

Università degli Studi di Milano-Bicocca – Facoltà di Economia
Esame di Analisi dei Dati (modulo B)

14 luglio 2004

N.B.: Commentare sempre i risultati ottenuti

1) Si illustrino gli obiettivi generali dell'analisi delle corrispondenze e si presentino le analogie e le diversità fra analisi delle corrispondenze e analisi delle componenti principali.

2) Gli aerei della compagnia Air France possono essere classificati in tre categorie, qui indicate con A, B e C. Con l'obiettivo di spiegare l'appartenenza di aereo della compagnia ad una delle 3 categorie menzionate, sono state rilevate su un campione di 10 aerei Air France (di cui 4 appartengono alla categoria A, 3 alla categoria B e 3 alla categoria C) le seguenti caratteristiche: $X_1 = \text{Capacità (numero di passeggeri a pieno carico, in centinaia)}$; $X_2 = \text{Autonomia (chilometri percorribili fra due rifornimenti di carburante, in migliaia)}$.

Nella seguente matrice sono riportate le osservazioni campionarie:

$$X^T = \begin{bmatrix} 1.07 & 1.54 & 1.7 & 1.85 & 2.1 & 2.5 & 1.8 & 2.7 & 2.4 & 4 \\ 6 & 4 & 10 & 2.9 & 7 & 6.9 & 5.5 & 12 & 14 & 13 \end{bmatrix}$$

da tale matrice risulta che media aritmetica e varianza per la variabile X_1 sono rispettivamente 2,166 e 0,5863, mentre per la variabile X_2 sono rispettivamente 8,13 e 13,5301.

Considerato che la matrice A degli autovettori normalizzati della matrice $W^{-1}B$ è pari a:

$$A = \begin{bmatrix} 0.7027 & -0.8582 \\ 0.7115 & 0.5134 \end{bmatrix}$$

- Si calcolino i punteggi discriminanti rispetto alle funzioni discriminanti estraibili.
- Si calcolino le correlazioni tra variabili originarie e funzioni discriminanti, commentando i risultati ottenuti.
- Si classifichi opportunamente in uno dei tre gruppi un nuovo aereo, caratterizzato da una capacità di 195 passeggeri e da un'autonomia di 14500 chilometri con un pieno di carburante.

3) Tipi di distribuzioni e forme di indipendenza in tabelle di contingenza riferite a tre caratteri.

4) Nella seguente tabella, 100 impiegati di una grande società sono stati classificati in base alle loro abitudini legate al fumo ed alla posizione occupata nella società:

Fumo	Posizione	Amministrativo	Venditore	Esecutivo	TOTALE
Non fumatore		7	15	8	30
Fumatore Leggero		5	18	14	37
Fumatore Accanito		6	17	10	33
TOTALE		18	50	32	100

Fissando in $\alpha = 0.05$ il livello di significatività, stabilire quale modello log-lineare gerarchico è idoneo a descrivere la relazione tra le variabili **Fumo** e **Posizione** nella popolazione corrispondente.

5) Criteri, descrittivi e inferenziali, per stabilire il numero di funzioni discriminanti da mantenere nell'analisi.