

**Università degli Studi di Milano-Bicocca - Facoltà di Economia**  
**Esame di Analisi dei Dati (modulo II)**  
 29 giugno 2006

*NB : Commentare sempre i risultati ottenuti.*

1. Si consideri la seguente distribuzione congiunta di frequenze assolute riferita ai caratteri  $X_1$  *tenore di vita* (*basso, medio, alto*),  $X_2$  *sexso* ( $M, F$ ),  $X_3$  *iscrizione ad una associazione culturale* (*si, no*), rilevati su un campione di 200 utenti estratti tra i visitatori del Cenacolo Vinciano dell'ultimo anno:

	$X_3$			
	<i>si</i>		<i>no</i>	
$X_2$	$M$	$F$	$M$	$F$
$X_1$				
<i>basso</i>	10	15	2	5
<i>medio</i>	18	35	14	6
<i>alto</i>	30	40	10	15

- a) Si verifichi che  $X_1$  non è indipendente dalla distribuzione doppia ( $X_2, X_3$ ).
- b) Si fornisca la distribuzione congiunta di  $X_1, X_2$  e  $X_3$  di frequenze assolute di indipendenza completa.
- 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 condizionatamente a  $X_3$ .
2. Nella seguente tabella, relativa all'anno 1995, è riportata, per un campione di 100 fondi comuni di investimento, la classificazione secondo il tasso percentuale di *rendimento annuale* e il *tipo di fondo*:

<i>Rendimento annuale</i>	<i>3</i>	<i>5</i>	<i>8</i>	<i>10</i>
<i>Tipo di fondo</i>				
<i>azionario</i>	11	6	4	7
<i>bilanciato</i>	14	10	4	6
<i>obbligazionario</i>	13	12	6	7

Si descriva il modello log-lineare di distribuzioni di colonna uniformi per la tabella sopra riportata, commentando adeguatamente rispetto al contesto di riferimento. Si forniscano le stime di massima verosimiglianza delle frequenze attese e dei parametri di tale modello.

3. I punteggi discriminanti: definizione, caratteristiche e impieghi.
4. Analisi delle corrispondenze: si dimostri la relazione tra la misura  $X^2$  di Pizzetti-Pearson e la media aritmetica ponderata di opportune distanze chi-quadrato.