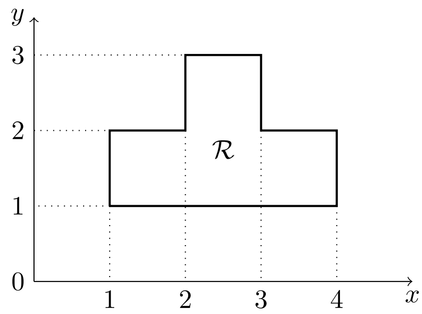
Informazione e stima – 11/09/2020

- Ogni esercizio viene valutato da 0 a 6 punti.
- Non riportare solo il risultato, ma cerca di argomentare sinteticamente la risposta.
- Riportare il proprio nome, cognome, e codice persona su ogni foglio consegnato.
- Nominare il file da caricare con il proprio codice persona.
- Esercizio da escludere dal punteggio finale:

1 Un docente deve formare 4 gruppi di 4 persone scegliendo tra 4 tutor e 12 studenti. Qual è la probabilità che in ogni gruppo ci sia esattamente un tutor, se le formazioni dei gruppi sono fatte in modo casuale?

2 Due variabili aleatorie X e Y hanno legge di probabilità $f_{X,Y}(x,y) = c$ nella regione $\mathcal{R}$ in figura. Determinare:

(a) il valore della costante c.
(b) Le leggi marginali di X e Y e graficarle.
(c) La legge $f_{Y|X}(y|1.5)$ e graficarla.



3 Sia $X \sim \text{Exp}(1)$ e $Y_n = X/n$ per $n \geq 1$. Dire se la successione di v.a. $\{Y_n\}_n$ converge in probabilità e, in caso affermativo, a quale valore.

4 I clienti di una libreria vengono serviti secondo un processo di Poisson con un ritmo di λ utenti all'ora. Ogni cliente compra un libro con probabilità p, indipendentemente da tutto il resto. Trovare:

(a) la legge di probabilità del tempo al primo acquisto;
(b) la probabilità che nessun libro venga venduto in un'ora;
(c) il numero atteso di clienti che comprano un libro tra le 12:00 e le 13:30.

5 Il numero Θ di carrelli di un supermercato è uniformemente distribuito tra 1 e 100. I carrelli sono numerati sequenzialmente da 1 a Θ. Quando entrate nel supermercato, osservate il numero X sul carrello che prendete, assunto uniformemente distribuito nell'intervallo da 1 a Θ, e usate questa informazione per stimare il valore di Θ. Determinare e graficare lo stimatore MAP $\widehat{\Theta}_{\text{MAP}}(X)$.

6 Si considerino $n = 10^8$ lanci indipendenti di una moneta sbilanciata con $P(\text{Testa}) = 0.8$. Si vogliono memorizzare i risultati dei lanci in un file in modo compresso, senza perdita di informazione. Qual è il minimo valore atteso possibile (in Byte) della lunghezza del file? Quale principio si può usare per ottenere un file di tale lunghezza?
Usare l'equivalenza 1 Byte = 8 bits.

Soluzioni

Problema 1

Il problema si può interpretare come una partizione in 4 gruppi. Ogni gruppo deve contenere necessariamente 1 tutor e 3 studenti. La probabilità cercata, intesa come rapporto tra casi favorevoli e casi totali, si può impostare grazie ai coefficienti multinomiali come segue:

$$\frac{\binom{4}{1,1,1,1}\binom{12}{3,3,3,3}}{\binom{16}{4,4,4,4}} = \frac{4^4}{\binom{16}{4}} \approx 0.14 \tag{1}$$

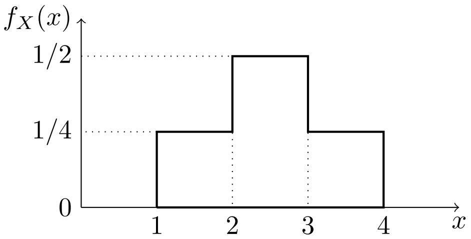
Problema 2

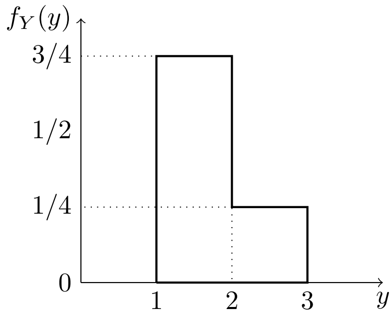
1. Essendo la distribuzione congiunta uniforme, si ha $c = 1/\text{Area}(\mathcal{R}) = 1/4$.

2. Le distribuzioni marginali sono:

$$f_X(x) = \int_1^3 f_{X,Y}(x,y)dy = \begin{cases} 1/4 & 1 < x < 2 \ \vee \ 3 < x < 4 \\ 1/2 & 2 < x < 3 \\ 0 & \text{altrove} \end{cases} \tag{2}$$

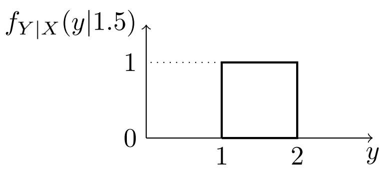
$$f_Y(y) = \int_1^4 f_{X,Y}(x,y)dx = \begin{cases} 3/4 & 1 < y < 2 \\ 1/4 & 2 < x < 3 \\ 0 & \text{altrove} \end{cases} \tag{3}$$





3. La legge condizionata è

$$f_{Y|X}(y|1.5) = \frac{f_{X,Y}(1.5,y)}{f_X(1.5)} = \begin{cases} 1 & 1 < y < 2 \\ 0 & \text{altrove} \end{cases} \tag{4}$$



Problema 3

Il sospetto è che la successione possa convergere al valore 0. Mostriamolo formalmente:

$$\Pr(|Y_n - 0| \geq \varepsilon) = \Pr(Y_n \geq \varepsilon) = \Pr(X \geq n\varepsilon) = 1 - F_X(n\varepsilon) = e^{-n\varepsilon} \tag{5}$$

dunque

$$\lim_{n\to\infty} \Pr(|Y_n - 0| \geq \varepsilon) = 0 \qquad \forall \varepsilon > 0 \tag{6}$$

che dimostra che $Y_n \longrightarrow 0$ in probabilità.

Problema 4

1. Si può considerare il processo split dei clienti che acquistano un libro, che è un processo di Poisson con ritmo λp clienti all'ora. Il tempo al primo arrivo (acquisto) in questo processo è Esponenziale con parametro λp.

2. Si può reinterpretare la richiesta come la probabilità che ci siano 0 arrivi durante un'ora del processo split degli acquirenti. Siccome il numero di arrivi in un intervallo di tempo specifico è una v.a. di Poisson, la probabilità cercata è $e^{-\lambda p}$.

3. Sfruttando il fatto che il numero di clienti acquirenti è una v.a. di Poisson, il valore atteso cercato è $3/2 \cdot \lambda p$.

Problema 5

Dal testo del problema sappiamo che $\Theta \sim \mathcal{U}\{1, 100\}$, e $\{X|\Theta = \theta\} \sim \mathcal{U}\{1, \theta\}$. Applichiamo la regola di Bayes per la ricerca della stima MAP di Θ basata sull'osservazione di $X = x$:

$$\begin{aligned}
\widehat{\Theta}_{\mathrm{MAP}}(X = x) &= \mathrm{argmax}_{\theta \in \{1,100\}} \quad p_{\Theta|X}(\theta|x) && (7) \\
&= \mathrm{argmax}_{\theta \in \{1,100\}} \quad p_{X|\Theta}(x|\theta) p_\Theta(\theta) && (8) \\
&= \mathrm{argmax}_{\theta \in \{1,100\}} \quad \frac{1}{\theta} \mathbb{1}(x \leq \theta) \cdot \frac{1}{100} && (9) \\
&= \mathrm{argmax}_{\theta \in \{x,100\}} \quad \frac{1}{\theta} && (10) \\
&= x && (11)
\end{aligned}$$

dove abbiamo ignorato le costanti moltiplicative indipendenti da θ, e usato la notazione $\mathbb{1}(x \leq \theta)$ per indicare il range di validità della legge $p_{X|\Theta}(x|\Theta)$.

Problema 6

Il teorema della codifica di sorgente dice che è possibile ottenere un file la cui lunghezza media minima è

$$\frac{10^8 H_2(0.8)}{8} \approx 9.02 \text{ MBytes} \tag{12}$$

dove 1 MByte = 10^6 Bytes. Un file di tale lunghezza si può ottenere usando un codice di sorgente, seguendo il principio generale di usare parole di codice corte per rappresentare sequenze di lanci molto probabili (ad es. una sequenza di Teste), e usare parole di codice lunghe per rappresentare sequenze di lanci poco probabili (ad esempio una sequenza di Croci).