

PROVA SCRITTA di STATISTICA\_a – 04/06/2019

Canale A–D,  
prof.ssa Guagnano

COGNOME

NOME

MATR.

**Prospetto riepilogativo dei risultati:**

1. a) popolazione di riferimento per: l'indagine A: **tutti gli studenti universitari europei;**  
l'indagine B: **tutti gli studenti universitari europei laureati da 3 anni;**  
b) numerosità: **9604**; c) livello di significatività: **0,0082**; tipo di modifica: **il valore critico va ridotto e avvicinato a 0,7**;
3. retta stimata:  $y^*=14,58 + 8,27 x$ ; consumo richiesto: **26,98**; grado di adattamento: **0,7674**;
4. a) risposta selezionata: **iv**; b) intervallo richiesto: **[0,3; 0,6]** (valutazione puntuale: **0,55**);  
c) moda: **2**; media aritmetica: **9,81**; mediana: **10** (oppure, valore centrale della terza classe, pari a **9**);  
varianza: **30,379**; d) IC richiesto: **[8,101; 11,519]**;
- e) p-value: **0,1515**; potenza: **0,9678**;
5. probabilità richiesta in: a) **0,0685**; b) **0,7226**.

**PARTE A (12 punti)**

1. In un dato anno vengono condotte due indagini campionarie, A e B, a livello europeo. La prima è volta a individuare gli ambiti curriculari maggiormente scelti dagli studenti universitari, mentre la seconda è volta ad analizzare la percentuale di occupati a tre anni dalla laurea.
- a) Indicare quali sono le popolazioni di riferimento in ciascuna delle due indagini.
- b) Determinare la numerosità campionaria necessaria nell'indagine B, affinché l'errore campionario sia inferiore a 0,01 con probabilità del 95%.
- c) Supponendo che, sempre nell'indagine B, sulla base dei dati raccolti in un campione di 8400 unità, si voglia verificare se la proporzione di occupati nella popolazione di riferimento sia pari al 70% oppure più elevata e che, con riferimento alla proporzione campionaria non standardizzata, sia pari a 0,712 il valore critico che separa la regione di accettazione da quella di rifiuto, determinare il livello di significatività del test. Indicare, inoltre, in che direzione dovrebbe essere modificato il suddetto valore critico, per ridurre la probabilità di commettere un errore di seconda specie.
2. Definire il concetto di concentrazione e una misura della sua intensità. Definire, inoltre, una misura di variabilità per caratteri qualitativi.

**PARTE B (23 punti)**

3. In un campione di 15 utilitarie vengono rilevati il consumo (variabile Y, espressa in chilometri percorsi per litro di carburante) e il peso (variabile X, espressa in tonnellate), pervenendo alle seguenti sintesi:

$M(X)=1,2$ ;  $Var(X)= 2,25$ ;  $M(Y)=24,5$ ;  $Var(Y)=200,36$ ;  $M(X*Y)=48$ .

Stimare la retta che esprime Y in funzione di X, indicare il consumo che dovremmo attenderci per un'autovettura che pesi 1,5 tonnellate e valutare il grado di adattamento ai dati della retta stimata.

4. A 50 neolaureati è stato proposto un contratto di formazione e lavoro e la distribuzione di frequenza dei giovani secondo la durata in mesi (e frazioni di mesi) è la seguente:

$X_{i, \text{inf}}   - X_{i, \text{sup}}$	$n_i$
1   - 3	10
3   - 6	5
6   - 12	15
12   - 18	18
18   - 25	2
Totale	50

- a) Indicare la tipologia del carattere considerato: i) qualitativo non ordinabile; ii) qualitativo ordinabile; iii) quantitativo discreto; iv) quantitativo continuo.
- b) Determinare l'intervallo di valori ammissibili (o anche una valutazione puntuale) per la funzione di ripartizione empirica in corrispondenza di una durata di 11 mesi.
- c) Fornire le valutazioni puntuali per la moda, la media aritmetica, la mediana e la varianza della durata dei contratti.
- d) Considerando i 50 neolaureati come un campione casuale estratto da una popolazione distribuita normalmente, costruire l'intervallo di confidenza al livello  $1 - \alpha = 0,97$  per la durata media dei contratti nella popolazione.
- e) Con riferimento al test unilaterale destro volto a verificare, al livello di significatività del 2,5%, l'ipotesi che la media di popolazione sia pari a 9 oppure più elevata, valutare il p-value del test e la potenza in corrispondenza dell'alternativa puntuale  $H_1: \mu=12$ .
5. Un test clinico consente di diagnosticare precocemente l'insorgere di una certa malattia, che colpisce il 5% della popolazione adulta. Si sa che l'esame può dar luogo a 'falsi positivi' (cioè viene diagnosticata la malattia, nonostante il soggetto sia sano), oppure a 'falsi negativi' (cioè non viene diagnosticata la malattia, nonostante il soggetto ne sia affetto); ciò avviene, rispettivamente, nel 2% e nell'1% dei casi.
- a) Valutare la probabilità che ad una persona qualsiasi venga diagnosticata la malattia.
- b) Valutare la probabilità che una persona, alla quale il test abbia diagnosticato la malattia, sia realmente affetta dalla malattia.