Probabilità e Statistica [CT0111]
Corso di Laurea triennale in Informatica, a.a. 2021/22

Docente: Isadora Antoniano Villalobos
Esame **Soluzioni**, 15 settembre 2022

Cognome: ____________________ Nome: ____________________

Matricola: ____________________ Firma: ____________________

ISTRUZIONI: DA LEGGERE CON ATTENZIONE!

Questo compito è composto di **6 domande**, per un totale di **30 punti** e dura complessivamente **90 minuti**.

Ai fini della valutazione si terrà conto solo ed esclusivamente di quanto riportato negli appositi spazi. Qualora si avesse bisogno di più spazio, sarà possibile continuare sulla prima e ultima pagina, indicandolo chiaramente.

Si richiede una traccia dello svolgimento di ogni esercizio e dei calcoli effettuati per rispondere alle domande. La sostituzione del calcolo e/o risposta numerica con un opportuno comando di R è consentita senza penalizzazione, soltanto dove indicato esplicitamente. Utilizzare almeno 4 decimali di precisione per tutti i calcoli numerici.

È consentito il solo uso della calcolatrice, della tavola della distribuzione Normale presente nel sito Moodle del corso e di un foglio A4 (entrambi lati) con formule e annotazioni. Non sono ammessi appunti, libri ed esercizi svolti.
Il compito non sarà corretto se ci sono informazioni mancanti: Cognome, Nome, Matricola o Firma

Question:	1	2	3	4	5	6	Total
Points:	5	5	5	5	5	5	30
Score:							

Domanda 1 (5 punti)

Per ognuna delle 5 domande a scelta, leggere attentamente e selezionare un'unica risposta, indicandola chiaramente. Le giustificazioni non sono richieste e, se fornite, non verranno valutate (quindi non perdere tempo).

(a) Si consideri la seguente matrice di transizione:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Quale dei seguenti vettori può essere la distribuzione stazionaria per la catena di Markov corrispondente?

i) Tutti.
ii) Nessuno.
iii) $\pi = (1, 0, 0)$
iv) $\pi = (0.5, 0.3, 0.3, -0.1)$
v) $\pi = (0.5, 0.3, 0.3, 0.1)$

Soluzione: *ii)*

(b) Si selezioni il commando di R che approssima, usando un metodo Monte Carlo, la probabilità che una viariabile con distribuzione Normale con media $\mu = -1.1$ e varianza $\sigma^2 = 4$ assuma valori maggiori o uguali a 2.9.

i) $\texttt{pnorm}(2.9, \texttt{mean} = -1.1, \texttt{sd} = \texttt{sqrt}(4))$
ii) $\texttt{dnorm}(2.9, \texttt{mean} = -1.1, \texttt{sd} = \texttt{sqrt}(4))$
iii) $\texttt{mean}(\texttt{rnorm}(1000, -1.1, \texttt{sqrt}(4)) >= 2.9)$
iv) $\texttt{mean}(\texttt{rbinom}(1000, -1.1, \texttt{sqrt}(4)) >= 2.9)$
v) $\texttt{mean}(\texttt{rnorm}(1000, -1.1, 4) <= 2.9)$

Soluzione: *iii)*

(c) Se F è una funzione di ripartizione per una variabile continua, quale delle seguenti affermazioni NON è necessariamente vera

i) Sono tutte affermazioni necessariamente vere
ii) $F(x) \geq 0, \quad \text{for all } x \in (-\infty, \infty).$
iii) $\lim_{x\to-\infty} F(x) = 0$
iv) $\lim_{x\to\infty} F(x) = 1$
v) $\lim_{x\to 0} F(x) = 1/2$

Soluzione: *v)*

(d) Se A e B sono due eventi con $A \cap B = \emptyset$ quale delle seguenti affermazioni è sicuramente vera?

i) $\mathbb{P}[A] > 1/2$ e $\mathbb{P}[B] > 1/2$
ii) $\mathbb{P}[A \cup B] > 1$

iii) $\mathbb{P}[A] = \mathbb{P}[B]$

iv) $\mathbb{P}[A \cap B] = 0$

v) Nessuna delle precedenti.

Soluzione: *iv)*

(e) Se X e Y sono due variabili casuali con media pari a 0, e $\mathbb{E}[XY] = 0$ quale delle seguenti espressioni è sicuramente vera?

i) $\mathbb{E}[aX + bY + c] = 0$ per qualsiasi valore di $a, b, c \in \mathbb{R}$

ii) $\mathbb{E}[X + Y - 1] = 1$

iii) $\text{Cov}[X, Y] = 0$.

iv) $\mathbb{E}(Y^2) = \mathbb{E}(X^2)$.

v) Sono tutte vere.

Soluzione: *iii)*

Domanda 2 (5 punti)

Un commerciante fa una ordinazione di 100 transistor. La sua politica consiste nel provarne 10 scelti a caso e rifiutare tutta l'ordinazione se almeno 2 di essi sono difettosi.

(a) Se una particolare ordinazione effettivamente contiene 20 pezzi difettosi qual è la distribuzione della variabile X = numero di pezzi difettosi scelti dal commerciante?

Soluzione: La variabile X ha una distribuzione ipergeometrica con parametri $N = 100$, $K = 20$, $n = 10$: $X \sim \text{Ig}(100, 20, 10)$

(b) Se una particolare ordinazione effettivamente contiene 20 pezzi difettosi qual è la probabilità che venga accettata?

Soluzione: La ordinazione viene rifiutata se $X \geq 2$, quindi la probabilità richiesta è

$$\mathbb{P}[X < 2] = \mathbb{P}[X = 0] + \mathbb{P}[X = 1] = \frac{\binom{20}{0}\binom{80}{10}}{\binom{100}{10}} + \frac{\binom{20}{1}\binom{80}{9}}{\binom{100}{10}} = 0.0951 + 0.2679 = 0.3630$$

Alternativamente, usando R,

$$\begin{aligned}\mathbb{P}[X < 2] &= \text{phyper}(1, 20, 80, 10)\\ &= \mathbb{P}[X = 0] + \mathbb{P}[X = 0] = \text{dhyper}(0, 20, 80, 10) + \text{dhyper}(1, 20, 80, 10)\end{aligned}$$

(c) Se 5 diverse ordinazioni contengono ognuna 20 pezzi difettosi, qual è la probabilità che soltanto 2 vengano accettate?

Soluzione: Sia Y il numero di ordenazioni accettate. Quindi Y ha una distribuzione binomiale con parametri $n = 5$ e $p = 0.3630$: $Y \sim \text{Bin}(5, 0.363)$. La probabilità richiesta è

$$\mathbb{P}[Y = 2] = \binom{5}{2}(0.363)^2(1 - 0.363)^3 = 0.3406$$

Alternativamente, in R: $P(Y = 2) = \text{dbinom}(2, 5, 0.363)$

Domanda 3 (5 punti)

Hai chiesto ad un vicino di innaffiare una piantina delicata mentre sei in vacanza. Pensi che senza acqua la piantina muoia con probabilità 0.8, mentre se innaffiata questa probabilità si ridurrebbe a 0.15. La tua fiducia che il vicino si ricordi di innaffiarla è del 90%

(a) Qual è la probabilità che la pianta sia ancora viva al tuo ritorno?

Soluzione: Definiamo i seguenti eventi:

A = Il vicino si è ricordato di innaffiare la pianta

V = Al tuo ritorno la pianta è ancora viva

Abbiamo $\mathbb{P}[A] = 0.9$, $\mathbb{P}[V|A] = 1 - 0.15 = 0.85$, $\mathbb{P}[V|\bar{A}] = 1 - 0.8 = 0.2$
Per la legge della probabilità totale, la probabilità richiesta è

$$\mathbb{P}[V] = \mathbb{P}[V|A]\mathbb{P}[A] + \mathbb{P}[V|\bar{A}]\mathbb{P}[\bar{A}] = 0.85(0.9) + 0.2(1 - 0.9) = 0.785$$

(b) Se al tuo ritorno la pianta fosse morta, quale sarebbe la probabilità che il vicino si fosse dimenticato di innaffiarla?

Soluzione: Per il teorema di Bayes, si ha che

$$\mathbb{P}[\bar{A}|\bar{V}] = \frac{\mathbb{P}[\bar{V}|\bar{A}]\mathbb{P}[\bar{A}]}{\mathbb{P}[\bar{V}]} = \frac{0.8(1-0.9)}{1-0.785} = 0.3721$$

Domanda 4 (5 punti)

Siano X e Y due variabili casuali discrete con distribuzione congiunta

$X \setminus Y$	1	2
1	0.13	0.17
2	0.09	0.17
3	0.22	0.22

Determinare:

(a) $P(X = 3)$

Soluzione: Le probabilità marginali della X si trovano sommando i valori sulle righe. In particolare,

$$\mathbb{P}[X = 3] = 0.22 + 0.22 = 0.44$$

.

(b) $P(Y = 1|X = 3)$

Soluzione: Le probabilità condizionate della Y data la X si trovano dividendo le probabilità congiunte per le marginali della X. In particolare,

$$\mathbb{P}[Y = 1|X = 3] = \frac{\mathbb{P}[X = 1, Y = 3]}{\mathbb{P}[X = 1]} = \frac{0.22}{0.44} = 0.5$$

(c) $\mathbb{E}[X]$

Soluzione:

$$\mathbb{E}[X] = 1\mathbb{P}[X=1]+2\mathbb{P}[X=2]+3\mathbb{P}[X=3] = (0.13+0.17)+2(0.09+0.17)+3(0.22+0.22) = 2.14$$

(d) $P(X \leq Y)$

Soluzione: $\mathbb{P}[X \leq Y]$ si ottiene sommando tutte le probabilità congiunte p_{ij} tali che $i \leq j$.

$$\mathbb{P}[X \leq Y] = 0.13 + 0.17 + 0.17 = 0.47$$

Domanda 5 (5 punti)

Sia $\{X_n\}$ una catena di Markov con spazio degli stati 1, 2, 3 e con le seguenti probabilità di transizione: $p_{11} = 30$, $p_{12} = 26$, $p_{21} = 71$, $p_{22} = 14$, $p_{31} = 33$, $p_{32} = 17$. La probabilità dello stato iniziale è data da $(1, 0, 0)$.

Si calcoli

(a) La matrice di transizione della catena

Soluzione: La matrice di transizione è

$$P = \begin{pmatrix} 0.3 & 0.26 & 0.44 \\ 0.71 & 0.14 & 0.15 \\ 0.33 & 0.17 & 0.50 \end{pmatrix}$$

(b) $\mathbb{P}[X_2 = `ri`]$

Soluzione: La matrice di transizione a due passi è

$$P^2 = PP = \begin{pmatrix} 0.4198 & 0.1892 & 0.3910 \\ 0.3619 & 0.2297 & 0.4084 \\ 0.3847 & 0.1946 & 0.4207 \end{pmatrix}$$

Se il vettore delle probabilità iniziali è $\pi^{(0)} = (1, 0, 0)$, la distribuzione marginale a due passi è $\pi^{(2)} = \pi^{(0)} P^2 = (0.4198, 0.1892, 0.3910)$. Dunque, $\mathbb{P}[X_2 = 1] = 0.4198$.

(c) $\mathbb{P}[X_2 = 1, X_3 = 3]$

Soluzione: La distribuzione congiunta di X_2 e X_3 si trova moltiplicando la marginale di X_2 per la condizionata di $X_3|X_2$:

$$\mathbb{P}[X_2 = 1, X_3 = 3] = \mathbb{P}[X_2 = 1]\mathbb{P}[X_3 = 3|X_2 = 1] = \mathbb{P}[X_2 = 1]p_{1,3} = 0.4198 \cdot 0.44 = 0.1847$$

Domanda 6 (5 punti)

(a) Si illustri brevemente il metodo di accettazione/rifiuto per la simulazione di valori casuali.

Soluzione: La risposta corretta non è unica...

(b) Si scrivano alcuni comandi di R utili per simulare 100 valori casuali da una distribuzione con funzione di densità $f(x) = 4x^3$ nell'intervallo $(0, 1)$, utilizzando il metodo di accettazione/rifiuto.

Soluzione: Si possono simulare valori casuali da una densità uniforme in $(0, 1)$, $g(x)$. La costante k tale che $f(x) \leq kg(x)$ è $k = 4$. Il codice R per la simulazione è:

```
x<-runif(10000)
y<-runif(10000,0,4)
x<-x[y <  4*x^3] # Si spera di accettare almeno 100
x<-x[1:100]
```