Informazione e stima – 10/06/2023 – Compito A

- 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, codice persona, e tipo di compito svolto su ogni foglio consegnato
- Esercizio da escludere dal punteggio finale:

1. Un calciatore sfida un portiere ai calci di rigore. Ad ogni tiro, si sa che il calciatore non centra la porta con probabilità 0.2. Quando il calciatore centra la porta, il portiere para il tiro con probabilità 0.3. Tutti i calci di rigore vengono tirati in maniera indipendente. Mediamente, dopo quanti tiri il calciatore sbaglia il suo primo rigore?

2. Ad un centralino telefonico arrivano chiamate secondo un processo di Poisson con intensità 10 chiamate ogni ora. Al centralino lavorano n operatori, e le chiamate sono assegnate in maniera casuale e indipendente tra gli n operatori. Quanti operatori dovrà avere il centralino affinché la probabilità che ogni operatore riceva più di una chiamata ogni 15 minuti sia minore di 0.05?

3. Sia $\Theta \sim \mathcal{U}[0,2]$ e $X = T \cdot \Theta$ dove $T \sim \text{Exp}(1)$ è indipendente da Θ. Trovare lo stimatore lineare a minimo errore quadratico medio $\widehat{\Theta}_{\text{Lin}}(X)$ di Θ avendo a disposizione la misura X.

4. Avendo a disposizione un generatore di campioni uniformi $U \sim \mathcal{U}[0,1]$ e un generatore di campioni W distribuiti con legge $f_W(w) = 2w$ per $0 \leq w \leq 1$ e $f_W(w) = 0$ altrimenti, descrivere un algoritmo di acceptance-rejection (pseudocodice) per generare dei campioni X con legge $f_X(x) = 3x^2$ per $0 \leq x \leq 1$ e $f_X(x) = 0$ altrimenti.

   Qual è la migliore efficienza dell'algoritmo?

5. Due squadre di calcio A e B segnano goal in modo indipendente secondo due processi di Poisson con intensità $\lambda_A = 2$ e $\lambda_B = 3$ goal a partita, rispettivamente. Sapendo che in una sfida diretta tra le due squadre sono stati segnati 3 goal, qual è la probabilità che abbia vinto la squadra A?

6. Si consideri un codice formato da 5 messaggi con codewords di lunghezze $\{1, 2, 3, 3, 4\}$. Questo codice può essere decodificabile in maniera univoca? Se sì, dare un esempio di parole codice. Se no, proporre una modifica per renderlo decodificabile in maniera univoca.

Soluzioni compito A

Problema 1

Si consideri come successo un calcio di rigore sbagliato. Allora i calci di rigore si possono intepretare come esperimenti di un processo di Bernoulli con probabilità di successo $0.2 + (1 - 0.2) \cdot 0.3 = 0.44$. Il numero di tiri fino al primo rigore sbagliato è una v.a. Geometrica con parametro 0.44, e dunque il numero medio di tiri fino al primo rigore sbagliato è $1/0.44 \approx 2.27$.

Problema 2

Applicando uno split del processo di Poisson originario, ogni operatore vedrà arrivare le chiamate secondo un processo di Poisson con intensità λ/n dove $\lambda = 10/60$ chiamate al minuto. Sia N il tempo di chiamate che un operatore riceve in 15 minuti. Allora $N \sim \text{Poisson}(15\lambda/n)$. La probabilità che un operatore riceva più di una chiamata in 15 minuti è

$$
\begin{aligned}
\Pr(N > 1) &= 1 - \Pr(N = 0) - \Pr(N = 1) && (1)\\
&= 1 - e^{-15/(6n)} - \frac{15}{6n} e^{-15/(6n)}. && (2)
\end{aligned}
$$

Il quesito impone di trovare il più piccolo intero n tale che $\Pr(N > 1) \overset{!}{\leq} 0.05$. Andando per tentativi, si trova che $n = 8$.

Problema 3

Lo stimatore lineare a minimo errore quadratico medio è

$$\widehat{\Theta}_{\text{Lin}}(X) = \frac{\mathsf{Cov}[X, \Theta]}{\mathsf{Var}[X]}(X - \mathsf{E}[X]) + \mathsf{E}[\Theta].$$

Siccome $\Theta \sim \mathcal{U}[0, 2]$, abbiamo $\mathsf{E}[\Theta] = 1$, $\mathsf{Var}[\Theta] = \frac{2^2}{12} = \frac{1}{3}$ e quindi $\mathsf{E}[\Theta^2] = \mathsf{Var}[\Theta] + \mathsf{E}[\Theta]^2 = \frac{4}{3}$. Siccome $T \sim \text{Exp}(1)$, abbiamo $\mathsf{E}[T] = 1$, $\mathsf{Var}[T] = 1$ e quindi $\mathsf{E}[T^2] = \mathsf{Var}[T] + \mathsf{E}[T]^2 = 2$. Inoltre, abbiamo che

$$\mathsf{E}[X] = \mathsf{E}[T \cdot \Theta] = \mathsf{E}[T]\mathsf{E}[\Theta] = 1$$

$$\mathsf{E}[X^2] = \mathsf{E}[T^2 \cdot \Theta^2] = \mathsf{E}[T^2]\mathsf{E}[\Theta^2] = \frac{8}{3}$$

grazie all'indipendenza tra T e Θ. Ora possiamo calcolare

$$\mathsf{Var}[X] = \mathsf{E}[X^2] - \mathsf{E}[X]^2 = \frac{5}{3}.$$

Infine, abbiamo

$$
\begin{aligned}
\mathsf{Cov}[X, \Theta] &= \mathsf{E}[X\Theta] - \mathsf{E}[X]\mathsf{E}[\Theta] && (3)\\
&= \mathsf{E}[T\Theta^2] - \mathsf{E}[X]\mathsf{E}[\Theta] && (4)\\
&= \mathsf{E}[T]\mathsf{E}[\Theta^2] - \mathsf{E}[X]\mathsf{E}[\Theta] && (5)\\
&= \frac{4}{3} - 1 = \frac{1}{3} && (6)
\end{aligned}
$$

e

$$\widehat{\Theta}_{\text{Lin}}(X) = \frac{1}{5}(X - 1) + 1 = \frac{1}{5}X + \frac{4}{5}.$$

Problema 4

Innanzitutto dobbiamo trovare un valore di m tale che $m f_W(x) \geq f_X(x)$ per ogni $0 \leq x \leq 1$. I grafici di f_W e f_X suggeriscono che è sufficiente imporre $m f_W(1) \geq f_X(1)$:

$$m2 \geq 3 \Longrightarrow m \geq \frac{3}{2}.$$

Il valore di m che darà massima efficienza è $m = \frac{3}{2}$. Lo pseudocodice è come segue:

1. Genero un campione $W \sim f_W$ e un campione $U \sim \mathcal{U}[0, 1]$ in maniera indipendente
2. Accetto e pongo $X = W$ se $m f_W(W) U \leq f_X(W)$, altrimenti torno al punto 1.

L'efficienza migliore dell'algoritmo è $\frac{1}{m} = \frac{2}{3} = 66.6\%$

Problema 5

I goal segnati nello scontro diretto tra le due squadre seguono un processo di Poisson con intensità $\mu = \lambda_A + \lambda_B = 5$ goal per partita. Siccome i goal vengono segnati in maniera indipendente, i tre goal segnati possono essere visti come 3 prove indipendenti di Bernoulli con probabilità di successo $p_A = \frac{\lambda_A}{\mu} = \frac{2}{5}$. Sia X il numero di goal segnati della squadra A, allora la probabilità che la squadra A abbia vinto è

$$\begin{aligned} \Pr(X \geq 2) &= \Pr(X = 2) + \Pr(X = 3) && (7) \\ &= \binom{3}{2} p_A^2 (1 - p_A) + \binom{3}{3} p_A^3 \approx 0.352 && (8) \end{aligned}$$

Problema 6

Verifichiamo se la diseguaglianza di Kraft-McMillan è soddisfatta:

$$\sum_{i=1}^{5} 2^{-\ell_i} = 2^{-1} + 2^{-2} + 2 \cdot 2^{-3} + 2^{-4} = 1.0625 \geq 1$$

dunque non esiste un codice decodificabile in maniera univoca con queste lunghezze di parola. Un modo per rendere il codice decodificabile è allungare qualche parola. Ad esempio, si potrebbero usare codewords con lunghezze $\{1, 3, 3, 3, 4\}$.

Informazione e stima – 10/06/2023 – Compito B

- 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, codice persona, e tipo di compito svolto su ogni foglio consegnato
- Esercizio da escludere dal punteggio finale:

1. In un villaggio ci sono 10 banche, 6 ristoranti, e 4 cinema. Una coppia decide di visitare 3 posti scelti in maniera casuale. Qual è la probabilità che la coppia non visiti sempre lo stesso tipo di attività?

2. Siano X e Y due variabili aleatorie indipendenti e distribuite uniformemente nell'intervallo $[0, 1]$. Trovare la legge della variabile aleatoria $Z = -\ln(XY)$.

3. Sia $X = \sum_{i=0}^{\infty} T_i$ dove le $T_i \sim \text{Exp}(2^i)$ per $i = 0, 1, \dots$ sono tutte variabili aleatorie indipendenti. Proporre un'approssimazione per le seguenti probabilità:

   (a) $\Pr(X > 20)$

   (b) $\Pr(|X - \mathsf{E}[X]| > 2)$

4. Ad un centralino telefonico arrivano chiamate secondo un processo di Poisson con intensità 10 chiamate ogni ora. Al centralino lavorano n operatori, e le chiamate sono assegnate in maniera casuale e indipendente tra gli n operatori. Quanti operatori dovrà avere il centralino affinché la probabilità che ogni operatore riceva più di una chiamata ogni 15 minuti sia minore di 0.05?

5. Sia $\Theta \sim \mathcal{U}[0, 2]$ e $X = T \cdot \Theta$ dove $T \sim \text{Exp}(1)$ è indipendente da Θ. Trovare lo stimatore lineare a minimo errore quadratico medio $\widehat{\Theta}_{\text{Lin}}(X)$ di Θ avendo a disposizione la misura X.

6. Avendo a disposizione un generatore di campioni uniformi $U \sim \mathcal{U}[0, 1]$ e un generatore di campioni W distribuiti con legge $f_W(w) = 2w$ per $0 \leq w \leq 1$ e $f_W(w) = 0$ altrimenti, descrivere un algoritmo di acceptance-rejection (pseudocodice) per generare dei campioni X con legge $f_X(x) = 3x^2$ per $0 \leq x \leq 1$ e $f_X(x) = 0$ altrimenti.

Soluzioni compito B

Problema 1

Lo spazio di probabilità è uniforme discreto. Il numero di casi totali è $\binom{20}{3}$. Conviene calcolare la probabilità finale come 1 meno la probabilità dell'evento complementare. I casi favorevoli per l'evento complementare sono la visita di 3 banche, oppure di 3 ristoranti, oppure di 3 cinema, dunque basta sommare $\binom{10}{3} + \binom{6}{3} + \binom{4}{3}$. La probabilità finale è

$$1 - \frac{\binom{10}{3} + \binom{6}{3} + \binom{4}{3}}{\binom{20}{3}} = 1 - \frac{120 + 20 + 4}{1140} = \frac{996}{1140} \approx 0.8737 \tag{9}$$

Problema 2

Notiamo innanzitutto che

$$Z = -\ln(XY) = -\ln(X) - \ln(Y).$$

Definendo $X' = -\ln(X)$ e $Y' = -\ln(Y)$, sappiamo dalla teoria che X' e Y' sono due variabili aleatorie indipendenti ed esponenziali di parametro $\lambda = 1$. Dunque $Z = X' + Y'$, somma di due v.a. esponenziali indipendenti, si può interpretare come il tempo al secondo successo in un processo di Poisson di intensità $\lambda = 1$. Pertanto, dalla teoria sappiamo che $Z \sim$ Erlang-2(1), con legge

$$f_Z(z) = ze^{-z}, \qquad z > 0$$

e $f_Z(z) = 0$ altrimenti.

Problema 3

1. Siccome X è una variabile aleatoria positiva, possiamo usare la diseguaglianza di Markov:

$$\Pr(X > 20) \leq \frac{\mathsf{E}[X]}{20} = \frac{1}{20}\sum_{i=0}^{\infty} \mathsf{E}[T_i] = \frac{1}{20}\sum_{i=0}^{\infty} \frac{1}{2^i} = \frac{1}{10} \tag{10}$$

2. Qui possiamo usare la diseguaglianza di Chebyshev:

$$\Pr(|X - \mathsf{E}[X]| > 2) \leq \frac{\mathsf{Var}[X]}{4} = \frac{1}{4}\sum_{i=0}^{\infty} \mathsf{Var}[T_i] = \frac{1}{4}\sum_{i=0}^{\infty} \frac{1}{4^i} = \frac{1}{4}\frac{1}{1-\frac{1}{4}} = \frac{1}{3}. \tag{11}$$

Problema 4

Applicando uno split del processo di Poisson originario, ogni operatore vedrà arrivare le chiamate secondo un processo di Poisson con intensità λ/n dove $\lambda = 10/60$ chiamate al minuto. Sia N il tempo di chiamate che un operatore riceve in 15 minuti. Allora $N \sim$ Poisson$(15\lambda/n)$. La probabilità che un operatore riceva più di una chiamata in 15 minuti è

$$\Pr(N > 1) = 1 - \Pr(N = 0) - \Pr(N = 1) \tag{12}$$
$$= 1 - e^{-15/(6n)} - \frac{15}{6n}e^{-15/(6n)}. \tag{13}$$

Il quesito impone di trovare il più piccolo intero n tale che $\Pr(N > 1) \overset{!}{\leq} 0.05$. Andando per tentativi, si trova che $n = 8$.

Problema 5

Lo stimatore lineare a minimo errore quadratico medio è

$$\widehat{\Theta}_{\text{Lin}}(X) = \frac{\mathsf{Cov}[X, \Theta]}{\mathsf{Var}[X]}(X - \mathsf{E}[X]) + \mathsf{E}[\Theta].$$

Siccome $\Theta \sim \mathcal{U}[0, 2]$, abbiamo $\mathsf{E}[\Theta] = 1$, $\mathsf{Var}[\Theta] = \frac{2^2}{12} = \frac{1}{3}$ e quindi $\mathsf{E}[\Theta^2] = \mathsf{Var}[\Theta] + \mathsf{E}[\Theta]^2 = \frac{4}{3}$. Siccome $T \sim \text{Exp}(1)$, abbiamo $\mathsf{E}[T] = 1$, $\mathsf{Var}[T] = 1$ e quindi $\mathsf{E}[T^2] = \mathsf{Var}[T] + \mathsf{E}[T]^2 = 2$. Inoltre, abbiamo che

$$\mathsf{E}[X] = \mathsf{E}[T \cdot \Theta] = \mathsf{E}[T]\mathsf{E}[\Theta] = 1$$

$$\mathsf{E}[X^2] = \mathsf{E}[T^2 \cdot \Theta^2] = \mathsf{E}[T^2]\mathsf{E}[\Theta^2] = \frac{8}{3}$$

grazie all'indipendenza tra T e Θ. Ora possiamo calcolare

$$\mathsf{Var}[X] = \mathsf{E}[X^2] - \mathsf{E}[X]^2 = \frac{5}{3}.$$

Infine, abbiamo

$$\begin{aligned}
\mathsf{Cov}[X,\Theta] &= \mathsf{E}[X\Theta] - \mathsf{E}[X]\mathsf{E}[\Theta] && (14)\\
&= \mathsf{E}[T\Theta^2] - \mathsf{E}[X]\mathsf{E}[\Theta] && (15)\\
&= \mathsf{E}[T]\mathsf{E}[\Theta^2] - \mathsf{E}[X]\mathsf{E}[\Theta] && (16)\\
&= \frac{4}{3} - 1 = \frac{1}{3} && (17)
\end{aligned}$$

e

$$\widehat{\Theta}_{\mathrm{Lin}}(X) = \frac{1}{5}(X-1) + 1 = \frac{1}{5}X + \frac{4}{5}.$$

Problema 6

Innanzitutto dobbiamo trovare un valore di m tale che $mf_W(x) \geq f_X(x)$ per ogni $0 \leq x \leq 1$. I grafici di f_W e f_X suggeriscono che è sufficiente imporre $mf_W(1) \geq f_X(1)$:

$$m2 \geq 3 \Longrightarrow m \geq \frac{3}{2}.$$

Il valore di m che darà massima efficienza è $m = \frac{3}{2}$. Lo pseudocodice è come segue:

1. Genero un campione $W \sim f_W$ e un campione $U \sim \mathcal{U}[0,1]$ in maniera indipendente
2. Accetto e pongo $X = W$ se $mf_W(W)U \leq f_X(W)$, altrimenti torno al punto 1.

L'efficienza migliore dell'algoritmo è $\frac{1}{m} = \frac{2}{3} = 66.6\%$