Compito scritto dell'esame di Statistica e analisi dei dati 15 febbraio 2022	Prof. Giuseppe Boccignone	Corso di Laurea
Cognome:	Nome:	Matricola:

Istruzioni

- Il tempo riservato alla prova scritta e' di 2 ore. Durante la prova e' possibile consultare il formulario ed utilizzare la calcolatrice. Non e' possibile consultare libri e appunti.
- Ogni foglio deve riportare il numero di matricola
- In ogni esercizio occorre indicare chiaramente, per ogni risposta, il numero della domanda corrispondente
- Riportare lo svolgimento degli esercizi per esteso (quando l'esercizio richiede piu' passaggi di calcolo, non sara' preso in considerazione se riporta solo le soluzioni). Se una serie di calcoli coinvolge una o piu' frazioni semplici (numeratore e denominatore interi), per chiarezza, si svolgano i calcoli mantenendo tali numeri in forma frazionaria fin dove possibile (non li si converta nelle loro approssimazioni con virgola e decimali: solo il risultato finale sara' eventualmente rappresentato in quest'ultima forma).

Problemi

Esercizio 1. Un dado é truccato in modo che la probabilitá di ogni faccia sia proporzionale al suo punteggio. .

(a) Determinare se é maggiore la probabilitá che in un lancio esca un numero primo o un numero non primo

Soluzione: Si ha che

$$P(x) = kx, \quad x = 1, \cdots, 6.$$

Determiniamo k con la condizione di normalizzazione $\sum_{x=1}^{6} P(x) = 1$:

$$\sum_{x=1}^{6} kx = 1 \implies 21k = 1 \implies k = \frac{1}{21}$$

Definiamo l'evento

- $E =$ "esce numero primo" $= \{1,2,3,5\}$

Allora

$$P(E) = P(1) + P(2) + P(3) + P(5) = \frac{1+2+3+5}{21} = \frac{11}{21}$$

$$P(\sim E) = 1 - P(E) = \frac{10}{21} \implies P(\sim E) < P(E)$$

(b) Sapendo che é uscito un numero primo, calcolare la probabilitá che sia uscito 2.

Soluzione:

$$P(2 \mid E) = \frac{P(2 \cap E)}{P(E)} = \frac{P(2)}{P(E)} = \frac{2/21}{11/21} = \frac{2}{11}$$

Esercizio 2. La ditta di Bob produce e vende un componente elettronico la cui vita media é di 2.5 anni. La ditta si impegna, per garanzia, a sostituirlo se esso cessa di funzionare entro due anni.

(a) Qual é la probabilitá che la ditta debba intervenire per un singolo pezzo?

Soluzione: Supponiamo un tempo di vita T del prodotto che segue una legge esponenziale $T \sim Exp(\lambda) = \lambda e^{-\lambda t}, \quad t > 0$

Sappiamo che $E[T] = 2.5$ anni. Da questo ricaviamo λ

$$E[T] = \frac{1}{\lambda} = 2.5 \implies \lambda = 0.4$$

Quindi la probabilitá di intervenire entro i due anni di garanzia per sostituire il pezzo é

$$P(T \leq 2) = \int_0^2 0.4 e^{-0.4t} dt = 1 - e^{-0.8} = 0.55$$

(b) Che tempo di garanzia dovrebbe stabilire la ditta in modo da dover intervenire in non piú del 10% dei casi?

Soluzione: Si deve determinare il tempo t per cui $P(T \leq t) \leq 0.1$, ovvero, usando la cumulativa dell'esponenziale:

$$1 - e^{-\lambda t} = 1 - e^{-0.4t} \leq 0.1$$

Risolvendo la disequazione, si ricava:

$$t < 2.5 \ln \frac{10}{9} = 0.263$$

Esercizio 3. Sia data una variabile aleatoria X.

(a) Calcolare quanto vale al massimo la probabilitá che lo scarto di X dalla propria media valga almeno 2 deviazioni standard.

Soluzione: Lo scarto di X dalla sua media μ é $|X - \mu|$.
Si chiede di valutare la probabilitá $P_X(|X - \mu| \geq 2\sigma)$.
La distribuzione di X é ignota dunque si usa la disuguaglianza di Chebychev nella forma

$$P_X(|X - \mu| \geq 2\sigma) \leq \frac{\sigma^2}{(2\sigma)^2} = \frac{1}{4} = 25\%$$

Esercizio 4.
In un videogioco Alice ha a disposizione 6 tiri, ognuno dei quali colpisce a caso uno di due bersagli (A e B). Si vince se il bersaglio A é colpito almeno 4 volte su 6.

(a) Qual é la probabilitá che Alice vinca?

Soluzione: (Distribuzione binomiale)
Ogni tiro (indipendente) ha successo (colpisce il bersaglio) con probabilitá che vale $p = 1/2$.
Usando il modello binomiale $X \sim Bin\left(n = 6, p = \frac{1}{2}\right)$ la probabilitá di vittoria é semplicemente

$$P(X \geq 4) = \sum_{x=4}^{6} \binom{6}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{6-x} = \sum_{x=4}^{6} \binom{6}{x} \left(\frac{1}{2}\right)^6 = \frac{11}{32} \approx 34.37\%$$

(b) Se la partita costa 1 euro e in caso di vittoria si ricevono 3 euro, quanto puó vincere in media Alice?

Soluzione:
Sia Y la v.a. che misura il guadagno di Alice. Allora Y si distribuisce come segue:

- in caso di vittoria: $P(Y = 3 - 1) = P(Y = 2) = \frac{11}{32}$, come determinato sopra;
- in caso di sconfitta: $P(Y = -1) = 1 - \frac{11}{32} = \frac{21}{32}$.

Il guadagno atteso di Alice é

$$E[Y] = 2 \times \frac{11}{32} - 1 \times \frac{21}{32} = \frac{1}{32}$$

Esercizio 5. Uno strumento produce mine per le matite di una certa marca con una deviazione standard della lunghezza delle mine pari a $\sigma = 0.1$ cm. Misurando dieci mine, si sono ottenuti i seguenti valori (in cm)

$$12.21, 12.33, 12.84, 12.97, 13.22, 12.93, 13.07, 13.52, 13.23, 13.01$$

(a) Determinare un intervallo di confidenza al 95% per la lunghezza media delle mine.

Soluzione:
Assumiamo il modello gaussiano per la lunghezza X delle mine: $X \sim \mathcal{N}(\mu, (0.1)^2)$ (σ nota).
Sappiamo che $n = 10$ e la media campionaria delle misure vale $\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = 12.933$.
L' IC al 95% si ottiene per

$$\begin{aligned} (100)(1 - \alpha) &= 95 \\ 1 - \alpha &= 0.95 \\ \alpha &= 0.05 \\ \alpha/2 &= 0.025 \end{aligned}$$

Il valore critico $z_{0.025}$ é il valore che lascia un'area a destra pari a 0.025 (a sinistra pari a $1 - 0.025 = 0.975$).

Usando la tabella della Normale standard lo z-value che lascia a destra 0.025 é $z_{0.025} = 1.96$.
L'IC cercato é:

$$12.933 - 1.96\frac{0.1}{\sqrt{10}} < \mu < 12.933 + 1.96\frac{0.1}{\sqrt{10}}$$

$$12.933 - 0.062 < \mu < 12.933 + 0.062$$

$$12.871 < \mu < 12.995$$

(b) Mantenendo il medesimo livello di confidenza, quante mine occorrerebbe misurare per avere un intervallo non piú ampio di 0.1 cm?

Soluzione:
La condizione che l'ampiezza dell'intervallo di confidenza al 95% non superi 0.1 cm, richiede che

$$2 \times 1.96\frac{0.1}{\sqrt{n}} \leq \frac{1}{10} \implies n \geq (2 \times 1.96)^2 \approx 15.4$$

Esercizio 6.
Bob ha calcolato un intervallo di confidenza per la media di una popolazione normale, a partire da un campione di 8 osservazioni, dopo averne stimato la varianza con la varianza campionaria $s^2 = 4.2$. Il risultato é $(16.08, 18.82)$.

(a) Con quale confidenza é stato ottenuto l'intervallo?

Soluzione:
La taglia del campione $n = 8$ é piccola e la varianza é ignota essendo stata stimata mediante la varianza campionaria $s^2 = 4.2$. Inoltre, la popolazione da cui proviene il campione é normale. Pertanto l'IC per la media deve essere stato stimato utilizzando la distribuzione t di Student con $\nu = n - 1 = 7$ gradi di libertá:

$$\mu = \bar{x} \pm \frac{s}{\sqrt{n}} t_{\frac{\alpha}{2}}(\nu = 7) = \bar{x} \pm \frac{\sqrt{4.2}}{\sqrt{8}} t_{\frac{\alpha}{2}}(\nu = 7)$$

Gli estremi dell'IC indicato corrispondono a

$$\bar{x} - \frac{\sqrt{4.2}}{\sqrt{8}} t_{\frac{\alpha}{2}}(\nu = 7) = 16.08$$

$$\bar{x} + \frac{\sqrt{4.2}}{\sqrt{8}} t_{\frac{\alpha}{2}}(\nu = 7) = 18.82$$

Risolvendo il sistema a due incognite si ottiene

$$\bar{x} = 17.45$$

$$t_{\frac{\alpha}{2}}(\nu = 7) = 1.895$$

Dalla tabella della t di Student con $\nu = 7$ leggiamo che $t_{\frac{\alpha}{2}} = 1.895$ per $\frac{\alpha}{2} = 0.05$, cioé $\alpha = 0.10$
Dunque l'intervallo é stato ottenuto con una confidenza pari al $(100)(1 - \alpha) = (100)(1 - 0.10) = 90\%$