Laurea Triennale in Scienze e Tecnologie per i Media, Università di Roma Tor Vergata
Calcolo delle Probabilità (ed insegnamenti mutuati)
Anno Accademico: `2023-2024`. *Titolare del corso:* `Claudio Macci`
Esame del 20 Febbraio 2024
L'Esercizio 6 è solo per coloro che sostengono l'esame da 8 crediti.

Esercizio 1. Un'urna ha 4 palline bianche e 3 nere.
`D1)` Si estraggono a caso 4 palline in blocco. Calcolare la probabilità di estrarre 2 palline bianche e 2 nere.
`D2)` Si estraggono a caso 3 palline in blocco. Calcolare la probabilità di estrarre 1 pallina nera e 2 bianche.
`D3)` Si estraggono a caso 2 palline, una alla volta e senza reinserimento. Calcolare la probabilità che la prima estratta sia nera e la seconda sia bianca (cioè la probabilità di estrarre la sequenza di colori $(\text{nera}, \text{bianca})$).

Esercizio 2. Si lancia un dado non equo: se esce il numero 1 (e questo accade con probabilità $\frac{1}{4}$), si lancia una moneta con due teste; se esce un numero diverso da 1, si lancia una moneta equa.
`D4)` Calcolare la probabilità di aver lanciato la moneta equa sapendo che è uscita testa.

Esercizio 3. Siano $q, r \in (0,1)$ arbitrariamente fissati. Consideriamo la seguente densità congiunta discreta:

$$p_{X_1,X_2}(x_1,x_2) = r^{x_1}(1-q)^{x_2}q(1-r) \quad \text{per } x_1, x_2 \geq 0 \text{ interi.}$$

`D5)` Verificare che $P(X_1 = X_2) = \frac{q(1-r)}{1-r(1-q)}$.
`D6)` Verificare che $P(X_1 = 0 | X_2 = 0) = P(X_1 = 0)$.

Esercizio 4. Sia $b > 1$ arbitrariamente fissato. Sia X una variabile aleatoria con distribuzione uniforme su (b, b^2).
`D7)` Trovare la funzione di distribuzione di $Y = e^{-X/b}$.
`D8)` Verificare che $\mathbb{E}[Y^b] = \frac{e^{-b}-e^{-b^2}}{b^2-b}$.

Esercizio 5. Poniamo $\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{x^2}{2}} dx$.
`D9)` Sia X una variabile aleatoria Normale con media 2 e varianza 16.
Calcolare $P(0 \leq X \leq 3)$ esprimendo il risultato con la funzione Φ con argomenti positivi.
`D10)` Siano $X_1, \ldots, X_{100}$ variabili aleatorie i.i.d. con media 1 e varianza 25.
Calcolare $P(X_1 + \cdots + X_{100} > 110)$ con l'approssimazione Normale esprimendo il risultato con la funzione Φ.

Esercizio 6. Sia $\{X_n : n \geq 1\}$ una catena di Markov omogenea su $E = \{1,2,3,4,5\}$ e con matrice di transizione

$$P = \begin{pmatrix} 7/10 & 3/10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ (1-q)/4 & (1-q)/4 & q & (1-q)/4 & (1-q)/4 \\ 0 & 0 & 0 & 1/4 & 3/4 \\ 0 & 0 & 3q/4 & 3(1-q)/4 & 1/4 \end{pmatrix},$$

dove $q \in [0,1]$.
`D11)` Sia $q = 0$. Giustificare l'esistenza di $\lim_{n\to\infty} p_{ij}^{(n)}$ per $i, j \in \{4,5\}$, e verificare che si ottiene sempre lo stesso valore (per ogni scelta di $i, j \in \{4,5\}$).
`D12)` Sia $q = 1$. Calcolare i tempi medi di primo arrivo nello stato 3 partendo da ciascuno degli stati $1,2,4,5$.

Cenno alle soluzioni (*Ogni segnalazione di errori o sviste (sempre possibili) è gradita*)

Esercizio 1.
`D1)` La probabilità richiesta è $\frac{\binom{4}{2}\binom{3}{2}}{\binom{7}{4}} = \frac{6\cdot 3}{35} = \frac{18}{35}$.
`D2)` La probabilità richiesta è $\frac{\binom{4}{2}\binom{3}{1}}{\binom{7}{3}} = \frac{6\cdot 3}{35} = \frac{18}{35}$.
`D3)` Con notazioni ovvie la probabilità condizionata richiesta è

$$P(N_1 \cap B_2) = P(B_2|N_1)P(N_1) = \frac{4}{6} \cdot \frac{3}{7} = \frac{2}{7}.$$

Esercizio 2.
`D4)` Con notazioni ovvie viene chiesto di calcolare una probabilità condizionata $P(E|T)$. Per la formula di Bayes (in quel che segue si tiene conto che $P(T) = P(T|E)P(E) + P(T|E^c)P(E^c)$ per la formula delle probabilità totali) si ha

$$P(E|T) = \frac{P(T|E)P(E)}{P(T)} = \frac{\frac{1}{2}\cdot\frac{3}{4}}{\frac{1}{2}\cdot\frac{3}{4} + 1\cdot\frac{1}{4}} = \frac{3/8}{(3+2)/8} = \frac{3}{5}.$$

Esercizio 3.
`D5)` Si ha

$$P(X_1 = X_2) = \sum_{k\geq 0} p_{X_1,X_2}(k,k) = q(1-r)\sum_{k\geq 0}(r(1-q))^k = \frac{q(1-r)}{1-r(1-q)}.$$

`D6)` Partiamo dal *primo membro* della uguaglianza da verificare:

$$P(X_1 = 0|X_2 = 0) = \frac{P(\{X_1 = 0\} \cap \{X_2 = 0\})}{P(X_2 = 0)} = \frac{p_{X_1,X_2}(0,0)}{P(X_2 = 0)} = \frac{q(1-r)}{P(X_2 = 0)}.$$

Poi osserviamo che

$$P(X_2 = 0) = \sum_{k\geq 0} p_{X_1,X_2}(k,0) = q(1-r)\sum_{k\geq 0} r^k(1-q)^0 = q(1-r)\frac{1}{1-r} = q$$

e, ritornando ai passaggi iniziali, per il primo membro possiamo dire che

$$P(X_1 = 0|X_2 = 0) = \frac{q(1-r)}{P(X_2 = 0)} = \frac{q(1-r)}{q} = 1 - r.$$

A questo punto completiamo la verifica richiesta calcolando il *secondo membro*. In effetti si ha

$$P(X_1 = 0) = \sum_{k\geq 0} p_{X_1,X_2}(0,k) = q(1-r)\sum_{k\geq 0} r^0(1-q)^k = q(1-r)\frac{1}{1-(1-q)} = \frac{q(1-r)}{q} = 1 - r.$$

Osservazione. L'uguaglianza da verificare segue da una affermazione più generale che si può dimostrare (e che per certi versi si può intuire dalla espressione della densità congiunta), cioè che le variabili aleatorie X_1 e X_2 sono indipendenti.

Esercizio 4.
`D7)` Si ha $P(e^{-b} \leq Y \leq e^{-1}) = 1$ e quindi

$$F_Y(y) = \begin{cases} 0 & \text{se } y \leq e^{-b}, \\ (*) & \text{se } e^{-b} < y < e^{-1}, \\ 1 & \text{se } y \geq e^{-1}. \end{cases}$$

Per $y \in (e^{-b}, e^{-1})$ si ha

$$(*) = P(e^{-X/b} \leq y) = P(X \geq -b\log y) = \int_{-b\log y}^{b^2} \frac{1}{b^2 - b}dx = \left[\frac{x}{b^2 - b}\right]_{x=-b\log y}^{x=b^2} = \frac{b^2 + b\log y}{b^2 - b} = \frac{b + \log y}{b - 1}.$$

D8) A partire dalla risposta precedente si ha che $f_Y(y) = \frac{1}{y(b-1)} 1_{(e^{-b}, e^{-1})}(y)$. Quindi

$$\mathbb{E}[Y^b] = \int_{e^{-b}}^{e^{-1}} y^b \frac{1}{y(b-1)} dy = \frac{1}{b-1} \int_{e^{-b}}^{e^{-1}} y^{b-1} dy = \frac{1}{b-1} \left[\frac{y^b}{b} \right]_{y=e^{-b}}^{y=e^{-1}} = \frac{e^{-b} - e^{-b^2}}{b^2 - b}.$$

Osservazione. In realtà non serve fare riferimento alla risposta precedente osservando che $Y^b = (e^{-X/b})^b = e^{-X}$. Infatti si ha

$$\mathbb{E}[Y^b] = \mathbb{E}[e^{-X}] = \int_b^{b^2} e^{-x} \frac{1}{b^2 - b} dx = \frac{1}{b^2 - b} \int_b^{b^2} e^{-x} dx = \frac{1}{b^2 - b} [-e^{-x}]_{x=b}^{x=b^2} = \frac{e^{-b} - e^{-b^2}}{b^2 - b}.$$

Esercizio 5.

D9) La variabile aleatoria standardizzata di X è $X^* = \frac{X-2}{\sqrt{16}} = \frac{X-2}{4}$. Allora

$$\begin{aligned} P(0 \leq X \leq 3) &= P\left(\frac{0-2}{4} \leq X^* \leq \frac{3-2}{4}\right) = \Phi\left(\frac{3-2}{4}\right) - \Phi\left(\frac{0-2}{4}\right) \\ &= \Phi\left(\frac{1}{4}\right) - \Phi\left(-\frac{1}{2}\right) = \Phi\left(\frac{1}{4}\right) - \left(1 - \Phi\left(\frac{1}{2}\right)\right) = \Phi\left(\frac{1}{4}\right) + \Phi\left(\frac{1}{2}\right) - 1. \end{aligned}$$

D10) Si ha

$$\begin{aligned} P(X_1 + \cdots + X_{100} > 110) &= P\left(\frac{X_1 + \cdots + X_{100} - 100 \cdot 1}{\sqrt{25}\sqrt{100}} > \frac{110 - 100 \cdot 1}{\sqrt{25}\sqrt{100}}\right) \\ &\approx 1 - \Phi\left(\frac{110 - 100 \cdot 1}{\sqrt{25}\sqrt{100}}\right) = 1 - \Phi\left(\frac{10}{50}\right) = 1 - \Phi\left(\frac{1}{5}\right). \end{aligned}$$

Esercizio 6.

D11) Essendo $q = 0$ si ha

$$P = \begin{pmatrix} 7/10 & 3/10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1/4 & 3/4 \\ 0 & 0 & 0 & 3/4 & 1/4 \end{pmatrix}.$$

La classe $\{4, 5\}$ è chiusa e irriducibile. Inoltre la catena ristretta a $\{4, 5\}$ è regolare (per l'Osservazione 5.16; infatti qui abbiamo $p_{44}, p_{55} > 0$). Quindi si può applicare il Teorema di Markov alla catena ristretta a $\{4, 5\}$, e si ha

$$\lim_{n \to \infty} p_{ij}^{(n)} = \pi_j \quad \text{per } i, j \in \{4, 5\}$$

dove (π_4, π_5) è l'unica distribuzione stazionaria per la catena ristretta a $\{4, 5\}$. Inoltre la catena ristretta ha matrice di transizione bistocastica perché si ha

$$\begin{pmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{pmatrix};$$

quindi si ha $\pi_4 = \pi_5 = \frac{1}{2}$. In conclusione il limite è sempre uguale ad $\frac{1}{2}$ (per ogni scelta di $i, j \in \{4, 5\}$).

D12) Essendo $q = 1$ si ha

$$P = \begin{pmatrix} 7/10 & 3/10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 3/4 \\ 0 & 0 & 3/4 & 0 & 1/4 \end{pmatrix}.$$

Lo stato 3 è assorbente. Inoltre $T = \{1, 2, 4, 5\}$ perché si può andare da qualsiasi stato in T nello stato 3, e non vale il viceversa. Allora indichiamo con $\mu_1, \mu_2, \mu_4, \mu_5$ i tempi medi richiesti e tali valori medi soddisfano il seguente sistema

$$\begin{cases} \mu_1 = 1 + \frac{7}{10}\mu_1 + \frac{3}{10}\mu_2 \\ \mu_2 = 1 + \mu_4 \\ \mu_4 = 1 + \frac{1}{4}\mu_4 + \frac{3}{4}\mu_5 \\ \mu_5 = 1 + \frac{1}{4}\mu_5. \end{cases}$$

Le soluzioni si trovano a partire dall'ultima equazione, e sostituendo ogni volta in ciascuna delle precedenti. In dettaglio:
- dall'ultima (la quarta) equazione si ha $\frac{3}{4}\mu_5 = 1$, e quindi $\mu_5 = 4/3$;
- dalla terza equazione, dopo aver sostituito $\mu_5 = 4/3$, si ha $\frac{3}{4}\mu_4 = 1 + 1$, e quindi $\mu_4 = 8/3$;
- dalla seconda equazione, dopo aver sostituito $\mu_4 = 8/3$, si ha $\mu_2 = 1 + \frac{8}{3}$, e quindi $\mu_2 = 11/3$;
- dalla prima equazione, dopo aver sostituito $\mu_2 = 11/3$, si ha $\frac{3}{10}\mu_1 = 1 + \frac{11}{10}$, e quindi $\mu_1 = 21/3 = 7$.

Osservazione. In effetti i tempi di primo arrivo in 3 esprimibili tramite opportune variabili aleatorie geometriche traslate indipendenti, e in corrispondenza i tempi medi richiesti si ottengono come opportune somme di speranze matematiche per linearità. Se usiamo la notazione $G_1(p)$ per una geometrica traslata di parametro p, per abuso di notazione si ha:
- $\mu_5 = \mathbb{E}[G_1(3/4)] = 4/3$;
- $\mu_4 = \mathbb{E}[G_1(3/4)] + \mathbb{E}[G_1(3/4)] = (4 + 4)/3 = 8/3$;
- $\mu_2 = 1 + \mathbb{E}[G_1(3/4)] + \mathbb{E}[G_1(3/4)] = 1 + (4 + 4)/3 = 11/3$;
- $\mu_1 = \mathbb{E}[G_1(3/10)] + 1 + \mathbb{E}[G_1(3/4)] + \mathbb{E}[G_1(3/4)] = 10/3 + 1 + (4 + 4)/3 = 21/3 = 7$.