Informazione e stima – 19/04/2023

- 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 codice persona su ogni foglio consegnato.
- Indicare l'esercizio da escludere dal punteggio finale.

1. Supponiamo di avere un mazzo di carte ben mescolato composto da 52 carte. Si estraggono tre carte a caso senza reinserimento. Qual è la probabilità che le tre carte estratte siano tutte di seme diverso?

   Nel mazzo di sono 4 *semi diversi e* 13 *carte per ogni seme.*

2. Una persona per andare sul posto di lavoro usa l'auto il 40% delle volte, il bus il 50% delle volte, e va a piedi il 10% delle volte. Nei tre casi arriva in ritardo il 3%, il 10%, e il 15% delle volte, rispettivamente.

   (a) Qual è la probabilità che abbia preso il bus sapendo che è arrivata in ritardo?

   (b) Qual è la probabilità che sia arrivata a piedi, sapendo che è in orario?

3. Siano $U \sim \mathcal{U}[-1, 1]$ e $Y = U + Z$ dove Z è una v.a. Gaussiana standard indipendente da U. Determinare

   (a) valore atteso e varianza di Y;

   (b) la covarianza tra Y e U;

   (c) il coefficiente di correlazione lineare tra U e Z.

4. Siano Y_1 e Y_n due v.a. continue con densità di probabilità congiunta $f_{Y_1,Y_n}(y_1, y_n) = n(n-1)(y_n - y_1)^{n-2}$ per $0 \leq y_1 \leq y_n \leq 1$ e zero altrove. Determinare la densità di probabilità di $R = Y_n - Y_1$.

   Consiglio: passare dalla cumulata di R e aiutarsi con una rappresentazione grafica dell'evento di interesse per il calcolo della probabilità.

5. La temperatura corporea di una persona pescata a caso può essere modellizzata con una v.a. Gaussiana con media 36.8°C e deviazione standard 0.4°C.

   (a) Qual è l'intervallo, centrato sulla media 36.8°C, che contiene il 90% delle persone?

   (b) Qual è la percentuale di popolazione con una temperatura superiore a 37.5°C?

6. Sia X_n una v.a. con densità di probabilità $f_{X_n}(x) = n(1-x)^{n-1}$ per $0 < x < 1$ e $f_{X_n}(x) = 0$ altrove. Stabilire se la sequenza $\{X_n\}_{n\in\mathbb{N}}$ converge in probabilità e, se sì, a quale valore.

Soluzioni

Problema 1

Lo spazio di probabilità è discreto uniforme. La probabilità cercata è

$$\Pr(3 \text{ semi diversi}) = \Pr(\text{Prime 2 carte hanno seme diverso}) \Pr(\text{terza carta seme diverso}|\text{prime 2 carte hanno seme diverso}) \tag{1}$$

$$= \frac{39}{51} \cdot \frac{26}{50} = 0.3976 \tag{2}$$

Problema 2

1. La soluzione si ottiene applicando il teorema di Bayes:

$$\Pr(\text{Bus}|\text{Ritardo}) = \frac{\Pr(\text{Ritardo}|\text{Bus}) \Pr(\text{Bus})}{\Pr(\text{Ritardo})} \tag{3}$$

$$= \frac{\Pr(\text{Ritardo}|\text{Bus}) \Pr(\text{Bus})}{\Pr(\text{Ritardo}|\text{Bus}) \Pr(\text{Bus}) + \Pr(\text{Ritardo}|\text{Auto}) \Pr(\text{Auto}) + \Pr(\text{Ritardo}|\text{Piedi}) \Pr(\text{Piedi})} \tag{4}$$

$$= \frac{0.1 \cdot 0.5}{0.1 \cdot 0.5 + 0.03 \cdot 0.4 + 0.15 \cdot 0.1} \approx 0.6494 \tag{5}$$

2. La soluzione si ottiene applicando il teorema di Bayes:

$$\Pr(\text{Piedi}|\text{Puntuale}) = \frac{\Pr(\text{Puntuale}|\text{Piedi}) \Pr(\text{Piedi})}{\Pr(\text{Puntuale})} \tag{6}$$

$$= \frac{\Pr(\text{Puntuale}|\text{Piedi}) \Pr(\text{Piedi})}{\Pr(\text{Puntuale}|\text{Piedi}) \Pr(\text{Piedi}) + \Pr(\text{Puntuale}|\text{Auto}) \Pr(\text{Auto}) + \Pr(\text{Puntuale}|\text{Bus}) \Pr(\text{Bus})} \tag{7}$$

$$= \frac{0.85 \cdot 0.1}{0.85 \cdot 0.1 + 0.97 \cdot 0.4 + 0.9 \cdot 0.5} \approx 0.0921 \tag{8}$$

Problema 3

1. Abbiamo che

$$\mathsf{E}[Y] = \mathsf{E}[U] + \mathsf{E}[Z] = 0 + 0 = 0 \tag{9}$$

siccome le leggi di U e Z hanno simmetria pari. Inoltre, siccome $U \perp Z$, abbiamo che

$$\mathsf{Var}[Y] = \mathsf{Var}[U + Z] = \mathsf{Var}[U] + \mathsf{Var}[Z] = \frac{1}{3} + 1 = \frac{4}{3} \tag{10}$$

dove abbiamo usato $\mathsf{Var}[U] = \frac{2^2}{12} = \frac{1}{3}$ e $\mathsf{Var}[Z] = 1$.

2. Siccome Y e U hanno media nulla, abbiamo che

$$\mathsf{Cov}[Y, U] = \mathsf{E}[YU] = \mathsf{E}[(U + Z)U] = \mathsf{E}[U^2] + \mathsf{E}[UZ] \tag{11}$$

$$= \mathsf{E}[U^2] + \mathsf{E}[U] \cdot \mathsf{E}[Z] \tag{12}$$

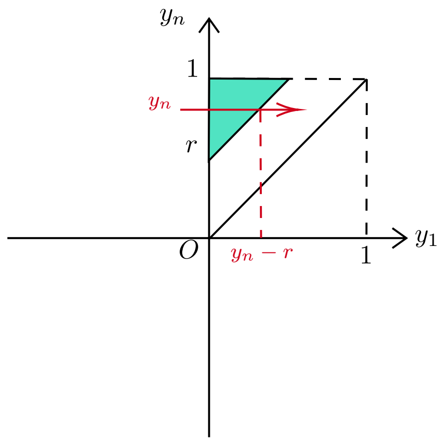
$$= \mathsf{Var}[U] = \frac{1}{3} \tag{13}$$

dove in (12) abbiamo usato l'indipendenza tra U e Z, e nell'ultimo passaggio abbiamo usato $\mathsf{E}[U] = 0$ per dire che $\mathsf{E}[U^2] = \mathsf{Var}[U]$.

3. Siccome U e Z sono indipendenti, il loro coefficiente di correlazione lineare è nullo.

Problema 4

Innanzitutto, si noti che $0 \leq R \leq 1$, e che è più comodo calcolare la cumulata di R tramite l'anticumulata:



$$\Pr(R \le r) = 1 - \Pr(R \ge r) = 1 - \Pr(Y_n - Y_1 \ge r) = 1 - \int\int_{r \le y_n - y_1 \le 1} f_{Y_1,Y_n}(y_1, y_n) dy_1 dy_n \tag{14}$$

$$= 1 - \int_r^1 \int_0^{y_n - r} n(n-1)(y_n - y_1)^{n-2} dy_1 dy_n, \tag{15}$$

dove l'evento di interesse $\{Y_n - Y_1 \ge r\}$ è stato colorato in azzurro in figura, e abbiamo visualizzato in rosso l'ordine di integrazione dell'integrale doppio. L'integrale si può calcolare come segue:

$$\int_r^1 \int_0^{y_n - r} n(n-1)(y_n - y_1)^{n-2} dy_1 dy_n = \int_r^1 \int_0^{y_n - r} n(n-1)(y_n - y_1)^{n-2} dy_1 dy_n \tag{16}$$

$$= \int_r^1 n(y_n^{n-1} - r^{n-1}) dy_n \tag{17}$$

$$= \left[y_n^n - n r^{n-1} y_n\right]_r^1 \tag{18}$$

$$= 1 - n r^{n-1} + (n-1) r^n. \tag{19}$$

Dunque abbiamo

$$F_R(r) = \begin{cases} 0 & r < 0 \\ n r^{n-1} - (n-1) r^n & 0 \le r \le 1 \\ 1 & r \ge 1 \end{cases} \tag{20}$$

e derivando si ottiene la densità di probabilità:

$$f_R(r) = \frac{d}{dr} F_R(r) = \begin{cases} n(n-1) r^{n-2} - n(n-1) r^{n-1} & 0 \le r \le 1 \\ 0 & \text{altrimenti} \end{cases} \tag{21}$$

Problema 5

1. Sia X la v.a. Gaussiana $\mathcal{N}(36.8, 0.4^2)$. Per rispondere al quesito, dobbiamo imporre

$$\Pr(36.8 - a \le X \le 36.8 + a) = 0.9 \tag{22}$$

dove a è il valore incognito da determinare. Procedendo con la standardizzazione di X, otteniamo

$$\Pr\left(-\frac{a}{0.4} \le \frac{X - 36.8}{0.4} \le \frac{a}{0.4}\right) = 0.9 \tag{23}$$

$$\Pr\left(-\frac{a}{0.4} \le Z \le \frac{a}{0.4}\right) = 0.9 \tag{24}$$

$$2 \Pr\left(0 \le Z \le \frac{a}{0.4}\right) = 0.9 \tag{25}$$

$$2\left(\Phi\left(\frac{a}{0.4}\right) - \frac{1}{2}\right) = 0.9 \tag{26}$$

$$\Phi\left(\frac{a}{0.4}\right) = 0.95 \tag{27}$$

$$\frac{a}{0.4} = \Phi^{-1}(0.95) \approx 1.645 \Longrightarrow a \approx 0.658 \tag{28}$$

dunque l'intervallo che contiene il 90% delle persone è $(36.142, 37.458)$.

2. Dobbiamo calcolare

$$
\begin{aligned}
\Pr(X > 37.5) &= \Pr\left(\frac{X - 36.8}{0.4} > \frac{37.5 - 36.8}{0.4}\right) && (29)\\
&= \Pr\left(Z > \frac{37.5 - 36.8}{0.4}\right) && (30)\\
&= 1 - \Phi(1.75) \approx 1 - 0.9599 = 0.0401, && (31)
\end{aligned}
$$

dunque la percentuale è 4.01%

Problema 6

All'aumentare di n ci aspettiamo che la densità di probabilità f_{X_n} si concentri attorno al valore $x = 0$, perché $(1-x)^{n-1}$ è una funzione decrescente in n per $0 < x < 1$. Al crescere di n la legge di probabilità si concentra sempre più attorno al valore $x = 0$. Dunque testiamo la convergenza in probabilità al valore $x = 0$:

$$
\begin{aligned}
\lim_{n\to\infty} \Pr(|X_n - 0| > \varepsilon) &= \lim_{n\to\infty} \Pr(X_n > \varepsilon) && (32)\\
&= \lim_{n\to\infty} \int_{\varepsilon}^{1} n(1-x)^{n-1} dx && (33)\\
&= \lim_{n\to\infty} -\left[(1-x)^n\right]_{\varepsilon}^{1} && (34)\\
&= \lim_{n\to\infty} (1-\varepsilon)^n && (35)\\
&= 0, \qquad \forall \varepsilon > 0, && (36)
\end{aligned}
$$

dove in (32) abbiamo usato la positività della v.a. X_n. Rimane dunque dimostrato che $X_n \xrightarrow{P} 0$.