Informazione e stima – 11/01/2022

- Ogni esercizio viene valutato da 0 a 6 punti.
- Verranno valutate solo le parti scritte in penna – non usare correttori.
- Non riportare solo il risultato, ma cercare di argomentare sinteticamente la risposta.
- Riportare il proprio nome, cognome, e codice persona su ogni foglio consegnato.
- Esercizio da escludere dal punteggio finale:

1. Si estraggono 5 carte da un mazzo di 52 ben mescolato. Si considerino gli eventi $A = \{$solo le prime due carte sono di cuori$\}$, $B = \{$prime due carte di cuori e le altre di fiori$\}$, e $C = \{$prime due di cuori, terza e quarta di fiori, e ultima di quadri o picche$\}$.

   (a) Senza fare calcoli, dire quale tra A e B è l'evento più probabile. Giustificare la risposta.

   (b) Calcolare le probabilità degli eventi A, B, e C.

   Le estrazioni sono senza reinserimento. Nel mazzo ci sono 13 *carte di ogni seme.*

2. Sia X la differenza tra il numero di teste e il numero di croci ottenute in due lanci di una moneta ben bilanciata. Si calcolino $\mathsf{E}[X]$ e $\mathsf{E}\left[X^2\right]$. Come cambierebbero i valori di $\mathsf{E}[X]$ e $\mathsf{E}\left[X^2\right]$ se si facessero 100 lanci?

3. Sia $X \sim \mathrm{Exp}(\lambda)$ e $Y = \arctan(X)$. Si calcoli la legge di probabilità di Y.
   Si presti attenzione ai valori assunti da Y.

4. Si consideri un processo di Poisson con intensità di 10 arrivi al minuto. Usando il teorema fondamentale del limite, dare un'approssimazione della probabilità che ci siano meno di 550 arrivi in un'ora.

5. Si consideri un esperimento aleatorio e un suo evento A di cui si vuole stimare la probabilità. Per stimare $p = \Pr(A)$ si ripete l'esperimento n volte, ottenendo una stima $\widehat{P}_n$ tramite media campionaria.

   (a) In quale caso bisogna usare un numero di prove n elevato? Quando A è un evento raro o frequente? Giustificare la risposta.

   (b) Quale tecnica si può usare per ridurre il valore di n senza peggiorare la qualità della stima?

6. Si consideri un mazzo <u>non</u> mescolato di 52 carte, vale a dire un mazzo di cui conoscete l'ordine esatto di tutte le carte. Chiedete ad un'altra persona di *tagliare* il mazzo in un punto a caso creando due mazzetti, e di ricomporlo invertendo l'ordine dei due mazzetti. Cominciate a svelare una carta alla volta, fino a svelarle tutte. Qual è la quantità totale di bit di informazione a voi rivelata?

Soluzioni

Problema 1

1. Siccome $A \supset B$, abbiamo $\Pr(A) > \Pr(B)$.
2. Possiamo risolvere tutti i casi ragionando su cosa accade in ogni stadio di scelta. Si hanno 5 stadi di scelta, e in ogni stadio rapportiamo casi favorevoli su casi totali.

$$\Pr(A) = \frac{13}{52}\frac{12}{51}\frac{39}{50}\frac{38}{49}\frac{37}{48} \approx 0.0274 \tag{1}$$

$$\Pr(B) = \frac{13}{52}\frac{12}{51}\frac{13}{50}\frac{12}{49}\frac{11}{48} \approx 0.000858 \tag{2}$$

$$\Pr(C) = \frac{13}{52}\frac{12}{51}\frac{13}{50}\frac{12}{49}\frac{26}{48} \approx 0.00202 \tag{3}$$

Problema 2

X vale 2, 0 oppure -2 con probabilità $\frac{1}{4}$, $\frac{1}{2}$ e $\frac{1}{4}$, rispettivamente. Dunque

$$\mathsf{E}[X] = 2\frac{1}{4} - 2\frac{1}{4} = 0$$

e

$$\mathsf{E}[X^2] = 4\frac{1}{4} + 4\frac{1}{4} = 2.$$

100 lanci equivalgono a 50 coppie di lanci (indipendenti) e quindi valore medio e varianza sono moltiplicati per 50.

Problema 3

Applichiamo il metodo della cumulata:

$$\begin{aligned} F_Y(y) = \Pr(Y \leq y) &= \Pr(\arctan(X) \leq y) && (4)\\ &= \Pr(X \leq \tan(y)) && (5)\\ &= F_X(\tan(y)) && (6)\\ &= \begin{cases} 1 & y \geq \pi/2 \\ 1 - e^{-\lambda \tan(y)}, & 0 \leq y < \pi/2 \\ 0 & y < 0 \end{cases} && (7) \end{aligned}$$

dove abbiamo usato $F_X(x) = 1 - e^{-\lambda x}$, e il fatto che Y assume valori tra $\arctan(0) = 0$ e $\arctan(+\infty) = \pi/2$.

La densità di probabilità si ottiene derivando:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{\lambda e^{-\lambda \tan(y)}}{\cos^2(y)}, \qquad 0 < y < \frac{\pi}{2} \tag{8}$$

e $f_Y(y) = 0$ altrove.

Problema 4

Sia Y_{550} il tempo al 550-esimo arrivo misurato in minuti. Il testo chiede di valutare approssimativamente la $\Pr(Y_{550} > 60)$. Innanzitutto si può notare che $Y_{550} = \sum_{i=1}^{550} T_i$, dove T_i sono i tempi di interarrivo del processo di Poisson, con $T_i \overset{iid}{\sim} \text{Exp}(10)$.

Siccome Y_{550} è la somma di un grande numero di v.a. iid, si può applicare il teorema fondamentale del limite come segue:

$$\begin{aligned} \Pr(Y_{550} > 60) &= \Pr\left(\frac{Y_{550} - \mathsf{E}[Y_{550}]}{\sqrt{\mathsf{Var}[Y_{550}]}} > \frac{60 - \mathsf{E}[Y_{550}]}{\sqrt{\mathsf{Var}[Y_{550}]}}\right) && (9)\\ &\overset{\text{CLT}}{\approx} \Pr\left(Z > \frac{60 - 550 \cdot \frac{1}{10}}{\sqrt{550 \cdot \frac{1}{100}}}\right) && (10)\\ &= \Pr(Z > 2.132) && (11)\\ &= 1 - \Phi(2.132) \approx 0.0165, && (12) \end{aligned}$$

dove abbiamo usato $\mathsf{E}[T_i] = \frac{1}{10}$ e $\mathsf{Var}[T_i] = \frac{1}{100}$.

Problema 5

Vedere la teoria.

Problema 6

Il fatto che il mazzo di carte sia inizialmente ordinato significa che non vi porta alcuna informazione. L'atto di tagliare il mazzo in un punto a caso è equivalente a scegliere una carta uniformemente a caso tra le 52. Infatti, una volta svelata la prima carta del nuovo mazzo, che è a voi ignota, tutte le altre carte sono prevedibili.

Dunque, l'unica fonte di informazione (o sorpresa) è la prima carta del nuovo mazzo, che è il risultato di una v.a. uniforme da 1 a 52. Il numero di bit di informazione così generato è

$$H(X) = \log_2(52) \approx 5.7 \text{ bit.} \tag{13}$$