

Università degli Studi di Milano-Bicocca - Facoltà di Economia
Esame di Analisi dei Dati (modulo B)
 17 febbraio 2006

NB : Commentare sempre i risultati ottenuti.

1. Una azienda produttrice di macchine utensili è interessata a valutare la soddisfazione dei propri clienti. A tal proposito viene selezionato un campione di 377 clienti utilizzatori delle macchine utensili e viene richiesto a tali clienti di esprimere un giudizio sul livello di soddisfazione (*I* insoddisfatto, *S* soddisfatto, *PS* più che soddisfatto, *MS* molto soddisfatto) per quel che riguarda l'*assistenza clienti* e la *formazione* del personale fornita dall'azienda. La tabella di contingenza riferita ai dati rilevati è la seguente:

<i>Formazione</i> <i>Assistenza clienti</i>	<i>I</i>	<i>S</i>	<i>PS</i>	<i>MS</i>	<i>Totale</i>
<i>I</i>	9	3	12	11	35
<i>S</i>	10	4	15	6	35
<i>PS</i>	22	10	81	31	144
<i>MS</i>	25	4	37	97	163
<i>Totale</i>	66	21	145	145	377

Attraverso l'analisi delle corrispondenze si sono ottenute le due seguenti matrici le cui colonne riportano rispettivamente le proiezioni dei profili *Assistenza Clienti* e le proiezioni dei profili *Formazione* sugli assi principali estraibili:

$$\left[\begin{array}{ccc} 0,0981 & 0,2497 & -0,0070 \\ 0,3927 & 0,3170 & 0,0054 \\ 0,3864 & -0,1173 & -0,0002 \\ -0,4467 & -0,0180 & 0,0005 \end{array} \right] \quad \left[\begin{array}{ccc} 0,0815 & 0,2523 & -0,0034 \\ 0,4715 & 0,2566 & 0,0096 \\ 0,3782 & -0,1165 & -0,0005 \\ -0,4836 & -0,0355 & 0,0007 \end{array} \right]$$

- a) Quanti assi principali si ritiene che sia utile mantenere nell'analisi?
- b) In base alla risposta al punto a) si effettui una opportuna rappresentazione grafica dei risultati e si commenti adeguatamente.
- c) Valutare la qualità della rappresentazione dei profili riga e colonna nello spazio degli assi principali mantenuti.
2. Criteri descrittivi e inferenziali, per stabilire il numero di funzioni discriminanti da mantenere nell'analisi.
3. Si presenti nel dettaglio il modello log-lineare saturo per una tabella di contingenza.
4. Si consideri la seguente distribuzione congiunta di frequenze assolute riferita ai caratteri X_1 classe di viaggio (*prima*, *seconda*, *terza*), X_2 esito del naufragio (*morto*, *sopravvissuto*), X_3 sesso, rilevati su un campione di 645 passeggeri del transatlantico Titanic affondato nel 1912:

	X_3			
	<i>uomini</i>		<i>donne</i>	
X_2	<i>morti</i>	<i>soprav.</i>	<i>morti</i>	<i>soprav.</i>
X_1				
<i>prima</i>	55	30	3	63
<i>seconda</i>	75	11	7	45
<i>terza</i>	210	43	53	50

- a) Si verifichi che non esiste indipendenza completa.
- b) Si verifichi che X_1 non è indipendente dalla distribuzione doppia (X_2, X_3) .
- c) Si fornisca la distribuzione congiunta di X_1, X_2 e X_3 di frequenze assolute in cui X_1 e X_2 sono indipendenti marginalmente condizionatamente a X_3 .