Soluzioni all'esercizio del 2022-05-25 creato per luigi.miazzo

Abbiamo quattro variabili aleatorie: X, Y, W, Z.

Le prime due, X e Y, sono indipendenti tra loro e hanno entrambe distribuzione poissoniana, di parametri $\lambda_X = 4$ e $\lambda_Y = 10$ rispettivamente.

La variabile Z è una normale (o Gaussiana) standard, mentre $W = 1.9 \cdot Z$.

Quesiti e soluzioni

Quesito 1

Dare la migliore possibile stima dall'alto della probabilità che la somma $X + Y$ sia maggiore o uguale a 17, usando la disuguaglianza di Markov.

La disuguaglianza di Markov, per una variabile aleatoria non negativa T, è:

$$P(T \geq a) \leq \frac{E[T]}{a}.$$

Per poter usare la disuguaglianza di Markov non abbiamo bisogno di conoscere (o ricordare) la funzione di ripartizione o la densità discreta. Abbiamo però bisogno di calcolare la speranza della variabile aleatoria $X+Y$: la speranza della somma è la somma delle speranze, quindi, siccome X e Y sono di Poisson e quindi riproducibili,

$$E[X+Y] = E[X] + E[Y] = \lambda_X + \lambda_Y = 14.$$

Ora possiamo scrivere la disuguaglianza nel nostro caso (osserviamo che entrambe le variabili aleatorie sono non negative e dunque tale è anche la loro somma):

$$P(X+Y \geq a) \leq \frac{E[X+Y]}{a}.$$

Questa è la miglior stima possibile purché sia minore o uguale di 1.

- La risposta corretta è: 0.8235294
- La risposta inserita è: 14/17

Quesito 2

Dare la migliore possibile stima dall'alto della probabilità che la somma $X + Y$ sia maggiore o uguale a 17, usando la disuguaglianza di Markov.

Sappiamo che la distribuzione di X sapendo che $X + Y = n$ è una binomiale di parametri n e $p = \frac{\lambda_X}{\lambda_X + \lambda_Y}$.

Il valore atteso condizionato è allora

$$E[X|X+Y=n] = n \cdot p = n\frac{\lambda_X}{\lambda_X + \lambda_Y}.$$

Possiamo ora andare a considerare la disuguaglianza di Markov:

$$P(X \geq 7.88 | X+Y=17) \leq \frac{E[X|X+Y=17]}{7.88} = \frac{17 \cdot 4}{7.88 \cdot (4+10)}.$$

Questa, però, è la miglior stima possibile solamente se è minore o uguale di 1 e se $7.88 \in \mathbb{N}$. Infatti siccome la stima data dalla disuguaglianza di Markov decresce al crescere del denominatore, ci conviene prendere il denominatore più grande possibile, a parità di probabilità a primo membro. In questo caso la cosa è rilevante, perché X assume solo valori in $\mathbb{N}$, quindi

$$P(X \geq y | X+Y=17) = P(X \geq \lceil y \rceil | X+Y=17)$$

e dunque possiamo stimare

$$P(X \geq 7.88 | X+Y=17) = P(X \geq 8 | X+Y=17) \leq \frac{E[X|X+Y=17]}{8} = \frac{17 \cdot 4}{8 \cdot (4+10)}.$$

- La risposta corretta è: 0.6071429
- La risposta inserita è: 17/28

Quesito 3

Dare una stima dal basso della probabilità $P(|W| < 3.7)$, usando la disuguaglianza di Chebychev.

Siccome $Z \sim \mathcal{N}(0,1)$, ha speranza 0 e varianza 1, W ha speranza 0 e varianza $1.9^2 = 3.61$ (si tratta di una trasformazione lineare).

Non solo, in realtà non ci occorre per questo quesito (pur essendo utile nel prossimo), ma conosciamo la distribuzione di W: $W \sim \mathcal{N}(0, 1.9)$.

Osserviamo che possiamo riscrivere la quantità che ci interessa nel modo seguente:

$$P(|W| < 3.7) = 1 - P(|W| \geq 3.7).$$

Scriverla in questa forma ci permette di mettere in evidenza il termine che possiamo stimare con la disuguaglianza di Chebychev:

$$P(|W| \geq 3.7) \leq \frac{3.61}{13.69},$$

dove abbiamo usato il fatto che la varianza di W è 3.61.

Questa è una stima dall'alto, ma siccome la stiamo sottraendo da 1, l'uguaglianza $P(|W| < 3.7) = 1 - P(|W| \geq 3.7)$ diventa la disuguaglianza

$$P(|W| < 3.7) = 1 - P(|W| \geq 3.7) \geq 1 - \frac{3.61}{13.69}.$$

Abbiamo dunque ottenuto la stima dal basso cercata.

- La risposta corretta è: 0.7363039
- La risposta inserita è: 1008/1369

Quesito 4

Qual è l'errore assoluto (rispetto al valore vero di $P(|W| < 3.7)$) fatto nella stima precedente?

Qui, a differenza del quesito precedente, vogliamo sfruttare il fatto che sappiamo la distribuzione di W. Possiamo infatti riscrivere la quantità cercata nel modo seguente:

$$P(|W| < 3.7) = P(-3.7 < W < 3.7) = F_W(3.7) - F_W(-3.7).$$

Osserviamo che la media di W è 0 (abbiamo solo riscalato Z, non l'abbiamo traslata). Questo ci dice che possiamo sfruttare questo per semplificare la differenza sopra. In realtà possiamo fare la stessa considerazione dopo aver standardizzato:

$$P(|W| < 3.7) = \Phi\left(\frac{3.7}{1.9}\right) - \Phi\left(-\frac{3.7}{1.9}\right) = \Phi\left(\frac{3.7}{1.9}\right) - 1,$$

quantità che possiamo calcolare in R, `2*pnorm(x)-1` con $x = \frac{3.7}{1.9}$ o con l'ausilio delle tavole.

Per finire, non ci resta che sottrarre a questo valore quello ottenuto al quesito 3 e prenderne il valore assoluto.

- La risposta corretta è: 0.2122055
- La risposta inserita è: 18071/85158