

Prova scritta del 08.02.2018

(tempo: 2 ore per gli studenti di *Calcolo delle Probabilità e Statistica*;
2,5 ore per gli studenti di *Matematica per l'Analisi dei Dati*)

Esercizio 1

Di seguito sono riportati i dati relativi all'età (x) e all'altezza (y) per un campione di 21 bambini.

Bambini	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21
x	3	10	9	7	4	7	6	5	5	6	2	8	10	4	3	8	9	6	10	5	8
y	98	160	155	140	110	150	130	112	125	125	68	148	140	100	88	150	165	118	130	105	135

1. Determinare i valori dei tre quartili per i dati relativi all'altezza.
2. Disegnare, per i dati relativi all'altezza, un boxplot, avente come baffi il 10° e il 90° percentile.
3. Si considerino esclusivamente i dati relativi ai bambini B2, B9 e B15. Calcolare le medie e le varianze sia per i dati relativi all'età che per quelli relativi all'altezza.
4. Calcolare, sempre in riferimento ai tre bambini del punto precedente, il coefficiente di correlazione e valutare se esiste qualche correlazione tra i dati, spiegando di che tipo di correlazione si tratta.
5. Determinare la retta di regressione (sempre in riferimento ai suddetti tre bambini).

Esercizio 2

Un gioco è composto da due urne: l'urna U1 contenente 4 palline bianche e 6 gialle e l'urna U2 contenente 5 palline nere e 10 bianche. Il giocatore lancia due dadi a sei facce. Se la somma delle facce è un multiplo di 3, si utilizza l'urna U1. Se la somma delle facce è un multiplo di 5, si utilizza l'urna U2. Negli altri casi il giocatore rilancia nuovamente i dadi fino a quando non si verifica uno dei due casi precedenti.

1. Qual è la probabilità che il giocatore rilanci i dadi?
2. Qual è la probabilità di estrarre una pallina bianca?
3. Sapendo che è stata estratta una pallina bianca, qual è la probabilità che provenga dall'urna U2?
4. Sapendo che la somma delle facce dei due dadi è 9, qual è la probabilità che su 4 estrazioni (senza rimpiazzo) vengano estratte al massimo 2 palline bianche?
5. Dopo quanti lanci dei due dadi, ci si aspetta di iniziare l'estrazione?

Esercizio 3

Un programmatore per simulare il comportamento di una fabbrica produttrice di chiodi implementa un costruttore che crea un oggetto di tipo chiodo assegnandogli in modo casuale la stringa "non difettoso" con probabilità pari a 0.98 e la stringa "difettoso" altrimenti. A questo punto viene simulato un processo di produzione generando un array di chiodi di dimensione N.

1. Si assuma di generare un array di $N = 150$ chiodi. Qual è il valore atteso di chiodi non difettosi?
2. Qual è la probabilità che ci sia esattamente il 2% di chiodi difettosi?
3. Si vuole simulare un processo di produzione oraria. Vengono generati $N = 20000$ chiodi in un'ora. Qual è la probabilità che si siano generati più di 380 chiodi difettosi. Si scriva la formula esatta senza calcolarla. Si calcoli invece la probabilità utilizzando l'approssimazione normale.
4. Come dovrebbe variare la probabilità di generare un chiodo difettoso per fare in modo che la probabilità di ottenere più di 380 chiodi difettosi sia inferiore al 50%?
5. Si determini il più piccolo valore da assegnare ad N in modo da generare almeno 19000 chiodi non difettosi con probabilità superiore al 99.9%?

Esercizio 4 (Soltanto per gli studenti di *Matematica per l'Analisi dei Dati*)

Si consideri la seguente funzione:

$$f(x) = \begin{cases} 2x+a & \text{se } 0 < x \leq 0.5 \\ 0.5 & \text{se } 0.5 < x \leq 1 \\ -x+3a & \text{se } 1 < x \leq 1.5 \\ 0 & \text{altrimenti} \end{cases}$$

1. Determinare per quale valore del parametro a la funzione f rappresenta una densità di probabilità.
2. Determinare la funzione di ripartizione della variabile aleatoria associata ad f .
3. Calcolare il valore atteso della variabile aleatoria associata ad f .