MASTER E^{2}C - Corso di STATISTICA
Prova finale - 2.2.2004

Esercizio 1. In un'urna sono contenute dieci palline numerate da 1 a 10. Vengono effettuate due estrazioni senza rimpiazzo. Si considerino le variabili aleatorie

$$X_i = \{\text{numero della pallina estratta all' } i\text{-esima estrazione}\}, \quad i = 1, 2.$$

a) Determinare le probabilità degli eventi $\{X_1 = i, X_2 = j\}$, per $i, j = 1, \ldots, 10$.

b) Determinare la probabilità dell'evento $\{X_1 + X_2 = 5\}$.

c) Determinare la probabilità dell'evento $\{X_2$ è pari$\}$. C'è differenza con la probabilità dell'evento $\{X_1$ è pari$\}$?

d) Determinare la probabilità dell'evento $\{X_2 = 9\}$, sapendo che la prima estrazione ha dato un esito maggiore di 5 (ovvero $X_1 > 5$).

Esercizio 2. Si consideri la variabile aleatoria X, la cui funzione di densità della probabilità è definita da:

$$f_X(x) = \begin{cases} \alpha & \text{se } \ 0 \le x < \theta \\ \frac{1}{2}\alpha & \text{se } \ \theta \le x \le 1 \\ 0 & \text{altrimenti} \end{cases}$$

dove θ è un parametro incognito compreso tra 0 e 1, e α è un coefficiente da determinare.

a) Determinare il valore di α in modo che la funzione $f_X(x)$ sia effettivamente una densità di probabilità.

b) Calcolare e disegnare la funzione di distribuzione della probabilità, $F_X(x)$.

c) Assumendo $\theta = \frac{1}{2}$, calcolare media e varianza della variabile aleatoria X.

d) Determinare la stima a massima verosimiglianza del parametro θ e dire se tale stima è corretta o polarizzata.

Esercizio 3. Siano a e b due parametri incogniti, e si considerino le misure:

$$\begin{aligned} y_1 &= a + b + e_1 \\ y_2 &= a - b + e_2 \end{aligned}$$

dove e_1 e e_2 sono realizzazioni delle variabili aleatorie indipendenti $E_1 \sim \mathcal{N}(0, \sigma_1^2)$, $E_2 \sim \mathcal{N}(0, \sigma_2^2)$. Siano inoltre Y_1 e Y_2 le variabili aleatorie associate alle misure y_1, y_2.

a) Determinare la funzione di densità di probabilità della variabile aleatoria $Y = Y_1 + Y_2$. Supponendo di disporre della sola misura $y = y_1 + y_2$, calcolare la stima di massima verosimiglianza del parametro a basata su y, $\hat{a} = T_{ML}(y)$.

b) Determinata la funzione di densità di probabilità congiunta delle variabili aleatorie Y_1 e Y_2, calcolare la stima di massima verosimiglianza del parametro a basata sulle misure y_1 e y_2, $\hat{a} = T_{ML}(y_1, y_2)$. C'è differenza tra le stime calcolate ai punti *a*) e *b*)?

c) Determinare la stima ai minimi quadrati $\hat{a}_{LS}$, $\hat{b}_{LS}$ dei parametri a e b.

d) Determinare se la stima $\hat{a}$ calcolata al punto *a*) è corretta, e calcolarne la varianza. Dimostrare che $\hat{a}$ è la miglior stima lineare non polarizzata (BLUE) di a, basata su y_1 e y_2.

Esercizio 4.
Il direttore finanziario della società di spedizioni FedEx ha escogitato il seguente sistema, per ridurre i tempi medi entro cui i propri clenti pagano le fatture mensili. Egli ha maturato la convinzione che includere nelle fatture inviate mensilmente ai propri clienti, una busta con già stampato l'indirizzo a cui spedire i pagamenti, potrebbe portare ad una riduzione dei ritardi che attualmente si riscontrano. Con il sistema di fatturazione in uso, in media i pagamenti avvengono entro 24 giorni, con una deviazione standard di 6 giorni, ed hanno una distribuzione assimilabile ad una Gaussiana.
È stato calcolato che ridurre i ritardi di pagamento ad una media di 22 giorni, consentirebbe di coprire i costi associati all'invio della busta prestampata. Per verificarne l'efficacia, il nuovo sistema è stato provato su un campione casuale di 220 clienti. Dall'indagine effettuata, il tempo medio entro cui i clienti hanno effettuato i pagamenti è risultato essere pari a 21.63 giorni.
Supponendo che la massima probabilità accettabile di introdurre inutilmente il nuovo sistema di fatturazione sia pari a 0.1, e che la deviazione standard dei tempi di pagamento rimanga invariata, si utilizzi un test per valutare la significatività statistica della apparente riduzione dei ritardi. In particolare si specifichino:

- l'ipotesi nulla $\mathbf{H_0}$ e l'ipotesi alternativa $\mathbf{H_1}$;
- la statistica di test da utilizzare;
- il livello di significatività α da adoperare;
- l'espressione della regione di rifiuto;
- l'esito del test.