Teoria dei fenomeni aleatori e della stima – 12/02/2019

- 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, e numero di matricola su ogni foglio consegnato.

1. Si consideri l'alfabeto italiano di 21 lettere e tutte le parole (anche di senso non compiuto) che si possono formare con 5 lettere (anche ripetute).

   (a) Qual è la probabilità che una parola formata a caso sia palindroma?

   (b) Ipotizzando di avere formato una parola palindroma con 3 lettere diverse, qual è la probabilità che permutando a caso le lettere si ottenga una parola palindroma?

   Una parola è palindroma se non è distinguibile letta da sinistra a destra o viceversa. Ad es. AVEVA.

2. Una v.a. continua X ha ddp $f_X(x) = \alpha\sqrt{1-x^2}$ per $-1 < x < 1$ e $f_X(x) = 0$ altrove. Determinare:

   (a) il valore di α e il valore atteso $\mathsf{E}[X]$.

   (b) senza svolgere conti, quale v.a. tra $X \sim f_X$ e $Y \sim \mathcal{U}[-1, 1]$ ha varianza maggiore. Giustificare la risposta.

3. Un segnale X è composto dalla somma di 10 componenti casuali N_i indipendenti e distribuite identicamente come $\mathcal{N}(0, 1)$. Prima di essere trasmesso, il segnale X viene amplificato da un amplificatore che inserisce pesanti distorsioni se $X > 5$. Qual è la probabilità di osservare un segnale non distorto all'uscita dell'amplificatore?

4. Si consideri un mazzo ben mescolato di carte con 20 carte rosse e 20 carte nere, ed un esperimento dove si estraggono (con reinserzione) delle carte una dopo l'altra da tale mazzo. L'esperimento si conclude quando vengono estratte 3 carte consecutive dello stesso colore.

   (a) Proporre una catena di Markov che descriva l'evoluzione dell'esperimento. Individuare gli stati, le transizioni di stato e le probabilità di transizione.

5. Sia $X \sim \text{Exp}(1)$, da cui si genera un valore x per ottenere un'osservazione distribuita come $Y \sim \mathcal{N}(0, x)$. Trovare lo stimatore MAP di X basato sull'osservazione Y.

6. Si vuole determinare numericamente il valore di $I = \int_0^1 e^{-x^2/2}\, dx$. Partendo da un generatore di campioni distribuiti come $U \sim \mathcal{U}[0, 1]$, proporre un algoritmo che produca una stima $\widehat{I}$ di I.

Soluzioni

Problema 1

1. L'evento cercato è equivalente a imporre prima e ultima lettera della parola uguali, e seconda e quarta lettera uguali. La probabilità dell'evento è:

$$\frac{21 \cdot 21 \cdot 21 \cdot 1 \cdot 1}{21^5} = \frac{1}{21^2} \approx 0.0023$$

2. Partendo da una parola palindroma con 3 lettere diverse si possono formare solo due parole palidrome, perché la lettera centrale non può cambiare la sua posizione, mentre le due lettere esterne e le due interne possono scambiarsi di ruolo. Quindi i casi favorevoli sono $2 \cdot 2 \cdot 2$ (considerando anche le permutazioni ottenute scambiando tra loro le 2 coppie di lettere uguali), mentre i casi totali sono tutte le permutazioni con 5 lettere. La probabilità è dunque

$$\frac{2^3}{5!} \approx 0.0667$$

Problema 2

1. Il valore di α si ottiene imponendo l'integrale della ddp a 1, o più semplicemente si può notare che $f_X(x)$ descrive un semicerchio di raggio 1 e centro nell'origine. L'area sottesa dal semicerchio è $\pi/2$, dunque deve essere $\alpha = 2/\pi$.

   Il valore atteso di X è zero perché $f_X(x)$ ha simmetria pari.

2. Anche Y è una v.a. a valor atteso zero e con una $f_Y(y)$ con simmetria pari. Senza fare calcoli si può vedere subito che $\mathsf{Var}[Y] > \mathsf{Var}[X]$, perché f_Y è più dispersa attorno all'origine rispetto a f_X.

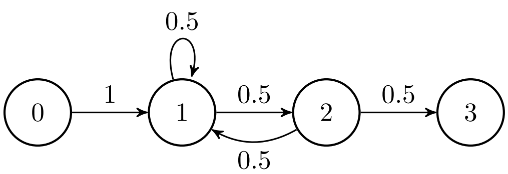
Problema 3

Siccome tutte le componenti sono Gaussiane indipedenti, anche X sarà Gaussiana con $\mathsf{E}[X] = 0$ e $\mathsf{Var}[X] = \sum_{i=1}^{10} \mathsf{Var}[N_i] = 10$. La probabilità si calcola come:

$$\Pr(X < 5) = \Pr\left(\frac{X}{\sqrt{10}} < \frac{5}{\sqrt{10}}\right) = \Pr\left(Z < \frac{5}{\sqrt{10}}\right) = \Phi(\frac{5}{\sqrt{10}}) \approx 0.9429.$$

Problema 4

1. Siccome le estrazioni sono con reinserimento, il numero di carte rosse e nere nel mazzo non cambia ad ogni estrazione. L'informazione che determina lo stato della catena di Markov è il numero di carte dello stesso colore consecutive che sono state osservate. All'inizio si può partire da uno stato '0' per poi usare gli stati $\{1, 2, 3\}$ che indicano il numero di carte consecutive dello stesso colore.



Problema 5

Seguendo la definizione di stimatore MAP si ha:

$$\begin{aligned}
\widehat{X}_{\mathrm{MAP}}(y) &= \arg\max_{x \in \mathbb{R}^+} \; f_{X|Y}(x|y) && (1)\\
&= \arg\max_{x \in \mathbb{R}^+} \; f_{Y|X}(y|x) f_X(x) && (2)\\
&= \arg\max_{x \in \mathbb{R}^+} \; \frac{1}{\sqrt{2\pi x}} e^{-y^2/(2x)} e^{-x} && (3)\\
&= \arg\max_{x \in \mathbb{R}^+} \; -\frac{1}{2}\log(x) - \frac{y^2}{2x} - x && (4)
\end{aligned}$$

dove in (2) abbiamo tralasciato i termini non dipendenti da x, e in (4) abbiamo applicato la trasformazione logaritmica che non cambia la posizione del massimo. Per trovare il minimo della (4) rispetto a x, poniamone a zero la derivata rispetto a x:

$$\frac{\partial}{\partial x}\left[-\frac{1}{2}\log(x)-\frac{y^2}{2x}-x\right]=-\frac{1}{2x}+\frac{y^2}{2x^2}-1\stackrel{!}{=}0 \quad\longrightarrow\quad x\stackrel{!}{=}\frac{\pm\sqrt{1+8y^2}-1}{4}.$$

Ci sono 2 punti estremanti, ma la soluzione negativa $x<0$ va esclusa. Pertanto, la soluzione è

$$\widehat{X}_{\mathrm{MAP}}(Y)=\frac{\sqrt{1+8Y^2}-1}{4}. \tag{5}$$

Problema 6

L'integrale I si può stimare numericamente ricorrendo ad una simulazione Monte Carlo. In particolare, si può notare che:

$$I=\int_0^1 e^{-x^2/2}dx=\int_0^1\frac{e^{-x^2/2}}{f_U(x)}f_U(x)dx=\mathsf{E}\left[\frac{e^{-U^2/2}}{f_U(U)}\right]=\mathsf{E}\left[e^{-U^2/2}\right]$$

dove $U\sim\mathcal{U}[0,1]$ e $f_U(u)=1$ per $0\le u\le 1$. L'algoritmo è fatto come segue:

1. Genero n campioni indipendenti $U_i\sim\mathcal{U}[0,1]$, per $i=1,\dots,n$.
2. Calcolo $\widehat{I}=\frac{1}{n}\sum_{i=1}^n e^{-U_i^2/2}$.