Teoria dei fenomeni aleatori e della stima – 31/08/2018

- 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 cerca di argomentare sinteticamente la risposta.
- Riportare il proprio nome, cognome, numero di matricola su ogni foglio consegnato
- Esercizio da escludere dal punteggio finale:

(1) La banca vi consegna la nuova tessera bancomat con il relativo codice PIN. Il codice ha 5 cifre: tutti i numeri da 00000 a 99999 sono possibili, e vengono assegnati a caso con legge di probabilità uniforme. Qual è la probabilità che vi sia stato assegnato un codice PIN contenente esattamente tre cifre 0?

(2) Vi viene proposto il seguente gioco: si compie una successione di lanci di una moneta bilanciata, dove effettuare ogni lancio vi costa 1 €. Se si ottengono 3 teste di fila allora si vincono 10 € e il gioco termina. Mediamente vi aspettate di guadagnare o perdere da questo gioco?

Suggerimento: aiutarsi con un diagramma ad albero per calcolare i lanci attesi da effettuare per osservare l'evento di interesse, e impostare il teorema dell'aspettativa totale.

(3) Sia X una v.a. continua con legge $f_X(x) = c/x^4$ per $x \geq 1$ e $f_X(x) = 0$ altrimenti. Determinare

(a) La costante c.

(b) La legge di probabilità cumulata.

(c) Quali momenti $\mathsf{E}[X^n]$ di X esistono, per $n = 1, 2, \cdots$. La varianza di X esiste?

(4) Una mattina dovete recarvi in banca: sapete che dalle ore 09:00 in poi i clienti cominceranno ad arrivare secondo un processo di Poisson con tasso $\lambda = 10$ clienti/ora. Decidete di partire da casa alle ore 09:00 per recarvi in banca, e sapete che il tempo per raggiungerla sarà di $X = 20 + Y$ minuti, dove $Y \sim \text{Exp}(2)$. Quanti clienti saranno già arrivati, in media, al momento del vostro arrivo? La media dei clienti arrivati va intesa su tutte le quantità aleatorie.

Suggerimento: l'intervallo di tempo X può essere spezzato in due parti, e la media dei clienti arrivati nell'intervallo somma è...

(5) Si consideri un parametro aleatorio ignoto $\Theta \sim \text{Exp}(\lambda)$ che viene osservato tramite la misura $X = \Theta + W$, dove $W \sim \mathcal{N}(0, 1)$ è indipendente da Θ. Determinare lo stimatore MAP di Θ basato sull'osservazione X.

(6) Avendo a disposizione un generatore di campioni indipendenti U_i distribuiti uniformemente in $[0, 1)$, proporre un algoritmo per simulare l'estrazione senza reinserzione di 3 carte da un mazzo ben mescolato di 40 carte.

Soluzioni

Problema 1

Siccome lo spazio di probabilità è uniforme, calcoliamo il rapporto tra casi favorevoli e casi totali:

- I casi totali sono 10^5
- Per contare i casi favorevoli dobbiamo innanzitutto scegliere le 3 posizioni che contengono gli zeri, tramite un coefficiente binomiale, e poi contare tutti i modi per riempire le 2 posizioni rimanenti. In tutto si ha $\binom{5}{3}9^2$, dove il 9 è dovuto al fatto che non possiamo scegliere la cifra 0 per le posizioni rimanenti.

La probabilità cercata è

$$\frac{\binom{5}{3}9^2}{10^5} \approx 8.1 \cdot 10^{-3}.$$

In alternativa si poteva usare una v.a. Binomiale per contare il numero di zeri nel codice.

Problema 2

Sia N il numero di lanci effettuati per ottenere 3 teste di fila, e si usino C e T per gli eventi croce e testa, rispettivamente. Siccome siamo interessati al numero medio di lanci, calcoliamo:

$$\begin{aligned}\mathsf{E}[N] &= 1 + \frac{1}{2}\mathsf{E}[N|C] + \frac{1}{2}\mathsf{E}[N|T] \\ &= 1 + \frac{1}{2}\mathsf{E}[N] + \frac{1}{2}\mathsf{E}[N|T] \qquad (1)\end{aligned}$$

$$\begin{aligned}\mathsf{E}[N|T] &= 1 + \frac{1}{2}\mathsf{E}[N|TC] + \frac{1}{2}\mathsf{E}[N|TT] \\ &= 1 + \frac{1}{2}\mathsf{E}[N] + \frac{1}{2}\mathsf{E}[N|TT] \qquad (2)\end{aligned}$$

$$\begin{aligned}\mathsf{E}[N|TT] &= 1 + \frac{1}{2}\mathsf{E}[N|TTC] + \frac{1}{2}\mathsf{E}[N|TTT] \\ &= 1 + \frac{1}{2}\mathsf{E}[N]. \qquad (3)\end{aligned}$$

Sostituendo la (3) nella (2), e il risultato di questa nella (1), si ottiene un'equazione nell'incognita $\mathsf{E}[N]$. Risolvendo l'equazione si trova $\mathsf{E}[N] = 14$. Pertanto mediamente si spendono 14 € per giocare, e si vincono solo 10 €, quindi giocare non conviene in media.

Problema 3

- La costante c deve essere tale che

  $$\int_1^\infty f_X(x)dx \overset{!}{=} 1$$

  $$\int_1^\infty \frac{c}{x^4}dx = c\left[-\frac{1}{3x^3}\right]_1^\infty = \frac{c}{3} \overset{!}{=} 1$$

  dunque $c \overset{!}{=} 3$.

- La legge di probabilità cumulata si trova come

  $$\begin{aligned}F_X(x) = \Pr(X \leq x) &= \int_1^x f_X(t)dt \\ &= \int_1^x \frac{3}{t^4}dt = \left[-\frac{1}{t^3}\right]_1^x = 1 - \frac{1}{x^3} \qquad x \geq 1\end{aligned}$$

  e $F_X(x) = 0$ per $x < 1$.

- Il momento n-esimo di X si trova come

  $$\mathsf{E}[X^n] = \int_1^\infty \frac{3}{x^4}x^n dx.$$

  L'integrale esiste solo per $n = 1$ e $n = 2$. Siccome la varianza è $\mathsf{Var}[X] = \mathsf{E}[X^2] - \mathsf{E}[X]^2$, essa esiste.

Problema 4

Bisogna determinare quanti clienti in media arrivano durante il tempo X. Il tempo X si può dividere in una parte deterministica (i 20 minuti) e una parte aleatoria (Y minuti):

- Durante i primi 20 minuti il numero di clienti arrivati sarà $N_1 \sim \text{Poisson}(10/60 \cdot 20)$, e quindi in media ne arriveranno $\mathsf{E}[N_1] = 10/3$.
- Durante l'intervallo di tempo Y arriveranno N_2 clienti, dove per la legge delle aspettazioni iterate abbiamo

$$\begin{aligned}
\mathsf{E}[N_2] &= \mathsf{E}[\mathsf{E}[N_2|Y]] && (4)\\
&= \mathsf{E}[10/60 \cdot Y] && (5)\\
&= 10/60 \cdot 1/2 = 1/12 && (6)
\end{aligned}$$

e nel passaggio (5) abbiamo sfruttato il fatto che $N_2|\{Y = y\}$ è una variabile aleatoria di Poisson di parametro $10/60 \cdot y$. Dunque in media arriveranno $\mathsf{E}[N] = \mathsf{E}[N_1] + \mathsf{E}[N_2] = 20 + 1/12$ clienti prima del vostro arrivo.

Problema 5

Seguendo la definizione di stimatore MAP si ha:

$$\begin{aligned}
\widehat{\Theta}_{\text{MAP}}(x) &= \arg\max_{\theta>0}\ f_{\Theta|X}(\theta|x) && (7)\\
&= \arg\max_{\theta>0}\ f_{X|\Theta}(x|\theta) f_\Theta(\theta) && (8)\\
&= \arg\max_{\theta>0}\ f_W(x-\theta) f_\Theta(\theta) && (9)\\
&= \arg\max_{\theta>0}\ \exp\left(-\frac{(x-\theta)^2}{2}\right)\exp(-\lambda\theta) && (10)\\
&= \arg\min_{\theta>0}\ \frac{(x-\theta)^2}{2} + \lambda\theta && (11)
\end{aligned}$$

dove in (8) e (10) abbiamo tralasciato i termini non dipendenti da θ, in (9) abbiamo sfruttato $X = \Theta + W$ e l'indipendenza di Θ e W, mentre in (11) abbiamo applicato il logaritmo e cambiato di segno. Per trovare il minimo della (11) rispetto a θ, poniamone a zero la derivata rispetto a θ:

$$\frac{\partial}{\partial\theta}\left[\frac{(x-\theta)^2}{2} + \lambda\theta\right] \stackrel{!}{=} 0 \quad \longrightarrow \quad \theta \stackrel{!}{=} x - \lambda.$$

Bisogna però prestare attenzione all'intervallo in cui si cerca il minimo, cioè $\theta > 0$. Pertanto, la soluzione sarà

$$\widehat{\Theta}_{\text{MAP}}(X) = \begin{cases} X - \lambda & X > \lambda, \\ 0 & X \leq \lambda. \end{cases} \tag{12}$$

Problema 6

Il problema si può dividere in 3 fasi. All'inizio è come se si volesse campionare da una distribuzione discreta e uniforme su 40 risultati possibili; nella seconda fase la distribuzione diventa uniforme su 39 risultati possibili, e si ha probabilità 0 in corrispondenza della carta pescata nella fase 1; nella terza fase la distribuzione da cui campionare ha solo 38 risultati con la stessa probabilità, e probabilità 0 in corrispondenza delle carte pescate nelle prime 2 fasi. Per simulare un'estrazione dalle distribuzioni della seconda e terza fase si può applicare un algoritmo simile ad acceptance/rejection.

1. Genero $U_1 \sim [0,1)$ e la prima carta pescata è $X_1 = \lceil 40 \cdot U_1 \rceil$, dove $\lceil x \rceil$ è l'intero più piccolo maggiore di x.
2. Genero $U_2 \sim [0,1)$ e pongo $T_2 = \lceil 40 \cdot U_2 \rceil$. Se $T_2 \neq X_1$ allora pongo $X_2 = T_2$, altrimenti torno al punto 2.
3. Genero $U_3 \sim [0,1)$ e pongo $T_3 = \lceil 40 \cdot U_3 \rceil$. Se $T_3 \neq X_1$ e $T_3 \neq X_2$ allora pongo $X_3 = T_3$, altrimenti torno al punto 3.