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

Isadora Antoniano Villalobos

Esame **Soluzioni**, 21 gennaio 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	4	5	6	6	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 funzioni non è una funzione di densità?

i) Sono tutte funzioni di densità.

ii) $f(x) = \frac{1}{\sqrt{8\pi}} e^{-\frac{(x-6)^2}{8}}, \quad x \in (-\infty, \infty)$.

iii) $f(x) = e^{-6} \frac{6^x}{x!} \quad x = 0, 1, 2, \ldots$.

iv) $f(x) = 4e^{-4x}, \quad x > 0$.

v) $\frac{1}{6^2}|x| \quad x \in (-6, 6)$.

Soluzione: *iii)*

(b) Sia X una variabile aleatoria con valore atteso 3 e varianza 5 e sia $Y = -2X + 1$ una sua trasformazione lineare. Quale delle seguenti affermazioni è sicuramente vera

i) $\mathbb{E}[Y] = -6$

ii) $\mathrm{Var}[Y] < \mathrm{Var}[X]$

iii) $Y < X$

iv) $\mathrm{Var}[Y] = 4\ \mathrm{Var}[X]$

v) $\mathrm{Var}[Y] = 21$

Soluzione: *iv)*

(c) Se A e B sono due eventi indipendenti, quale delle seguenti affermazioni è sicuramente vera?

i) $A \cup B = \emptyset$

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

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

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

v) $\mathrm{Cov}[A, B] = 0$

Soluzione: *v)*

(d) Per una variabile casuale discreta Y uniforme con valori $\{`r-41`, 0, `r41`\}$ quale delle seguenti espressioni è vera?

i) Il valore atteso $E(Y)$ è pari a -0.5.

ii) Il valore atteso del quadrato di Y, $E(Y^2)$ è pari a -0.5.

iii) La differenza $E(Y^2) - E(Y)$ vale 0.

iv) $P(Y > -41) = 2/3$.

v) $P(Y > 41) = 1/3$.

Soluzione: *iv)*

(e) Un'urna contiene 23 palline di cui 7 bianche e 16 nere. Si estraggono in blocco 4 palline dall'urna. Qual è la probabilità che fra le palline estratte ve ne siano al massimo 2 bianche?

i) `phyper(2,7,16,4)`
ii) `pbinom(2,4,7/23)`
iii) `ppois(2,4*7/23)`
iv) `dbinom(2,4,7/23)`
v) `dhyper(2,7,16,4)`

Soluzione: *i)*

Domanda 2 (4 punti)
La durata in ore delle pile prodotte da una ditta segue una distribuzione esponenziale di media 800 ore. Si approssimi la probabilità che la durata media di un campione di $n = 200$ pile sia maggiore di 750 ore. (è possibile sostituire il calcolo con un opportuno comando di R)

Soluzione: Sia $\bar{X}_{200}$ la media campionaria su un campione di 200 pile. Allora, per il teorema limite centrale, possiamo approssimare la distribuzione di $\bar{X}_{200}$ con una normale $N(\mu = 800, \sigma = 800/\sqrt{200})$. La probabilità richiesta si approssima con

$$P\left(\bar{X}_{200} > 750\right) \approx$$
$$\approx P\left(Z > \sqrt{2}\frac{750 - 800}{80}\right) = P\left(Z > -0.884\right) =$$
$$= P\left(Z \leq 0.884\right) = 0.812.$$

La probabilità richiesta si può anche ottenere in R con il comando:

`1-pnorm(750,800,80/sqrt(2))`.

Domanda 3 (5 punti)
Il 22% dei microchip utilizzati da un'azienda di assemblaggio di computer è prodotto dal Fornitore A e il restante 78% dal Fornitore B. Il peso, (in grammi) dei microchip prodotti dal Fornitore A segue una distribuzione normale di media $\mu = 3.1$ e scarto quadratico medio $\sigma = 0.23$, ed è compreso tra 2.8 e 3.2 con probabilità 0.5721. Il peso dei microchip prodotti dal Fornitore B è distribuito uniformemente tra 2.8g e 3.3g.

(a) Qual è la probabilità che un microchip prodotto del Fornitore B abbia un peso compreso tra 2.8 e 3.2 grammi?

Soluzione: Si definiscano gli eventi:
$G = \{$Il peso del microchip è compreso tra 2.8 e 3.2 grammi$\}$
$A = \{$Il microchip è stato prodotto dal fornitore A$\}$
$B = \{$Il microchip è stato prodotto dal fornitore B$\}$
Sia Y il peso di un microchip prodotto dal fornitore B. Quindi, $Y \sim \mathrm{U}(2.8, 3.3)$, e

$$\mathbb{P}[G|B] = \mathbb{P}[2.8 < Y < 3.2] = \frac{3.2 - 2.8}{3.3 - 2.8} = 0.8$$

(b) Se un microchip utilizzato dall'azienda di assemblaggio ha un peso compreso tra 2.8g. e 3.2g, qual è la probabilità che provenga dal fornitore B?

Soluzione: Per il Teorema di Bayes si ha

$$\mathbb{P}[B|G] = \frac{\mathbb{P}[B]P[G|B]}{\mathbb{P}[G]} = \frac{\mathbb{P}[B]P[G|B]}{\mathbb{P}[A]P[G|A] + \mathbb{P}[B]P[G|B]} = \frac{0.78 \cdot 0.8}{0.22 \cdot 0.5721 + 0.78 \cdot 0.8}$$
$$= \frac{0.624}{0.7499} = 0.8321$$

Domanda 4 (6 punti)

Siano X e Y due variabili casuali discrete con:

x	$p_X(x)$
0	0.2
1	0.5
2	0.3

	$y = 1$	$y = 2$
$p_{Y\|X}(y\|0)$	0.5	0.5
$p_{Y\|X}(y\|1)$	0.1	0.9
$p_{Y\|X}(y\|2)$	0.4	0.6

(a) Si calcoli la funzione di probabilità congiunta di X e Y.

Soluzione: Per ogni $i = 0, 1, 2$ e ogni $j = 1, 2$,

$$\mathbb{P}[X = i, Y = j] = \mathbb{P}[Y = j|X = i]\mathbb{P}[X = i].$$

Ad esempio,

$$\mathbb{P}[X = 2, Y = 1] = \mathbb{P}[Y = 1|X = 2]\mathbb{P}[X = 1] = 0.4 \cdot 0.3 = 0.12.$$

La distribuzione contiunta di X e Y è, dunque,

$X \setminus Y$	1	2
0	0.1	0.1
1	0.05	0.45
2	0.12	0.18

(b) Si calcoli la funzione di probabilità marginale di Y.

Soluzione: Le probabilità marginali della Y si trovano sommando i valori sulle colonne della distribuzione congiunta. Dunque,

y	1	2
$p_Y(y)$	0.27	0.73

(c) Trovare $\mathbb{P}[X < Y]$

Soluzione:

$$\mathbb{P}[X < Y] = \mathbb{P}[X = 0, Y = 1] + \mathbb{P}[X = 0, Y = 2] + \mathbb{P}[X = 1, Y = 2] = 0.65.$$

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

Soluzione: La distribuzione di $Y|X = x$ cambia dipendendo dal valore di x, dunque le due variabili non sono indipendenti.

Domanda 5 (6 punti)

In una città ogni giorno il tempo è soleggiato (1) o piovoso (2). Un giorno di sole è seguito da un altro giorno di sole con probabilità 0.2, mentre un giorno di pioggia è seguito da uno di sole con probabilità 0.6.

(a) Se lunedì c'è il sole, qual è la probabilità che ci sia il sole anche mercoledì?

Soluzione: La matrice di transizione è

$$P = \begin{pmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{pmatrix}$$

e la matrice di transizione a due passi è

$$P^2 = PP = \begin{pmatrix} 0.52 & 0.48 \\ 0.36 & 0.64 \end{pmatrix}.$$

Quindi, la probabilità richiesta è

$$\mathbb{P}[X(2) = 1|X(0) = 1] = p_{1,1}^{(2)} = 0.52$$

(b) Se lunedì c'è il sole, qual è la probabilità che ci sia il sole sia martedì che mercoledì?

Soluzione: La probabilità richiesta è

$$\begin{aligned}\mathbb{P}[X(2) = 1, X(1) = 1|X(0) = 1] &= \mathbb{P}[X(2) = 1|X(1) = 1]\mathbb{P}[X(1) = 1|X(0) = 1] \\ &= p_{1,1}p_{1,1} = (0.2)^2 = 0.04\end{aligned}$$

(c) Si calcoli una distribuzione stazionaria della catena.

Soluzione: Le distribuzioni stazionarie della catena soddisfano la condizione $\pi = \pi P$. Aggiungendo la condizione $\sum_i \pi_i = 1$, si ottiene un sistema lineare con due equazioni e due incognite:

$$\begin{cases} 0.2\pi_1 + 0.6\pi_2 = \pi_1 \\ \pi_1 + \pi_2 = 1 \end{cases}$$

Il sistema ha un'unica soluzione: $\pi = (0.4286, 0.5714)$

Domanda 6 (4 punti)

(a) Cos'è una passeggiata aleatoria con barriere assorbenti?

Soluzione: La risposta corretta non è unica...

(b) Si scriva una funzione di `R` che simuli una traiettoria di lunghezza 500 della passeggiata aleatoria con probabilità di salire (+1) pari a 0.6 ad ogni passo, e che trovi il primo tempo in cui la traiettoria simulata tocca la barriera 30.

Soluzione: Il codice R per la simulazione è:

```
x <- 2*rbinom(500,1,0.6)-1
y <- cumsum(x)
```

```
t <- min(which(y==30))
```

Il vettore `y` contiene gli stati visitati dalla catena (la traiettoria) e il valore `t` corrisponde al primo tempo in cui la traiettoria simulata tocca la barriera 30.