Esercizi di Inferenza Statistica
Blocco IV
a.a. 2024 – 2025

1. Si vuole valutare se ci siano differenze nel numero medio di ore mensili, X, spese al cellulare tra adolescenti che vivono in aree rurali (R) e in città (C). I dati relativi a due campioni sono riportati di seguito, dove con σ^2 si denota la varianza (supposta nota) nella popolazione:

	n	$\sum_{i=1}^n x_i$	σ^2
R	39	2330.6	25
C	33	2316.3	27

Si assuma che i due campioni siano indipendenti e provenienti da popolazioni normali.

a. Si fornisca un intervallo di confidenza al livello 0.95 per la differenza fra le medie del numero di ore spese al cellulare tra adolescenti che vivono in aree rurali e in città essendo le variante note e pari a quelle in tabella;

b. Che cosa si può dire con un grado di fiducia del 99% circa l'errore massimo nella stima della differenza fra le medie immaginando le varianze ignote e uguali e supponendo che le varianze campionarie (corrette) siano pari a a 21.5 per R e 23.3 per C?

2. La probabilità che un albicocco sia produttivo è pari a p. In un campione casuale di 200 piante si trovano 170 alberi che hanno dato frutti.

a. Si determini un intervallo di confidenza per p a livello del 95%;

b. Si determini il numero minimo di alberi da esaminare affinché, con probabilità del 98%, l'errore della stima di p sia inferiore al 2%?

c. Se si ripete il controllo per altri 150 terreni scegliendo per ogni terreno un campione di 200 alberi e contando quanti di essi hanno dato frutti, quanti tra i 150 intervalli di confidenza a livello del 95% ci si attende coprano il vero valore di p?

3. In uno studio farmacologico si vuole stimare il tasso di guarigione, p, connesso con l'uso di un nuovo medicinale. Su 500 pazienti che hanno seguito la nuova terapia è stata riscontrata la guarigione in 280 individui. Sia noto che il medicinale in uso in precedenza garantisce una proporzione di guarigione pari a 0.5. Si adotti una procedura statistica appropriata e si verifichi sulla base dei dati raccolti se il nuovo medicinale è più efficace di quello standard.

4. Un carattere X ha presso una popolazione distribuzione normale con media μ incognita e varianza $\sigma^2 = 4$. Un campione casuale è estratto dalla popolazione e fissato il livello di significatività $\alpha = 0.1$, si vuole verificare il sistema di ipotesi: $H_0 : \mu = 3$ contro $H_1 : \mu = 3.5$. Determinare l'ampiezza campionaria n necessaria affinché la potenza del test sia pari a 0.5.

5. Una macchina è calibrata per produrre listelli di legno di lunghezza pari a 10 cm. Viene acquistata una nuova macchina, e si vuole verificare se essa, a parità di calibrazione, sia più precisa. Pertanto si analizzano un campione di $n = 10$ misurazioni sulle lunghezze dei listelli da entrambe le macchine ottenendo i seguenti risultati

Macchina	n	$\sum_{i=1}^n x_i$	$\sum_{i=1}^{10} x_i^2$
Vecchia	10	100.93	1021.53
Nuova	10	100.56	1011.67

Assumendo la normalità per entrambe le popolazioni, si verifichi al livello del 5% se la seconda macchina è più precisa, cioè se la varianza delle lunghezze dei pezzi prodotti può ritenersi minore.

6. In due stabilimenti di un'azienda sono stati registrati per alcuni giorni i rispettivi numeri di pezzi prodotti:

stabilimento A	83	81	80	84	75	80	83	91	83	92
stabilimento B	102	92	91	82	101	99	95	88	91	

Assumendo che il numero di pezzi sia distribuito secondo una normale con varianze ignote e uguali.

a. Si ottenga un intervallo di confidenza a livello 0.90 per la differenza tra le medie.

b. Si decida sull'eguaglianza tra le medie mediante un test al livello del 5%.

c. Si determini il p-value per l'ipotesi di eguaglianza tra le medie

7. Si voglia sottoporre a verifica il seguente sistema di ipotesi relativo alla media di una gaussiana $H_0 : \mu = 10$ contro $H_1 : \mu \neq 10$. Avendo a disposizione un campione casuale di 10 unità che ha fornito i seguenti risultati:

$$\sum_{i=1}^{10} y_i = 97.6 \quad \sum_{i=1}^{10} y_i^2 = 962.2$$

a. Con $\alpha = 0.02$, cosa si decide sull'ipotesi nulla?

b. Il livello di significatività osservato è maggiore di 0.4?

8. Si osserva un campione di 8 valori, $x_1, x_2, \ldots, x_8$, da una esponenziale $X \sim \text{Esp}(\lambda)$. Si vuole verificare l'ipotesi H_0 che $\lambda = 0.5$ contro l'alternativa che sia maggiore. Si rifiuta l'ipotesi nulla se la media è maggiore di 4.

a. Si calcoli la probabilità dell'errore di I tipo.

b. Se la media campionaria risulta essere 1.9175, quanto vale il valore-p per la suddetta verifica d'ipotesi?

9. Si dispone dei dati su un indice del livello di produttività di un gruppo di 9 lavoratori, misurato avendo e non avendo effettuato una pausa caffè.

Senza	28.7	30.4	28.3	33.2	30.7	28.4	31.0	31.5	31.2
Con	30.2	34.8	32.0	29.4	25.5	33.8	30.9	31.0	33.4

a. Determinare un intervallo di confidenza per la differenza tra gli indici di produttività per il gruppo di lavoratori avendo e non avendo effettuato la pausa caffè (al livello 0.95).

b. Verificare l'ipotesi che la pausa caffè sia in grado di aumentare il livello di produttività utilizzando un livello di significatività pari a 0.05.

10. Un'indagine campionaria svolta su $n = 50$ telefonate effettuate a un centro di assistenza clienti ha dato luogo alla seguente distribuzione di frequenza delle durate in minuti.

Durate	≤ 3	[3,5)	[5,8)	[8,10)	≥ 10
Frequenze	5	12	27	5	1

Verificare al livello $\alpha = 0.01$ l'ipotesi che la durata delle telefonate al centro di assistenza clienti abbia distribuzione $\mathcal{N}(\mu = 6, \sigma^2 = 4)$.

11. Un sondaggio svolto tra gli studenti dell'Università di Trieste vuole esaminare la propensione a diminuire il numero di prove di esame in un anno accademico (< 6) contro la propoensione a lasciarle invariate o aumentarle (≥ 6). La seguente tabella mostra le frequenze assolute sulla base della distinzione degli studenti iscritti a corsi *umanistici* (U) e *scientifici* (S).

	U	S
< 6	60	120
≥ 6	100	90

   a. Si può ritenere, al livello del 1%, che vi sia indipendenza tra la propensione a diminuire o aumentare/lasciare invariate e il macro-settore di appartenenza?

   b. Si utilizzino i dati forniti dalla tabella sopra riportata per verificare l'ipotesi $H_0 : p_1 - p_2 \leq 0$ contro l'alternativa $H_1 : p_1 - p_2 > 0$, dove p_1 e p_2 sono le proporzione di studenti dei corsi *umanistici* e *scientifici*, rispettivamente, favorevoli a ridurre le prove di esame. Cosa si decide utilizzando $\alpha = 0.05$?

   c. Qual è il livello di significatività osservato per la verifica dell'ipotesi $H_0 : p_1 - p_2 = 0$ contro l'alternativa $H_1 : p_1 - p_2 \neq 0$? (p_1 e p_2 definiti come al punto b.)

12. L'altezza di 10 atleti (in m) è riportata di seguito

$$1.74 \;\; 1.79 \;\; 1.65 \;\; 1.76 \;\; 1.73 \;\; 1.72 \;\; 1.82 \;\; 1.78 \;\; 1.75 \;\; 1.85$$

Verificare al livello $\alpha = 0.1$ se l'altezza si possa ritenere distribuita con legge normale di media 1.75 m e deviazione standard 0.06 m utilizzando il test di Kolmogorov-Smirnov.