Compito scritto dell'esame di Statistica e analisi dei dati 30 giugno 2021	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. In un recente sondaggio sul sindaco di una certa cittá, il 62% dei rispondenti ha fiducia nel sindaco. Le donne costituiscono il 53% del campione, e tra queste il 46% ha fiducia nel sindaco.

(a) Si seleziona a caso una persona tra quelle intervistate. Qual é la probabilitá che la persona selezionata sia maschio?

Soluzione: Esercizio banale se si filtrano i dati che non servono. Gli unici eventi che ci interessano sono

- $D =$ "rispondente donna"
- $M =$ "rispondente maschio"

Sappiamo che $P(D) = 0.53$. Dunque:

$$P(M) = 1 - P(D) = 1 - 0.53 = 0.47$$

Esercizio 2. La tabella riporta le probabilitá congiunte di un insieme di nuove aziende che operano nel settore del

Table 1: Tabella aziende

Crescita	Nord-Est	Sud	Centro	Nord-Ovest
Bassa	0.04	0.12	0.14	0.19
Media	0.05	0.08	0.06	0.12
Alta	0.03	0.05	0.08	0.04

commercio via internet, classificate per regione di ubicazione e prospettiva di crescita.

(a) Se l'azienda ha una crescita attesa media o alta, qual é la probabilitá che sia ubicata nel Nord-Ovest?

Soluzione:
Indichiamo con M l'evento "l'azienda ha una crescita media", con A l'evento "l'azienda ha una crescita alta" e con NO l'evento "azienda ubicata nel Nord-Ovest".
Si tratta di determinare la probabilitá condizionata $P(NO \mid (M \cup A))$:

$$P(NO \mid (M \cup A)) = \frac{P(NO \cap (M \cup A)}{P(M \cup A)} \tag{1}$$

Essendo A, M disgiunti:

$$P(M \cup A) = P(M) + P(A)$$

Dalla tabella abbiamo le seguenti probabilitá marginali. :

$$P(M) = 0.05 + 0.08 + 0.06 + 0.12 = 0.31$$

$$P(A) = 0.03 + 0.05 + 0.08 + 0.04 = 0.20$$

Dunque

$$P(M \cup A) = P(M) + P(A) = 0.31 + 0.20 = 0.51.$$

Usando la proprieta' distributiva dell'intersezione rispetto all' unione, la disgiunzione di A, M, e leggendo in tabella

$$P(NO \cap M) = 0.12$$

e

$$P(NO \cap A) = 0.04$$

si ottiene:

$$P(NO \mid (M \cup A)) = \frac{P(NO \cap M) + P(NO \cap A)}{P(M \cup A)} = \frac{0.12 + 0.04}{0.51} = 0.31. \tag{2}$$

Esercizio 3. Una fabbrica di cioccolato promuove una campagna pubblicitaria per la vendita di un nuovo tipo di uova pasquali, caratterizzate dal contenuto della sorpresa, costituita da bracciali di ottone. Su una partita di 1000 uova, 5 di esse contengono dei bracciali d'oro, del valore di 100000 euro cadauno. Le uova contenenti bracciali d'oro, a loro volta, vengono inserite in modo casuale in scatole da 20, e vendute ai negozianti.

(a) Qual é la probabilitá che un negoziante che acquista una scatola di uova, venda al pubblico una o piú uova contenenti bracciali d'oro?

Soluzione:
La risposta richiederebbe l'uso della distribuzione binomiale $Bin(n,p)$, con parametri $n = 20$ e $p = 5/1000 = 0.005$. In tal caso, il numero atteso di bracciali d'oro per scatola é dato da $np = 20 \times 0.005 = 0.1$. Tuttavia, in questo caso la distribuzione di Poisson di parametro $\mu = np$ puó essere una buona approssimazione della binomiale, essendo $p = 0.005$ molto piccolo, e semplifica il calcolo della probabilitá di avere $1, 2, \cdots$ uova contenenti bracciali d'oro. Pertanto:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - Pois(X = 0) = 1 - \frac{0.1^0}{0!} e^{(-0,1)} = 1 - 0.904837418 = 0.09516258$$

(con la Binomiale si otterrebbe 0.0953895197)

Esercizio 4. Una variabile aleatoria X, con valori nell'intervallo $1 \leq x \leq 2$, ha densitá

$$f_X(x) = \frac{k}{x^2}$$

.

(a) Calcolare valore atteso $E[X]$ e deviazione standard σ_X

Soluzione: Si calcola la costante k utilizzando la proprietá di normalizzazione:

$$1 = \int_1^2 f_X(x)dx = \int_1^2 \frac{k}{x^2} dx = k \left[-\frac{1}{x} \right]_1^2 = k(-\frac{1}{2} + 1) = \frac{k}{2},$$

da cui $k = 2$

$$E[X] = \int_1^2 x \frac{2}{x^2} dx = 2 \log 2 \approx 1.386$$

$$\sigma_X^2 = E\left[X^2\right] - E[X]^2 = E\left[X^2\right] - 4 \log^2 2 = 2 \int_1^2 dx - 4 \log^2 2 = 2(1 - 2 \log^2 2) \approx 0.078$$

$$\sigma_X = \sqrt{0.078} \approx 0.28.$$

Esercizio 5. Una popolazione assume un farmaco per mantenere il valor medio del battito cardiaco (battito/min, bpm) tra 85 e 95 bpm. Il bpm medio rilevato su un campione di 49 soggetti é pari a 90 bpm. Da dati storici, si sa che la variazione del bpm della popolazione segue una legge normale con deviazione standard $\sigma = 10$.

(a) Si verifichi se, relativamente al bpm medio, il farmaco possa ritenersi efficace con livello di confidenza al 90% e 95%

Soluzione:
Sappiamo che $n = 49$ e $\bar{x} = 90$.
L' IC al 90% si ottiene per

$$\begin{aligned}(100)(1-\alpha) &= 90\\ 1-\alpha &= 0.9\\ \alpha &= 0.1\\ \alpha/2 &= 0.05\end{aligned}$$

Il valore critico $z_{0.05}$ é il valore che lascia un'area a destra pari a 0.05 (a sinistra pari a $1 - 0.05 = 0.95$).
Usando la tabella della Normale standard lo z-value che lascia a destra 0.05 é $z_{0.05} = 1.64$.
L'IC cercato é:

$$90 - 1.64\frac{10}{\sqrt{49}} < \mu < 90 + 1.64\frac{10}{\sqrt{49}}$$
$$87.66 < \mu < 92.34$$

Analogamente, per l'IC al 95%, $\alpha = 0.05$ e $z_{\alpha/2} = 1.96$.
Pertanto,

$$90 - 1.96\frac{10}{\sqrt{49}} < \mu < 90 + 1.96\frac{10}{\sqrt{49}}$$
$$87.2 < \mu < 92.8$$

Sulla base di tali risultati il farmaco si puó ritenere efficace.

Esercizio 6. Un produttore di batterie per auto garantisce che il suo prodotto dura in media 3 anni con una deviazione standard di 1 anno. Si campionano casualmente 5 batterie e ne risulta che abbiano una durata di $1.9, 2.4, 3.0, 3.5, 4.2$ anni.

(a) Assumendo che la vita della popolazione di batterie sia distribuita con legge normale, si tragga una conclusione con un livello di confidenza al 95% sull'asserzione del produttore che $\sigma = 1$.

Soluzione:
Utilizziamo l'intervallo di confidenza

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$$

con $\sigma^2 = \sigma = 1$
Determiniamo la varianza del campione s^2 con

$$s^2 = \frac{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)}$$

Per $n = 5$:

$$\sum_{i=1}^5 x_i = 1.9 + 2.4 + 3.0 + 3.5 + 4.2 = 15$$

e

$$\sum_{i=1}^5 x_i^2 = 1.9^2 + 2.4^2 + 3.0^2 + 3.5^2 + 4.2^2 = 48.26$$

Pertanto

$$\begin{aligned}s^2 &= \frac{n\sum_{i=1}^5 x_i^2 - (\sum_{i=1}^5 x_i)^2}{n(n-1)}\\ &= \frac{5 \cdot 48.26 - (15)^2}{5(5-1)}\\ &= \frac{16.3}{20}\\ &= 0.815\end{aligned}$$

Determiniamo $\alpha/2$. Da $95\% = 100(1-\alpha)\%$, si ha:

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

I g.d.l. sono $\nu = n-1 = 4$. Usando la tabella della distribuzione Chi-quadro, il valore critico per 0.025 con 4 g.d.l. é $\chi^2_{0.025} = 11.143$, e $\chi^2_{1-0.025} = \chi^2_{0.975} = 0.484$.
L'intervallo di interesse é

$$\begin{aligned}\frac{(n-1)s^2}{\chi^2_{\alpha/2}} &< \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}\\ \frac{(5-1)\cdot 0.815}{\chi^2_{0.025}} &< \sigma^2 < \frac{(5-1)\cdot 0.815}{\chi^2_{0.975}}\\ \frac{4\cdot 0.815}{11.143} &< \sigma^2 < \frac{4\cdot 0.815}{0.484}\\ 0.29 &< \sigma^2 < 6.74\end{aligned}$$

Poiché 1 cade nell'intervallo di confidenza al 95% si puó concludere che quanto dichiarato dal produttore é plausibile statisticamente.