1}{n}\}_{n\in\mathbb{N}}$. Dal punto precedente è chiaro che $P(X=n)=P(X=\frac{1}{n})$, per ogni $n\in\mathbb{N}$; di conseguenza $p_Y(n)=p_X(\frac{1}{n})=p_X(n)$ e $p_Y(\frac{1}{n})=p_X(n)=p_X(\frac{1}{n})$.

(c) Si ha

$$\begin{aligned} E(X) &= \sum_{x\in X(\Omega)} x\cdot p_X(x) = \sum_{n=2}^{\infty}\frac{1}{n}p_X\left(\frac{1}{n}\right) + 1\cdot p_X(1) + \sum_{n=2}^{\infty} n\,p_X(n)\\ &= \sum_{n=2}^{\infty}\frac{1}{n}\frac{1}{2}p(1-p)^{n-1} + p + \sum_{n=2}^{\infty} n\,\frac{1}{2}p(1-p)^{n-1}\\ &= \frac{p}{2(1-p)}\sum_{n=1}^{\infty}\frac{1}{n}(1-p)^n + \frac{p}{2}\sum_{n=1}^{\infty} n\,(1-p)^{n-1} = \frac{1}{2}\left(\frac{p}{(1-p)}\log\frac{1}{1-p} + \frac{1}{p}\right)\,. \end{aligned}$$

Esercizio 17 Lancio un dado regolare a sei facce una prima volta: se esce un numero in $\{1,2,3,4,5\}$ mi fermo, altrimenti rilancio il dado; se nel secondo lancio esce un numero in $\{1,2,3,4,5\}$ mi fermo, altrimenti rilancio il dado una terza volta; e così via.

Indichiamo con T il numero totale di lanci effettuati, con X_i il risultato dell'i-esimo lancio (per $i\in\mathbb{N}:=\{1,2,\ldots\}$) e con $Y:=X_T$ il risultato dell'ultimo lancio.

(a) Si determini la legge (cioè i valori assunti e la densità discreta) delle variabili T e X_i.

(b) Per ogni valore di $n\in\mathbb{N}$ e $a\in\{1,2,3,4,5\}$, si esprima l'evento $\{T=n, Y=a\}$ in termini delle variabili aleatorie $X_1,\ldots,X_n$. Si calcoli dunque la densità congiunta delle variabili casuali (T,Y).

(c) Si determini la legge di Y. Le variabili T e Y sono indipendenti?

Soluzione.

(a) T è l'istante in cui avviene il primo successo in uno schema di prove ripetute e indipendenti, in cui la probabilità di successo in ogni prova vale $\frac{5}{6}$. Di conseguenza, $T(\Omega)=\mathbb{N}=\{1,2,\ldots\}$ e per $n\in\mathbb{N}$ si ha $p_T(n)=\frac{5}{6}\,(\frac{1}{6})^{n-1}$.

(b) Per $n=1$ si ha $\{T=1,Y=a\}=\{X_1=a\}$, per cui

$$p_{T,Y}(1,a)=\frac{1}{6}\,.$$

Per $n\geq 2$ possiamo scrivere $\{T=n,Y=a\}=\bigcap_{i=1}^{n-1}\{X_i=6\}\cap\{X_n\leq 5\}$ e per l'indipendenza delle variabili X_i si ottiene

$$p_{T,Y}(n,a)=\left(\frac{1}{6}\right)^{n-1}\cdot\frac{1}{6}=\left(\frac{1}{6}\right)^{n}\,.$$

In definitiva, si ha $p_{T,Y}(n,a) = (\frac{1}{6})^n$ per ogni $n \in \mathbb{N}$ e $a \in \{1,2,3,4,5\}$.

(c) Chiaramente $Y(\Omega) = \{1,2,3,4,5\}$. La legge marginale della variabile Y è data da

$$p_Y(a) = \sum_{n\in\mathbb{N}} P_{T,Y}(n,a) = \sum_{n=1}^{\infty}\left(\frac{1}{6}\right)^n = \sum_{n=0}^{\infty}\left(\frac{1}{6}\right)^n - 1 = \frac{1}{1-\frac{1}{6}} - 1 = \frac{6}{5} - 1 = \frac{1}{5}\,,$$

per ogni $a \in \{1,2,3,4,5\}$. Per quanto determinato ai punti precedenti, per ogni $n \in \mathbb{N}$ e $a \in \{1,2,3,4,5\}$ si ha

$$p_{T,Y}(n,a) = \left(\frac{1}{6}\right)^n = \frac{1}{5}\cdot\frac{5}{6}\left(\frac{1}{6}\right)^{n-1} = p_Y(a)\cdot p_T(n)\,,$$

per cui le variabili T e Y sono indipendenti.

Esercizio 18 Un gioco a premi ha il seguente funzionamento. Il montepremi iniziale è $C > 0$. Il concorrente lancia due volte una moneta. Se l'esito del primo e del secondo lancio sono uguali, il concorrente vince l'intero montepremi. In caso contrario il montepremi si dimezza. Il concorrente lancia la moneta una terza volta. Se l'esito del terzo lancio è uguale a quello del secondo, allora il concorrente vince $C/2$. In caso contrario il montepremi si dimezza ulteriormente, e così via. Sia X il montepremi vinto dal concorrente al termine del gioco. Si accetti il fatto che il gioco ha termine, cioè che non sia possibile avere una successione infinita di esiti in cui ogni esito è diverso dal precedente.

(a) Si assuma che la moneta sia equilibrata. Determinare la densità discreta di X.

(b) Sempre nell'ipotesi che la moneta sia equilibrata, si determini il valor medio di X (ricordare la serie geometrica $\sum_{n\geq 0} a^n = \frac{1}{1-a}$, per $|a| < 1$).

(c) * Supponiamo ora che la moneta non sia equilibrata: la probabilità che esca *testa* è $p \in (0,1)$. Determinare la densità e il valor medio di X.

Soluzione.

(a) Si tratta di usare lo schema delle prove ripetute indipendenti. Si noti che l'evento $\left\{X = \frac{C}{2^{k-1}}\right\}$ è costituito dalle sequenze di esiti di lanci di una moneta in cui nei primi k esiti le teste e le croci si alternano, mentre l'esito $k+1$-mo è uguale al k-mo. Limitandosi ai primi $k+1$ lanci, ci sono due sequenze con questa proprietà, a seconda che l'esito del primo lancio sia testa o croce. Ognuna di queste sequenze ha probabilità $\frac{1}{2^{k+1}}$. Quindi

$$P\left(X = \frac{C}{2^{k-1}}\right) = \frac{2}{2^{k+1}} = \frac{1}{2^k}.$$

(b) Ne segue che

$$E(X) = \sum_{k\geq 1}\frac{C}{2^{k-1}}\frac{1}{2^k} = \frac{C}{2}\sum_{k\geq 1}\frac{1}{4^{k-1}} = \frac{C}{2}\frac{1}{1-\frac{1}{4}} = \frac{2C}{3}.$$

(c) È ancora vero che l'evento $\left\{X = \frac{C}{2^{k-1}}\right\}$, limitandosi ai primi $k+1$ lanci, è costituito da due sequenze, quella che inizia con una testa e quella che inizia con una croce. Non necessariamente queste due sequenze hanno la stessa probabilità. Conviene distinguere due casi:

-$k = 2n$, cioè k pari. La sequenza di $k+1 = 2n+1$ lanci che inizia per testa contiene n teste e $n+1$ croci, e ha dunque probabilità $p^n(1-p)^{n+1}$. Analogamente, la sequenza che comincia per croce ha probabilità $p^{n+1}(1-p)^n$. Pertanto

$$P\left(X = \frac{C}{2^{2n-1}}\right) = p^{n+1}(1-p)^n + p^n(1-p)^{n+1} = p^n(1-p)^n.$$

-$k = 2n-1$, cioè k dispari. In questo caso le due sequenze dei primi $k+1 = 2n$ lanci contengono, rispettivamente, $n+1$ teste e $n-1$ croci e viceversa, da cui segue che

$$P\left(X = \frac{C}{2^{2n-2}}\right) = p^{n+1}(1-p)^{n-1} + p^{n-1}(1-p)^{n+1} = p^{n-1}(1-p)^{n+1}\left[p^2 + (1-p)^2\right].$$

Infine

$$E(X) = \sum_{n\geq 1} \frac{C}{2^{2n-1}} p^n(1-p)^n + \left[p^2 + (1-p)^2\right] \sum_{n\geq 1} \frac{C}{2^{2n-2}} p^{n-1}(1-p)^{n-1}$$
$$= C\left[\frac{1}{2}p(1-p) + p^2 + (1-p)^2\right] \sum_{n\geq 1} \left(\frac{p(1-p)}{4}\right)^{n-1} = C\frac{\frac{1}{2}p(1-p) + p^2 + (1-p)^2}{1 - \frac{p(1-p)}{4}}$$

Esercizio 19 Si consideri l'insieme $\Omega := \{0,1\}^n \times \{0,1\}^n$, munito della probabilità uniforme P. Gli elementi di Ω sono dunque coppie (σ, θ), dove $\sigma, \theta \in \{0,1\}^n$. Si consideri la variabile aleatoria $X : \Omega \to \mathbb{R}$ definita da

$$X(\sigma, \theta) := |\{i = 1, 2, \ldots, n : \ \sigma(i) = \theta(i)\}|.$$

Si determini la distribuzione, la media e la varianza di X.

Soluzione. Cominciamo col contare gli elementi dell'evento $\{X = k\}$, dove $k = 0, 1, \ldots, n$. Un elemento (σ, θ) di $\{X = k\}$ può essere scelto tramite il seguente schema di scelte successive:

a) scelgo un generico τ: ho 2^n scelte possibili;

b) scelgo k componenti in cui σ deve coincidere con τ: $\binom{n}{k}$ scelte possibili. Ovviamente questa scelta identifica σ univocamente.

Infine:

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

da cui segue che $X \sim B(n, 1/2)$, e perciò $E(X) = n/2$ e $Var(X) = n/4$.

Esercizio 20 Da un mazzo di 50 carte numerate da 1 a 50, si estraggono a caso 3 carte. Siano

$$\begin{aligned} X &= \text{numero più basso estratto} \\ Z &= \text{numero più alto estratto} \\ Y &= \text{terzo numero estratto} \end{aligned}$$

(a) Determinare le distribuzioni marginali di X, Y e Z.

(b) Determinare la distribuzione di (X, Y), e mostrare che $Y - X$ ha la stessa distribuzione di X.

Soluzione.

(a) Notare che $X \in \{1, 2, \ldots, 48\}$. Se $n \in \{1, 2, \ldots, 48\}$, l'evento $\{X = n\}$ significa "oltre a n, gli altri due numeri estratti sono maggiori di n". Quindi

$$P(X = n) = \frac{\binom{50-n}{2}}{\binom{50}{3}}.$$

Con argomenti analoghi: per $n \in \{3, 4, \ldots, 50\}$

$$P(Z = n) = \frac{\binom{n-1}{2}}{\binom{50}{3}}$$

e per $n \in \{2, 3, \ldots, 49\}$

$$P(Y = n) = \frac{(n-1)(50-n)}{\binom{50}{3}}.$$

(b) Se $n \in \{1, 2, \ldots, 48\}$, $m \in \{2, 3, \ldots, 49\} = n < m$

$$P(X = n, Y = m) = \frac{50 - m}{\binom{50}{3}}.$$

Notare che $Y - X \in \{1, \ldots, 48\}$. Se $k \in \{1, \ldots, 48\}$

$$P(Y - X = k) = \sum_n P(X = n, Y = n + k) = \sum_{n=1}^{49-k} \frac{50 - n - k}{\binom{50}{3}}$$
$$= \frac{1}{\binom{50}{3}} \left[(50 - k)(49 - k) - \sum_{n=1}^{49-k} n \right] = \frac{\binom{50-k}{2}}{\binom{50}{3}}$$

Esercizio 21 Sia (Ω, P) uno spazio di probabilità discreto.

(a) Mostrare che per ogni scelta di eventi A e B

$$|P(A) - P(B)| \leq P(A \Delta B).$$

(b) Siano $X, Y : \Omega \to E$ due variabili aleatorie. Mostrare che per ogni $C \subseteq E$

$$|P(X \in C) - P(Y \in C)| \leq P(X \neq Y).$$

(c) Nelle stesse ipotesi del punto precedente, mostrare che

$$\sum_{x \in E} |P(X = x) - P(Y = x)| \leq 2P(X \neq Y).$$

(*Sugg.*: porre $C := \{x \in E : P(X = x) \geq P(Y = x)\}$, e usare il risultato del punto precedente per C e C^c).

(d) Per $n \geq 1$ siano $X_n, X : \Omega \to \mathbb{R}$ variabili aleatorie tali che

$$\lim_{n \to +\infty} P(X_n \neq X) = 0,$$

e, per ogni $\omega \in \Omega$ e $n \geq 1$,

$$|X_n(\omega)| \leq M, \quad |X(\omega)| \leq M,$$

dove $M > 0$. Mostrare che

$$\lim_{n \to +\infty} E(X_n) = E(X).$$

Soluzione.

(a) Senza perdita di generalità possiamo supporre che $P(A) \geq P(B)$ (altrimenti scambiamo i ruoli di A e B).

$$|P(A) - P(B)| = P(A) - P(B) \leq P(A) - P(A \cap B) = P(A \setminus B) \leq P(A \Delta B).$$

(b) Posto $A := \{X \in C\}$, $B := \{Y \in C\}$, dal punto precedente abbiamo che

$$|P(X \in C) - P(Y \in C)| \leq P(\{X \in C\}\Delta\{Y \in C\}).$$

Per concludere basta mostrare che

$$\{X \in C\}\Delta\{Y \in C\} \subseteq \{X \neq Y\}.$$

Ma questo è ovvio, poiché se $\omega \in \{X \in C\}\Delta\{Y \in C\}$ allora o $X(\omega) \in C, Y(\omega) \notin C$ o $Y(\omega) \in C, X(\omega) \notin C$; in entrambi i casi $X(\omega) \neq Y(\omega)$.

(c) Posto C come nel suggerimento, abbiamo

$$\begin{aligned}\sum_{x \in E} |P(X = x) - P(Y = x)| &= \sum_{x \in C} [P(X = x) - P(Y = x)] + \sum_{x \in C^c} [P(Y = x) - P(X = x)] \\ &= [P(X \in C) - P(Y \in C)] + [P(Y \in C^c) - P(X \in C^c)] \leq 2P(X \neq Y).\end{aligned}$$

(d) Basta osservare che

$$\begin{aligned}|E(X_n) - E(X)| &= \left| \sum_{x \in E} xP(X_n = x) - \sum_{x \in E} xP(X = x) \right| \leq \sum_{x \in E} |x||P(X_n = x) - P(X = x)| \\ &\leq M \sum_{x \in E} |P(X_n = x) - P(X = x)| \leq 2MP(X_n \neq X).\end{aligned}$$