Compito scritto dell'esame di Statistica e analisi dei dati 22 febbraio 2024	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, appunti, cellulari.
- 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. Aldo ha una moneta di probabilitá p che in un lancio esca Testa. Aldo fa questo gioco: lancia la moneta e se esce Testa il gioco finisce; altrimenti, lancia la moneta una seconda volta e se esce Testa il gioco finisce; altrimenti, lancia la moneta per la terza e ultima volta.

(a) Descrivere lo spazio campionario e la funzione di probabilitá associati all'esperimento.

Soluzione:
Posto:

- $T_1 =$ `esce Testa al primo lancio`,
- $C_1 =$ `esce Croce al primo lancio`
- ecc. $\cdots$

si ha

$$P(T_i) = p, \quad P(C_i) = 1 - p = q, \quad i = 1{,}2{,}3.$$

Allora i possibili esiti ω dell'esperimento e le corrispondenti probabilitá sono:

$$P(\omega = T_1) = p, \quad P(\omega = C_1 \cap T_2) = qp \quad P(\omega = C_1 \cap C_2 \cap T_3) = q^2 p \quad P(\omega = C_1 \cap C_2 \cap C_3) = q^3$$

Per la $P(\omega)$, dato che $p + q = 1$, vale la condizione di normalizzazione $\sum_\omega P(\omega) = p + qp + q^2 p + q^3 = 1$ e dunque é distribuzione ammissibile.

(b) Calcolare la probabilitá che Aldo lanci la moneta tre volte.

Soluzione:
Si ha che la probabilitá che Aldo lanci la moneta tre volte é

$$P(\texttt{tre lanci}) = P(\omega = C_1 \cap C_2 \cap T_3) + P(\omega = C_1 \cap C_2 \cap C_3) = q^2 p + q^3 = q^2$$

Esercizio 2.
Una linea di produzione fabbrica resistori da 1000 ohm (Ω). La fabbrica adotta una tolleranza del 10% (ovvero sono scartati i resistori con resistenza $R < 900\ \Omega$ e $R > 1100\ \Omega$).

(a) Sotto l'ipotesi che la resistenza R di un resistore sia una variabile aleatoria distribuita con legge normale di media $1000\ \Omega$ e varianza 2500, si calcoli la probabilitá che un resistore preso a caso venga scartato

Soluzione: Sia S l'evento in cui un resistore viene scartato. Allora:

$$S = \{R < 900\} \cup \{R > 1100\}$$

Poiché $\{R < 900\} \cap \{R > 1100\} = \emptyset$, si ha che

$$P(S) = P(\{R < 900\}) + P(\{R > 1100\}) = F_X(900) + (1 - F_X(1100)) = \Phi\left(\frac{900 - \mu}{\sigma}\right) + \left(1 - \Phi\left(\frac{1100 - \mu}{\sigma}\right)\right)$$

Poiché $\mu = 1000$, e $\sigma^2 = 2500 \implies \sigma = 50$, le variabili normalizzate sono

$$\frac{900 - 1000}{50} = -2$$

$$\frac{1100 - 1000}{50} = 2$$

.
Sfruttando la proprietá della cumulativa standardizzata $\Phi(-z) = 1 - \Phi(z)$, la precedente diventa:

$$P(S) = 2\,(1 - \Phi(2))$$

.
Accedendo alla tabella della Normale standard per $z = 2$, $\Phi(2) = 0.9772$ da cui $P(S) = 0.0455$

Esercizio 3. Si vuole studiare la durata di un processo produttivo che dal materiale grezzo porta al prodotto finito. Il venditore del meccanismo di produzione sostiene che la durata del processo si distribuisce normalmente con media pari a 11 ore e deviazione standard pari a 4 ore. L'acquirente, sulla base delle valutazioni di un esperto, sospetta invece che, pur distribuendosi normalmente e con deviazione standard 4, la durata media del processo sia 14 ore. Si mettono allora in produzione 16 pezzi e si decide che il meccanismo di produzione verrá acquistato soltanto se la durata media della produzione nel campione é inferiore a 13.

(a) Si calcoli la probabilità di rischio dell'acquirente ovvero dell'errore del I tipo associato al criterio di decisione sopra riportato

Soluzione: La durata T é distribuita con legge normale $T \sim \mathcal{N}(\mu_0 = 11, \sigma^2 = 4^2)$.
Siccome, fino a prova contraria, non si puó sostenere che il venditore stia barando, il problema si puó mettere nella forma di un test di ipotesi unilaterale

$$\begin{cases} H_0 & : \mu = 11 = \mu_0 \\ H_1 & : \mu > 11 (\mu_1 = 14). \end{cases} \tag{1}$$

Dalla teoria dei test di ipotesi sappiamo che il rischio ovvero la probabilitá di errore di tipo I associato al criterio di decisione é

$$P(\texttt{rifiuto } H_0 \mid H_0 \texttt{ vera}) = \alpha.$$

Considerato il criterio di decisione $\bar{T} < 13$, la regione di rifiuto di H_0 é $\bar{T} \geq 13$. Nell'ipotesi che H_0 sia vera, $\mu = \mu_0 = 11$; dunque:

$$\alpha = P(\texttt{rifiuto } H_0 \mid H_0 \texttt{ vera}) = P(\bar{T} \geq 13 \mid \mu = \mu_0)$$

Standardizzando e usando le tavole della distribuzione normale standard:

$$\alpha = P(\bar{T} \geq 13 \mid \mu = \mu_0) = P(\frac{\bar{T} - \mu}{\sigma/\sqrt{n}} \geq \frac{13 - \mu}{\sigma/\sqrt{n}} \mid \mu = \mu_0) = P(\frac{\bar{T} - \mu_0}{\sigma/\sqrt{n}} \geq \frac{13 - \mu_0}{\sigma/\sqrt{n}}) = P(Z \geq \frac{13 - 11}{4/\sqrt{16}}) = P(Z \geq 2) \simeq 0.0228.$$

(Analogamente, si poteva ragionare sul valore critico z_α che determina la regione critica $P(Z \geq z_\alpha)$ con

$$Z = \frac{\bar{T} - 11}{4/\sqrt{16}} \geq z_\alpha \quad \implies \quad \bar{T} \geq 11 + z_\alpha \frac{4}{\sqrt{16}} \tag{2}$$

sotto la condizione che l'acquirente accetta un rischio a patto che $\bar{T} < 13$, ovvero con valore critico determinato dall'equazione

$$11 + z_\alpha \frac{4}{\sqrt{16}} = 13; \tag{3}$$

dunque $z_\alpha = \frac{13-11}{4/\sqrt{16}} = 2$ come sopra)

(b) Si valuti la potenza del test associato al criterio di decisione sopra riportato

Soluzione:
La potenza del test é definita nella teoria dei test come complemento della probabilitá β dell'errore di tipo II

$$P = 1 - \beta = 1 - P(\texttt{non rifiuto } H_0 \mid H_0 \texttt{ falsa}) = P(\texttt{rifiuto } H_0 \mid H_1 \texttt{vera}).$$

Nell'ipotesi che H_1 sia vera, $\mu = \mu_1 = 14$. Pertanto

$$\beta = P(\bar{T} \leq 13 \mid \mu = \mu_1) = P(\frac{\bar{T} - \mu}{\sigma/\sqrt{n}} \leq \frac{13 - \mu}{\sigma/\sqrt{n}} \mid \mu = \mu_1) = P(\frac{\bar{T} - \mu_1}{\sigma/\sqrt{n}} \leq \frac{13 - \mu_1}{\sigma/\sqrt{n}}) = P(Z \leq \frac{13 - 14}{4/\sqrt{16}}) = P(Z \leq -1) \simeq 0.1587.$$

Dunque la potenza é notevole, pari a $P = 1 - \beta \simeq 84{,}13\%$

Esercizio 4. Ogni settimana sulla ruota del Lotto di Milano vengono estratti in blocco 5 numeri tra i 90 a disposizione. Nel corso di un anno vengono effettuate 52 estrazioni.

(a) Calcolare la probabilitá che nel corso di un anno il numero 47 esca almeno due volte sulla ruota di Milano

Soluzione: Si verifica il "successo" quando nel blocco di 5 numeri esce il 47. Ne segue che la variabile X che conta il numero di successi nel corso di un anno ha distribuzione binomiale

$$P(X = x) = Bin(x \mid n = 52, p)$$

dove la probabilitá p di successo nell'estrazione settimanale si calcola con la distribuzione ipergeometrica

$$p = \frac{\binom{89}{4})\binom{1}{1}}{\binom{90}{5}} = \frac{1}{18}.$$

Pertanto,

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \sum_{x=0}^{1} \binom{52}{x} \left(\frac{1}{18}\right)^x \left(\frac{17}{18}\right)^{52-x} \approx 79.22\%$$

Esercizio 5. Le telefonate in arrivo ad un centralino seguono un processo di Poisson con una media di 5 telefonate/minuto in entrata.

(a) Qual é la probabilitá che trascorra fino a un minuto dal momento che sono arrivate due telefonate al centralino?

Soluzione:
Dalla teoria sappiamo che il processo descrive una situazione in cui l'intervallo di tempo al verificarsi di due eventi di Poisson segue una distribuzione $Gamma(x \mid \alpha, \beta)$ con

$$\beta = \frac{1}{5} \qquad \alpha = 2.$$

Sia dunque X la V.A. che denota il tempo (minuti) che trascorre prima di osservare 2 telefonate in arrivo. Allora,

$$P(X \leq 1) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^1 x^{\alpha-1} e^{-\frac{x}{\beta}} dx = 25 \int_0^1 x e^{-5x} dx = 1 - e^{-5}(1+5) \approx 0.96$$

dove l'ultimo passaggio si é ottenuto integrando per parti.

Esercizio 6. I punteggi di un test d'ingresso posto alle matricole di un college negli ultimi cinque anni sono distribuiti con legge normale a media $\mu = 74$ e varianza $\sigma^2 = 8$.

(a) Si può considerare ancora $\sigma^2 = 8$ un valore di varianza valido se un campione casuale di 25 studenti che eseguono il test d'ingresso quest'anno ottenesse un valore di $S^2 = 15$?

Soluzione: Se S^2 è la varianza di un campione casuale di dimensione n preso da una popolazione normale avente varianza σ^2, allora la statistica:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

ha una distribuzione chi-quadro con $v = n - 1$ gradi di libertà.
Quindi,

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{24 \cdot S^2}{8} = 3 \cdot S^2$$

ha una distribuzione chi-quadro con 24 gradi di libertà.
Verificheremo se $\sigma^2 = 8$ è un'assunzione ragionevole determinando la probabilità

$$P(S^2 > 15) = P(\chi^2 > 3 \cdot 15) = P(\chi^2 > 45) < P(\chi^2 > 42.980) \approx 0.01$$

quindi, abbiamo trovato che sotto l'ipotesi $\sigma^2 = 8$, c'è meno di 1% di possibilità che $P(S^2 > 15)$, quindi l'ipotesi non è ragionevole.