Probabilità e Statistica [CT0111]
Corso di Laurea triennale in Informatica, a.a. 2022/23

Isadora Antoniano Villalobos

Esame A **Soluzioni**, 10 febbraio 2023

Cognome: ______________________ Nome: ______________________

Matricola: ______________________ Firma: ______________________

ISTRUZIONI: DA LEGGERE CON ATTENZIONE!

Questo compito è composto di **5 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	Total
Points:	5	9	6	4	4	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) Quale delle seguenti è una funzione di densità per una variabile casuale continua?

i) Tutte.

ii) Nessuna.

iii)
$$f(x) = e^{-\frac{y^2}{2}}\,dy; \quad x \in (-\infty, \infty).$$

iv)
$$f(x) = \begin{cases} \frac{2}{3}, & -0.75 \leq x \leq 0.75 \\ 0, & \text{altrimenti.} \end{cases}$$

v)
$$f(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-6x}, & x \geq 0 \end{cases}$$

Soluzione: *iv)*

(b) Se X e Y sono due variabili casuai continue con funzione di densità congiunta $f_{X,Y}(x, y) = g(x)h(y)$. Quale delle seguenti affermazioni è sicuramente vera?

i) X e Y sono indipendenti con densità marginali $f_X(x) = g(x)$ e $f_Y(y) = h(y)$

ii) X e Y sono indipendenti ma $g(x)$ e $h(y)$ potrebbero non essere le densità marginali.

iii) $f_X(x) = g(x)$ e $f_Y(y) = h(y)$ sono le densità marginali ma X e Y potrebbero non essere indipendenti

iv) $Y = f_{X,Y}(x, y)/g(x)$

v) Nessuna delle precedenti.

Soluzione: *i)*

(c) Se A e B formano una partizione dello spazio campionario, quale delle seguenti affermazioni è sicuramente vera?

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

ii) $\mathbb{P}[A] \geq 1/2$

iii) $\mathbb{P}[B] \geq 1/2$

iv) $\mathbb{P}[A] \geq 1/2$ o $\mathbb{P}[B] \geq 1/2$

v) Nessuna delle precedenti.

Soluzione: *iv)*

(d) Per una variabile casuale Y con possibili valori nell'intervallo $(0, 4)$ quale delle seguenti affermazioni è sicuramente vera?

i) $\text{Var}[Y] = 4/3$

ii) $\mathbb{E}[Y] = 2$

iii) $\text{Var}[Y] \geq 0$ ma non si può sapere nulla sul valore di $\mathbb{E}[Y]$
iv) $\mathbb{E}[Y] \geq 0$ e $\text{Var}[Y] \geq 0$
v) $\mathbb{E}[Y] > 0$ ma non si può sapere nulla sul valore di $\text{Var}[Y]$

Soluzione: *iv)*

(e) Un'urna contiene 20 palline tra cui 5 bianche e 10 nere. Si estraggono 3 palline con reinserimento. Qual è la probabilità che almeno 2 delle palline estratte non siano ne bianche ne nere?

i) `phyper(1,5,10,3)`
ii) `1-pbinom(1,3,5/20)`
iii) `1-phyper(2,5,15,3)`
iv) `1-pbinom(1,3,5/10)`
v) `1-phyper(1,5,10,3)`

Soluzione: *ii)*

Domanda 2 (9 punti)

I componenti elettronici di un modello e di una marca specifici vengono spediti a un fornitore in lotti da dieci. 70% dei lotti non contengono componenti difettosi, 20% ne contengono essattamente uno e il restante 10% contiene due componenti difettosi. Si sceglie a caso un lotto e due componenti dello stesso vengono selezionati in modo casuale e testati. Si considerino le variabili $X =$ *Numero di componenti difettosi nel lotto selezionato* e $Y =$ *Numero di componenti difettosi testati.*

(a) Qual è la distribuzione marginale di X?

Soluzione:

x	0	1	2
$p_X(x)$	0.7	0.2	0.1

(b) Qual è la distribuzione condizionata di Y dato che il lotto selezionato ha due componenti difettosi?

Soluzione: Se il lotto ha due componenti difettosi, allora il numero di componenti difettosi selezionati può essere 0, 1 o 2. Infatti, dato $X = 2$, Y è una variabile ipergeometrica: $Y|X = 2\text{Ig}(N = 10, K = 2, n = 2)$. Quindi, la distribuzione richiesta è:

y	$p_{Y!X}(y\|2)$
0	$\frac{\binom{2}{0}\binom{8}{2}}{\binom{10}{2}} = \frac{1\cdot 28}{45} = 0.6222$
1	$\frac{\binom{2}{1}\binom{8}{1}}{\binom{10}{2}} = \frac{2\cdot 8}{45} = 0.3556$
2	$\frac{\binom{2}{2}\binom{8}{0}}{\binom{10}{2}} = \frac{1\cdot 1}{45} = 0.0222$

(c) Qual è la distribuzione marginale di Y?

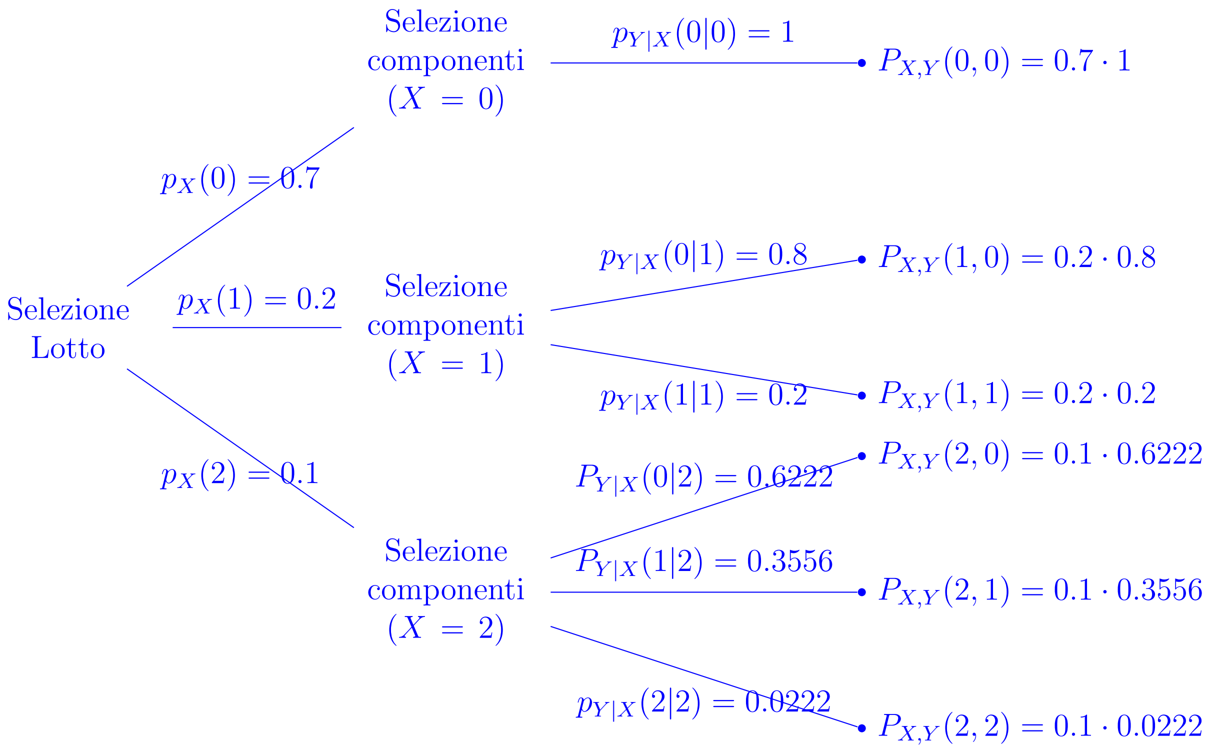
Soluzione: Consideriamo prima che, se $X = 0$ allora $Y = 0$ con probabilità 1.

Se $X = 1$, allora $Y = 0$ con probabilità $\binom{9}{2}/\binom{10}{2} = 36/45 = 0.8$ e $Y = 1$ con probabilità $1 - 0.8 = 0.2$.

Quindi, la distribuzione marginale di Y è

y	$p_Y(y)$
0	$0.7 \cdot 1 + 0.2 \cdot 0.8 + 0.1 \cdot 0.6222 = 0.9222$
1	$0.2 \cdot 0.2 + 0.1 \cdot 0.3556 = 0.0756$
2	$0.1 \cdot 0.0222 = 0.0022$

L'informazione fornita dall'esercizio si può rappresentare nel seguente diagramma:



(d) Le due variabili sono indipendenti? Si giustifichi adeguatamente la risposta.

Soluzione: I possibili valori di Y dipendono del valore di X, quindi le variabili non sono indipendenti.

(e) Se nessuno dei componenti testati risulta difettoso, qual è la probabilità che non ci siano componenti difettosi nel lotto?

Soluzione: Dal teorema di Bayes abbiamo:

$$\mathbb{P}[X = 0|Y = 0] = \frac{\mathbb{P}[Y = 0|X = 0]\mathbb{P}[X = 0]}{\mathbb{P}[Y = 0]} = \frac{1 \cdot 0.7}{0.9222} = 0.7590$$

Domanda 3 (6 punti)

Si consideri una variabile casuale R con funzione di densità

$$f_R(r) = \begin{cases} 0, & x < 0 \\ Ce^{-5x}, & x \geq 0 \end{cases}$$

(a) Si trovi il valore di C.

Soluzione: Possiamo riconoscere f_R come la funzione di densità di una variabile esponenziale di parametro $\lambda = 5$, quindi $C = \lambda = 5$.

(b) Si calcoli la media e la varianza di R

Soluzione: $R \sim$Exp(5), quindi

$$\mathbb{E}[R] = 1/5 = 0.2; \quad \text{Var}[R] = 1/5^2 = 0.04$$

(c) Si trovi il valore atteso dell'area, A, del cerchio di raggio R

Soluzione: L'area del cerchio di raggio R è $A = \pi R^2$, quindi $\mathbb{E}[A] = \mathbb{E}[\pi R^2] = \pi \mathbb{E}[R^2]$
Sappiamo che $\text{Var}[R] = \mathbb{E}[R^2] - \mathbb{E}[R]^2$, quindi

$$\mathbb{E}[R^2] = \text{Var}[R] + \mathbb{E}[R]^2 = \frac{1}{\lambda^2} + \left(\frac{1}{\lambda}\right)^2 = \frac{2}{\lambda^2} = \frac{2}{25} = 0.08$$

Finalmente,

$$\mathbb{E}[A] = \frac{2\pi}{25} = 0.08\pi = 0.2513$$

(d) Sia P il perimetro del cerchio di raggio R. Si trovi il valore atteso e la varianza della variabile $X = P/R$

Soluzione: Il perimetro del cerchio è $P = 2\pi R$, quindi $X = 2\pi R/R = 2\pi$ è una costante, con $\mathbb{E}[X] = 2\pi = 6.2832$ e $\text{Var}[X] = 0$.

Domanda 4 (6 punti)

Ogni giorno Alessandro percorre la stessa strada per andare da casa all'università. Ci sono 4 semafori lungo la strada, ed Alessandro ha notato che se vede un semaforo verde a un incrocio, il 60% delle volte anche il semaforo successivo è verde, e il 40% delle volte il semaforo successivo è rosso. Tuttavia, se vede un semaforo rosso, il 70% delle volte anche quello successivo è rosso e il 30% delle volte è verde.

(a) Si determini la matrice di transizione per la catena di Markov che rappresenta i colori dei semafori, specificando chiaramente gli stati della catena.

Soluzione: Sia $X(n)$ la catena di Markov con stati $X(n) = 1$ se l'ennesimo semaforo è verde e $X(n) = 2$ se è rosso. Allora, la matrice di transizione per il processo è:

$$P = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}$$

(b) Se il primo semaforo è verde, qual è la probabilità che il terzo sia rosso?

Soluzione: La matrice di transizione a due passi è

$$P^2 = PP = \begin{pmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{pmatrix}$$

Quindi, la probabilità richiesta è

$$\mathbb{P}[X(3) = 2 | X(1) = 1] = p_{1,2}^{(2)} = 0.52$$

(c) Tommaso, il compagno di classe di Alessandro, ha tantissimi semafori lungro il percorso da casa sua all'università. Se il primo semaforo è verde, qual è la probabilità che l'ultimo sia rosso?

Hint: Utilizzare la distribuzone stazionaria

Soluzione: La distribuzione stazionaria della catena deve soddisfare la condizione $\pi = \pi P$. Aggiungendo la condizione $\sum_i \pi_i = 1$, si ottiene un sistema lineare con due equazioni e due incognite:

$$\begin{cases} 0.6\pi_1 + 0.3\pi_2 = \pi_1 \\ \pi_1 + \pi_2 = 1 \end{cases}$$

Il sistema ha un'unica soluzione: $\pi = (3/7, 4/7) = (0.4286, 0.5714)$. Dopo tantissimi semafori, il colore di quello iniziale non è più rilevante e la probabilità che l'ultimo sia rosso è $\pi_2 = 4/7 = 0.5714$

Domanda 5 (4 punti)

Si spieghi la proprietà Markoviana e si fornisca un esempio, giustificato, di una situazione per la quale può risultare utile come modello.

Soluzione: La risposta corretta non è unica... Si deve menzionare l'evoluzione temporale di una quantità casuale per la quale, dato il presente, il futuro risulta indipendente del passato.