

Università degli Studi di Milano-Bicocca – Facoltà di Economia
Esame di Analisi dei Dati (modulo B)
 10 giugno 2004
N.B.: Commentare sempre i risultati ottenuti

1) Per scopi pubblicitari, una società di marketing è interessata a valutare se esiste una relazione fra condizione professionale delle donne e la tendenza delle stesse ad effettuare acquisti al mercato, nei negozi oppure al supermercato. Ad un campione di 100 donne viene chiesto di indicare la propria condizione professionale (classificata per semplicità in *casalinga*, *impiegata* e *dirigente*) e di rispondere alla domanda: “in quale punto vendita, tra *mercato*, *negozio* e *supermercato*, effettua più frequentemente i suoi acquisti?”. Dalla tabella di contingenza riferita alle associazioni ottenute si hanno i seguenti profili medi, rispettivamente per i caratteri **Condizione professionale** e **Punto vendita**:

$$\bar{c}' = (0.3 \quad 0.45 \quad 0.25) \qquad \bar{r}' = (0.35 \quad 0.15 \quad 0.5)$$

Attraverso l'analisi delle corrispondenze si sono ottenute le seguenti matrici, le cui colonne riportano rispettivamente le proiezioni dei profili **Condizione professionale** e le proiezioni dei profili **Punto vendita** sugli assi principali estraibili:

$$\begin{bmatrix} 0.0933 & 0.1816 \\ -0.3189 & -0.0572 \\ 0.4620 & -0.1150 \end{bmatrix} \qquad \begin{bmatrix} -0.0301 & 0.1647 \\ 0.7367 & -0.0697 \\ -0.2000 & -0.0944 \end{bmatrix}$$

- a) Quanti assi principali si ritiene sia utile mantenere nell'analisi?
- b) In base alla risposta fornita al punto (a) si effettui un'opportuna rappresentazione grafica dei risultati e si commenti adeguatamente.
- c) Si calcolino i contributi assoluti e relativi ritenuti utili per il tipo di analisi, e si commentino i risultati ottenuti.

2) La distanza *chi-quadrato*: definizione, proprietà ed impieghi.

3) Si illustrino gli obiettivi dell'analisi discriminante, si definisca e si determini la prima funzione lineare discriminante, si estenda quindi il criterio alle altre funzioni discriminanti. Quante sono le funzioni discriminanti estraibili?

4) All'uscita da una libreria, un campione di 100 soggetti che hanno acquistato un solo libro è stato intervistato sulla propria professione e sul genere di libro acquistato. I dati sono riportati nella seguente tabella:

Genere	<i>giallo</i>	<i>rosa</i>	<i>avventura</i>	<i>saggio</i>	<i>comico</i>
<i>studente</i>	10	2	1	2	7
<i>impiegato</i>	11	2	3	1	9
<i>pensionato</i>	13	0	2	1	8
<i>autonomo</i>	11	1	4	2	10

Si descriva il modello log-lineare di distribuzioni di colonna uniformi per la tabella sopra riportata e si forniscano le stime di massima verosimiglianza delle frequenze attese e dei parametri di tale modello.

5) Si illustri l'origine della statistica test impiegata per effettuare verifiche d'ipotesi sull'adattamento di un modello log-lineare per tabelle di contingenza a duplice entrata, specificando l'ipotesi nulla e l'ipotesi alternativa.