

**SIMULAZIONE D'ESAME DI METODI STATISTICI PER LA BIOINGEGNERIA 25/26**

## PARTE 2: DOMANDE APERTE

**Durata: 30 minuti**

**Domanda 1 (6 punti):**

- 1) Data una matrice  $X \in \mathbb{R}^{N \times M}$ , dove N è il numero di osservazioni ed M il numero di variabili, si descriva il metodo di clustering K-means riportando in particolare:
  - a. la funzione obiettivo che viene utilizzata nel caso di utilizzo della distanza euclidea, specificando il significato di tutti i termini utilizzati
  - b. gli step principali dell'algoritmo iterativo che consente di minimizzare tale funzione obiettivo.
- 2) Supponendo che i dati vengano suddivisi in K cluster, quanti saranno i centroidi e che dimensione avranno?

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

**Domanda 2 (5 punti):**

- 1) Si scriva l'equazione del modello di regressione logistica univariata, ovvero la sua versione più semplice in cui compare una sola variabile indipendente  $X_1$ , specificando il significato dei termini coinvolti.
- 2) Si scriva la formula dell'odds ratio associato alla variabile  $X_1$ .
- 3) Come si interpretano i valori dell'odds ratio associato alla variabile  $X_1$ ?

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.