ESERCITAZIONE 6– DOCENTE: MAURO PICCIONI, TUTOR: HLAFO ALFIE MIMUN

April 11, 2020

1. ESERCIZI

Ex. 1: Alice e Bob si sono appena incontrati e si chiedono se hanno amici in comune. Alice e Bob vivono nello stesso paese, che ha 1000 abitanti (esclusi Alice e Bob). Entrambi hanno 50 amici che vivono nella loro stessa città. Si assuma che i 50 amici di Alice siano scelti uniformemente a caso tra i 1000 abitanti del paese. Si assuma anche che la conoscenza di chi siano gli amici di Alice non dia alcuna informazione su chi siano gli amici di Bob. Sia X il numero di amici che Alice e Bob hanno in comune. Si calcoli $\mathbb{E}[X]$.

(1a) Si calcoli $\mathbb{E}[X]$.

(1b) Si calcoli la PMF di X e si espliciti se si tratta di una PMF nota.

Ex. 2: Per sconfiggere gli spam, Bob ha installato 2 programmi anti-spam, che chiamiamo P_1 e P_2. Ogni email che arriva è marcata da ognuno dei due programmi come legittima o spam. Definiamo gli eventi

$$L = \{\text{la mail è legittima}\},$$

$$M_j = \{\text{la mail è marcata come legittima dal programma } P_j\}, \text{ for } j = 1, 2.$$

Assumiamo che il 10% delle mail di Bob sia legittimo e che i due programmi siano accurati al 90%, nel senso che

$$\mathbb{P}(M_j \mid L) = \mathbb{P}(M_j^c \mid L^c) = \frac{9}{10}, \quad \text{for } j = 1, 2.$$

Assumiamo anche che per ogni mail che arriva, i due programmi decidono se essa sia spam o no indipendentemente l'uno dall'altro.

(2a) Si calcoli la probabilità che la mail sia legittima sapendo che viene marcata come legittima dal programma P_1.

(2b) Si calcoli la probabilità che la mail sia legittima sapendo che entrambi i programmi la marchino come legittima.

Ex. 3: In una lotteria si estraggono 5 numeri tra i primi 35 interi (ovvero $1, 2, \dots, 34, 35$). I numeri si estraggono in blocco. Si sta scommettendo sui 5 numeri estratti. Si trovi la probabilità di indovinare 3 numeri sapendo che ne è stato indovinato almeno 1.

Ex. 4: Il proprietario di un sito web sta studiando la distribuzione del numero di visitatori del suo sito. Ogni giorno un milione di persone decide di visitare o meno il sito indipendentemente dalle altre persone. La probabilità che una persona visiti il sito è pari a $p = 2 \cdot 10^{-6}$.

(4a) Si dia una buona approssimazione della probabilità di avere almeno 2 visitatori in un dato giorno.

(4b) Quanti giorni si dovranno aspettare in media per avere un giorno in cui le visite al sito siano almeno 2?

Ex. 5: Una statistica sull'altezza media di una classe di 16 persone ha provato che con probabilità $\frac{1}{4}$ uno studente è più alto 1.75 metri.

(5a) Si calcoli il numero medio di studenti che è più alto di 1.75 metri;

(5b) si calcoli la probabilità che il numero di studenti che è più alto di 1.75 metri sia maggiore di 2 e minore di 6.

2. SOLUZIONI

Ex. 1: Soluzione (1a):

Definiamo per $j = 1, \ldots, 1000$ la variabile

$$I_j := \begin{cases} 1, & \text{se il } j\text{-esimo abitante è un amico in comune;} \\ 0, & \text{altrimenti.} \end{cases}$$

Si noti che le variabili aleatorie $\{I_j\}_{j=1}^{1000}$ sono i.i.d. ognuna con distribuzione di Bernoulli di parametro p dato da

$$p = \mathbb{P}(I_j = 1) = \mathbb{P}(\text{il } j\text{-esimo abitante è un amico sia di Alice sia di Bob})\,.$$

Per le ipotesi fatte, gli eventi $\{\text{il } j\text{-esimo abitante è un amico di Alice}\}$ e $\{\text{il } j\text{-esimo abitante è un amico di Bob}\}$ sono indipendenti e

$$\mathbb{P}(\{\text{il } j\text{-esimo abitante è un amico di Alice}\}) =$$
$$= \mathbb{P}(\{\text{il } j\text{-esimo abitante è un amico di Bob}\}) = \frac{50}{1000} = \frac{1}{20}\,.$$

Dunque

$$\mathbb{P}(I_j = 1) = \mathbb{P}(\{\text{il } j\text{-esimo abitante è un amico sia di Alice sia di Bob}\}) =$$
$$= \mathbb{P}(\{\text{il } j\text{-esimo abitante è un amico di Alice}\}) \cdot \mathbb{P}(\{\text{il } j\text{-esimo abitante è un amico di Bob}\}) =$$
$$= \left(\frac{1}{20}\right)^2\,.$$

Dunque, poiché

$$X = \sum_{j=1}^{1000} I_j\,,$$

si ha che

$$\mathbb{E}[X] = \sum_{j=1}^{1000} \mathbb{E}[I_j] = \sum_{j=1}^{1000} \mathbb{P}(I_j = 1) = \sum_{j=1}^{1000} \left(\frac{1}{20}\right)^2 = 1000 \cdot \frac{1}{400} = 2.5\,.$$

Soluzione (1b):

Fissiamo chi sono gli amici di Alice. Vogliamo calcolare la probabilità

$$\mathbb{P}(X = k)$$

per $k = 0, 1, \ldots, 50$. Se $X = k$ vuol dire che quando scegliamo gli amici di Bob, dobbiamo sceglierne k dai 50 amici di Alice ed i restanti $50 - k$ dal resto della popolazione (ovvero 1000-50=950 persone). Ogni modo di formare questo gruppo (secondo queste regole) ha stessa probabilità. Dunque possiamo usare la formula

$$\frac{\text{numero di casi favorevoli}}{\text{numero di casi possibili}}\,.$$

Il numero di casi favorevoli per questo evento è dato dal numero di modi di creare un gruppo di k persone della forma appena descritta. Il numero di casi possibili invece è dato dal numero di formare un gruppo di 50 persone (gli amici di Bob) da una popolazione di 1000 persone. Dunque per $k = 0, 1, \dots, 50$

$$\mathbb{P}(X = k) = \frac{\binom{50}{k}\cdot\binom{1000-50}{50-k}}{\binom{1000}{50}}\,.$$

Notiamo dunque che X ha distribuzione Ipergeometrica di parametri 1000,50,50.

Ex. 2: Soluzione (2a):

Dobbiamo calcolare $\mathbb{P}(L \,|\, M_1)$. Applicando la formula di Bayes

$$\mathbb{P}(L \,|\, M_1) = \frac{\mathbb{P}(M_1 \,|\, L)\cdot\mathbb{P}(L)}{\mathbb{P}(M_1 \,|\, L)\mathbb{P}(L) + \mathbb{P}(M_1 \,|\, L^c)\mathbb{P}(L^c)}\,.$$

Sappiamo che

$$\mathbb{P}(M_1 \,|\, L) = \frac{9}{10}\,, \quad \mathbb{P}(M_1 \,|\, L^c) = 1 - \mathbb{P}(M_1^c \,|\, L^c) = 1 - \frac{9}{10} = \frac{1}{10}\,,$$

$$\mathbb{P}(L) = \frac{10}{100} = \frac{1}{10}\,,$$

in quanto sappiamo che il 10% delle mail è legittimo. Dunque

$$\mathbb{P}(L \,|\, M_1) = \frac{\mathbb{P}(M_1 \,|\, L)\cdot\mathbb{P}(L)}{\mathbb{P}(M_1 \,|\, L)\mathbb{P}(L) + \mathbb{P}(M_1 \,|\, L^c)\mathbb{P}(L^c)} = \frac{\frac{9}{10}\cdot\frac{1}{10}}{\frac{9}{10}\cdot\frac{1}{10}+\frac{1}{10}\cdot\frac{9}{10}} = \frac{1}{2}\,.$$

Soluzione (2b):

Dobbiamo calcolare $\mathbb{P}(L \,|\, M_1 \cap M_2)$. Usando sempre prima la formula di Bayes e poi il fatto che i due programmi lavorano indipendentemente, abbiamo

$$\begin{aligned}\mathbb{P}(L \,|\, M_1 \cap M_2) &= \frac{\mathbb{P}(M_1 \cap M_2 \,|\, L)\cdot\mathbb{P}(L)}{\mathbb{P}(M_1 \cap M_2 \,|\, L)\mathbb{P}(L) + \mathbb{P}(M_1 \cap M_2 \,|\, L^c)\mathbb{P}(L^c)} = \\ &= \frac{\mathbb{P}(M_1 \,|\, L)\mathbb{P}(M_2 \,|\, L)\cdot\mathbb{P}(L)}{\mathbb{P}(M_1 \,|\, L)\mathbb{P}(M_2 \,|\, L)\mathbb{P}(L) + \mathbb{P}(M_1 \,|\, L^c)\mathbb{P}(M_2 \,|\, L^c)\mathbb{P}(L^c)} = \\ &= \frac{\frac{9}{10}\cdot\frac{9}{10}\cdot\frac{1}{10}}{\frac{9}{10}\cdot\frac{9}{10}\cdot\frac{1}{10}+\frac{1}{10}\cdot\frac{1}{10}\cdot\frac{9}{10}} = \frac{9}{10}\,.\end{aligned}$$

Ex. 3: Definiamo gli eventi

$$E = \{\text{si indovinano esattamente 3 numeri estratti}\},$$
$$F = \{\text{si indovina almeno un numero}\}.$$

Dobbiamo calcolare

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Notiamo che $E \cap F = E$ ed E si verifica se si scelgono 3 numeri dai 5 numeri fortunati e 2 numeri dai restanti 35-5=30. Dunque

$$\mathbb{P}(E \cap F) = \mathbb{P}(E) = \frac{\binom{5}{3} \cdot \binom{35-5}{2}}{\binom{35}{5}}.$$

Per quanto riguarda $\mathbb{P}(F)$ è più facile analizzare $\mathbb{P}(F^c)$, dove

$$F^c := \{\text{nessun numero viene indovinato}\}.$$

Per realizzare F^c significa che estraiamo i 5 numeri dai 35-5=30 numeri sfortunati. Dunque

$$\mathbb{P}(F^c) = \frac{\binom{30}{5}}{\binom{35}{5}}.$$

Quindi

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)} = \frac{\frac{\binom{5}{3} \cdot \binom{35-5}{2}}{\binom{35}{5}}}{1 - \frac{\binom{30}{5}}{\binom{35}{5}}} = \frac{\binom{5}{3} \cdot \binom{35-5}{2}}{\binom{35}{5} - \binom{30}{5}}.$$

Ex. 4: Soluzione (4a):

Potremmo pensare al problema tramite una variabile aleatoria binomiale di parametri $n = 10^6$ e $p = 2 \cdot 10^{-6}$, in quanto sappiamo che abbiamo un insieme di 10^6 persone ed ognuna di esse con probabilità di visita p visita il sito. Se provassimo però una trattazione con tale variabile otterremmo valori intrattabili. Sappiamo che quando in una binomiale il numero n di prove (ovvero $n = 10^6$) è grande e la probabilità di successo (ovvero la probabilità di visita p) è piccola, si può approssimare tale variabile con una variabile di Poisson di densità pari a $\lambda = n \cdot p = 10^6 \cdot 2 \cdot 10^{-6} = 2$. Sia dunque $X \sim \text{Pois}(2)$ la variabile che rappresenta il numero di visite al giorno del sito web. Dobbiamo calcolare $\mathbb{P}(X \geq 2)$. Si ha

$$\mathbb{P}(X \geq 2) = 1 - [\mathbb{P}(X = 0) + \mathbb{P}(X = 1)] = 1 - e^{-2} - 2e^{-2} = 1 - 3e^{-2}. \qquad (1)$$

Soluzione (4b):

Sia T il numero di giorni che bisogna attendere per avere il primo giorno in cui le visite al sito siano almeno 2. Dunque $T = 5$ significa che abbiamo aspettato 5 giorni, ovvero per 4 giorni abbiamo avuto meno di 2 visite ed il quinto giorno abbiamo avuto almeno 2 visite. Poichè i visitatori di un giorno sono indipendenti dai visitatori di un altro giorno, abbiamo che $T \sim \text{Geom}(q)$, dove $q = \mathbb{P}(\text{almeno due visite in un dato giorno})$. Si noti che il valore di q è dato dalla soluzione alla parte (4a) dell'esercizio, ovvero (si veda (1)) si ha

$$q = \mathbb{P}(X \geq 2) = 1 - 3e^{-2} .$$

Poiché $T \sim \text{Geom}(q)$ we have that

$$\mathbb{E}[T] = \frac{1}{q} = \frac{1}{1 - 3e^{-2}} .$$

Ex. 5: Soluzione (5a):

Sia X il numero di studenti che è più alto di 1.75 metri. Si noti che ogni persona ha probabilità $\frac{1}{4}$ di essere più alto di 1.75 metri ed è ragionevole assumere che le altezze delle persone siano indipendenti tra loro. Dunque $X \sim \text{Bin}\left(16, \frac{1}{4}\right)$ e dunque

$$\mathbb{E}[X] = 16 \cdot \frac{1}{4} = 4 .$$

Soluzione (5b):

Se continuiamo a denotare con $X \sim \text{Bin}\left(16, \frac{1}{4}\right)$ il numero di studenti che è più alto di 1.75 metri, dobbiamo calcolare $\mathbb{P}(2 < X < 6)$.

$$\begin{aligned}
\mathbb{P}(2 < X < 6) = \mathbb{P}(3 \leq X \leq 5) &= \mathbb{P}(X = 3) + \mathbb{P}(X = 4) + \mathbb{P}(X = 5) = \\
&= \binom{16}{3}\left(\frac{1}{4}\right)^3\left(1 - \frac{1}{4}\right)^{16-3} + \binom{16}{4}\left(\frac{1}{4}\right)^4\left(1 - \frac{1}{4}\right)^{16-4} + \\
&+ \binom{16}{5}\left(\frac{1}{4}\right)^5\left(1 - \frac{1}{4}\right)^{16-5} \approx \\
&\approx 0.21 + 0.23 + 0.18 = 0.62 .
\end{aligned}$$