

Università degli Studi di Milano-Bicocca - Facoltà di Economia
Esame di Analisi dei Dati (modulo B)
26 gennaio 2005

NB : Commentare sempre i risultati ottenuti.

1. Lo schema di campionamento multipoissoniano nel caso di una tabella di contingenza riferita a due caratteri.

2. Dopo aver descritto la natura dei dati di cui bisogna disporre per applicare l'analisi delle corrispondenze ed aver spiegato gli obiettivi dell'analisi stessa ricorrendo eventualmente a qualche esempio, si definisca il primo asse principale nello spazio dei profili riga (o colonna), illustrandone il significato geometrico e la soluzione analitica.

3. Dopo aver illustrato il contesto applicativo e gli obiettivi dell'analisi discriminante si definisca e si ricavi la prima funzione discriminante.

4. Con l'obiettivo di spiegare le preferenze di un gruppo di giovani rispetto a tre tipi di bevande, si rilevano presso un campione di 10 teenager, dei quali 3 consumatori della bevanda *A*, 4 della bevanda *B* e 3 della bevanda *C*, le seguenti informazioni: *consumo settimanale* (in litri), *tipo di confezione consumata* (piccola, media, grande), *versione di bibita* (gassata, non gassata).
I centri dei punteggi discriminanti dei tre gruppi rispetto alle funzioni discriminanti estraibili sono contenuti nei seguenti vettori:

$$\mathbf{c}_1 = \begin{bmatrix} +1,783 \\ -0,572 \\ -1,021 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} +0,161 \\ -0,755 \\ +0,845 \end{bmatrix}.$$

Considerato che gli autovalori della matrice $\mathbf{W}^{-1}\mathbf{B}$ sono $\lambda_1 = 1,9959$ e $\lambda_2 = 0,6434$ e la matrice \mathbf{R}^* delle correlazioni tra le variabili originarie (secondo l'ordine di presentazione precedentemente descritto) e le funzioni discriminanti è:

$$\mathbf{R}^*_{(3 \times 2)} = \begin{bmatrix} 0,630 & 0,564 \\ 0,168 & -0,189 \\ -0,872 & 0,489 \end{bmatrix};$$

- a) si stabilisca il numero di funzioni discriminanti da mantenere nell'analisi secondo un criterio puramente descrittivo;
 - b) si commentino i risultati contenuti nella matrice \mathbf{R}^* e si effettui una rappresentazione grafica dell'analisi interpretando opportunamente.
-
5. Nel condurre uno studio per il rinnovo delle confezioni delle scatole di cioccolatini, una ditta dolciaria ha analizzato un campione di 400 consumatori rilevando le loro preferenze secondo la forma dei cioccolatini e il colore della confezione di ognuno. I risultati sono riportati nella seguente tabella:

<i>Forma</i>	<i>Colore</i>			
	<i>rosso</i>	<i>verde</i>	<i>blu</i>	Totale
<i>quadrata</i>	44	46	50	140
<i>rettangolare</i>	36	46	48	130
<i>esagonale</i>	45	40	45	130
Totale	125	132	143	400

Fissando $\alpha = 0,05$, stabilire quale modello log-lineare gerarchico sia idoneo a descrivere la relazione fra la forma del cioccolatino e il colore della relativa confezione nella popolazione dei consumatori corrispondente.