Probabilità e Statistica (Informatica) 2023/24, Foglio I

4 marzo 2024

Sia $\boldsymbol{x} \doteq (x_i)_{i\in\{1,\ldots,n\}} \subset \mathbb{R}$ un campione di numerosità n. Per $k \in \mathbb{N}$, la quantità

$$\mu_k \doteq \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

si dice *momento campionario di ordine* k. Nota che $\mu_1 = \bar{x}$, la media campionaria di $\boldsymbol{x}$. Si dice *momento campionario centrato di ordine* k la quantità

$$\hat{\mu}_k \doteq \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^k .$$

Esercizio 1. Sia s^2 la varianza campionaria (come definita a lezione) di un campione $\boldsymbol{x} \doteq (x_i)_{i\in\{1,\ldots,n\}} \subset \mathbb{R}$. Si esprima s^2 prima in funzione dei momenti campionari del primo e secondo ordine, poi solo in funzione del momento campionario centrato del secondo ordine.

Esercizio 2. Si considerino i quattro campioni di numerosità 60 indicati con Tmin, Tmed, Tmax (temperature giornaliere minima, media e massima in gradi Celsius) e Ptot (quantità di pioggia in milimetri o, equivalentemente, litri al metro quadro) nel file `Meteo_Chioggia60.ods`. Si creino diagrammi di dispersione ("scatterplot") sul modello del foglio "DispGraf" del file per le seguenti coppie di campioni: Tmin versus Tmed, Tmed versus Tmax, Tmin versus Ptot, Tmed versus Ptot. Si dia un'interpretazione dei grafici confrontandoli con la tabella delle correlazioni campionarie (cf. foglio "Statistiche" del file).

Esercizio 3. (a) In quanti modi nove persone si possono disporre in fila a uno sportello?

(b) In quanti modi nove persone possono sedersi attorno a un tavolo rotondo? Consideriamo uguali due modi di sedersi se non cambiano i vicini di tavolo.

Esercizio 4. Immaginiamo di avere 200 cassetti abbastanza capienti, numerati da 1 a 200, e 300 palline *indistinguibili*. In quanti modi si possono distribuire le 300 palline sui 200 cassetti in maniera che ogni cassetto contenga almeno una pallina?

Esercizio 5. Immaginiamo di avere 100 cassetti, numerati da 1 a 100, e 5050 oggetti *distinti*. Si calcoli il numero di modi in cui gli oggetti possono essere disposti nei cassetti in modo che esattamente i oggetti finiscano nel cassetto i-esimo.

Esercizio 6. Sia $\Omega \neq \emptyset$. Interpretiamo sottoinsiemi di Ω come affermazioni. Siano $A, B \subseteq \Omega$. Usando operazioni insiemistiche si rappresentino le seguenti affermazioni:

(i) Non A.

(ii) A e B.

(iii) A o B.

(iv) A ma non B.

(v) O A o B.

(vi) Se A allora B.

(vii) Né A né B.

Esercizio 7. Supponiamo di estrarre a caso cinque carte (una "mano") da un mazzo di carte da poker. Vi sono quindi 52 carte determinate dal loro seme (picche, fiore, quadro, cuore) e dal loro tipo ($2, \ldots, 10$, J, Q, K, A). Le carte dal seme picche o fiore sono nere, le altre rosse. Si calcolino le probabilità delle seguenti combinazioni di cinque carte :

(i) almeno due carte dello stesso tipo;

(ii) un poker ("four of a kind"): quattro carte dello stesso tipo e una quinta carta;

(iii) un full ("full house"): tre carte di un tipo e due carte di un altro tipo;

(iv) un full con una sola carta rossa.

Esercizio 8. Distribuiamo le 52 carte di un mazzo da poker tra quattro giocatori A, B, C, D; ogni giocatore riceve quindi 13 carte. Si calcoli la probabilità che A o B (o entrambi) abbiano almeno due assi.

Esercizio 9. Il codice segreto di una carta di credito è formato da una sequenza di 4 cifre distinte tra 0 e 9.

(i) Effettuando un tentativo di indovinare il codice, qual è la probabilità di indovinarlo?

(ii) Qual è la probabilità di indovinare le 4 cifre che formano il codice, ma non necessariamente nel giusto ordine?

(iii) Qual è la probabilità di indovinare almeno una cifra del codice, ma non necessariamente nella giusta posizione?

(iv) Qual è la probabilità di indovinare almeno due cifre nella giusta posizione?

Esercizio 10. Sia Ω un insieme finito, e sia $H : \Omega \to \mathbb{R}$ una funzione. Per $\beta > 0$, definiamo una densità discreta su Ω attraverso

$$p_\beta(\omega) \doteq \frac{1}{Z_\beta} e^{-\beta H(\omega)}, \quad \omega \in \Omega,$$

dove Z_β è la costante di normalizzazione: $Z_\beta \doteq \sum_{\omega\in\Omega} e^{-\beta H(\omega)}$. Denotiamo con $\mathbf{P}_\beta$ la misura di probabilità su $\mathcal{P}(\Omega)$ indotta da p_β. Poniamo

$$A \doteq \{\omega \in \Omega : H(\omega) \leq H(\tilde{\omega}) \text{ per ogni } \tilde{\omega} \in \Omega\}.$$

Per ogni $\omega \in \Omega$, si determinino

$$\lim_{\beta\to\infty} \mathbf{P}_\beta(\{\omega\}) \qquad \text{e} \qquad \lim_{\beta\to 0+} \mathbf{P}_\beta(\{\omega\}).$$

Esercizio 11. Sia $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità, e siano $B \in \mathcal{F}$, $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$. Si verifichino le seguenti implicazioni:

(i) Se $\mathbf{P}(A_n) = 0$ per ogni $n \in \mathbb{N}$, allora $\mathbf{P}\left(\bigcup_{n=1}^\infty A_n\right) = 0$.

(ii) Se $\mathbf{P}(A_n) = 1$ per ogni $n \in \mathbb{N}$, allora $\mathbf{P}\left(\bigcap_{n=1}^\infty A_n\right) = 1$.

(iii) Se $\mathbf{P}(B) = 0$, allora $\mathbf{P}(B \cap A) = 0$ per ogni $A \in \mathcal{F}$.

(iv) Se $\mathbf{P}(B) = 1$, allora $\mathbf{P}(B \cap A) = \mathbf{P}(A)$ per ogni $A \in \mathcal{F}$.

Esercizio 12. Siano Ω_1, Ω_2 due insiemi finiti. Poniamo $\Omega \doteq \Omega_1 \times \Omega_2$. Sia $\mathbf{P}$ la distribuzione uniforme su Ω. Definiamo la funzione $Q\colon \mathcal{P}(\Omega_1) \to [0,1]$ tramite

$$Q(A) \doteq \mathbf{P}(A \times \Omega_2), \quad A \subseteq \Omega_1.$$

Si mostri che allora Q è la distribuzione uniforme su Ω_1.

Esercizio 13. Sia $\Omega \neq \emptyset$, e siano $A, B \subseteq \Omega$. Si costruisca la più piccola σ-algebra in Ω che contenga sia A che B.

Esercizio 14 ($\star$). Sia $(x_i, y_i)_{i\in\{1,\ldots,n\}}$ un campione bi-variato di numerosità n. Poniamo

$$\phi(a,b) \doteq \sum_{i=1}^{n} \left(y_i - (a \cdot x_i + b)\right)^2, \quad (a,b) \in \mathbb{R}^2.$$

Si scriva un programma che, dato $(x_i, y_i)_{i\in\{1,\ldots,n\}}$, calcoli un punto di minimo di ϕ, cioè un punto $(a_*, b_*) \in \mathbb{R}^2$ tale che

$$\phi(a_*, b_*) = \min\{\phi(a,b) : a, b \in \mathbb{R}\}.$$

Si considerino poi i due campioni bi-variati (Tmin, Tmed) e (Tmin, Ptot) del file `Meteo_Chioggia60.ods` e si crei, per ciascuno dei due campioni bi-variati, il corrispondente diagramma di dispersione col grafico della retta $t \mapsto a_* t + b_*$ determinata dal punto di minimo (a_*, b_*) calcolato col programma.

Per la consegna servono:

- una giustificazione matematica della procedura utilizzata per il calcolo di un punto di minimo di ϕ;
- lo pseudo-codice del programma e il codice commentato in un linguaggio standard come `C++` o Python;
- i grafici dei due diagrammi di dispersione con le rette di regressione in formato `pdf` e i valori numerici dei punti di minimo utilizzati.

Contatto: Markus Fischer (fischer@math.unipd.it)