

Corso di Valutazione di Campagne di Comunicazione – 2007/08

Esercitazione n. 5

Es. A) Si consideri la seguente tabella, estratta dal *Compendio Statistico Italiano* (1998), che riporta il numero di occupati (x1.000) secondo il tipo di occupazione (Agricoltura, Industria e Altre attività) per alcune regioni italiane.

Regione	Tipo di occupazione			Totale
	Agricoltura	Industria	Altre attività	
Piemonte	87	676	930	1693
Emilia Romagna	121	592	981	1694
Lazio	80	362	1375	1817

- Si ricavi la distribuzione marginale del carattere *tipo di occupazione*, riportando anche le frequenze assolute cumulate;
- si derivino le distribuzioni percentuali condizionate del carattere *tipo di occupazione* per ogni modalità del carattere *regione* e si confrontino tali distribuzioni commentando opportunamente le eventuali differenze;
- si ricavi la tabella di indipendenza e quella delle contingenze (c_{ij});
- si calcolino gli indici di connessione χ^2 e C di Cramér.

Es. B) Si consideri la tabella seguente che riporta il numero cittadini italiani condannati in procedimenti penali celebrati nel 1999 a seconda del tipo di delitto commesso e della durata della reclusione (*Compendio Statistico Italiano*, 2000).

Tipo di reato	Reclusione		
	0-5 anni	5-10 anni	10-20 anni
Omicidio volontario	137	134	460
Omicidio colposo	3.601	2	0

- Si calcolino e si confrontino tra loro le distribuzioni condizionate percentuali del carattere *reclusione* per ogni modalità del carattere *tipo di reato*;
- si costruisca la corrispondente tabella con le frequenze teoriche sotto indipendenza e, per ogni cella della tabella, si calcoli la contingenza;
- si calcoli l'indice χ^2 di connessione e l'indice C di Cramer;
- si rappresenti la spezzata di regressione e si calcoli il rapporto di correlazione η^2 commentando opportunamente il risultato ottenuto.

Es. C) Nell'ambito di un'indagine sulla distribuzione del reddito nelle Marche, sono state classificate 240 famiglie a seconda del reddito percepito annualmente dai due coniugi:

Reddito dello sposo (in migliaia di €)	Reddito della sposa (in migliaia di €)			
	[0; 15)	[15; 30)	[30; 60)	[60; 120)
[0; 15)	25	5	3	0
[15; 30)	15	17	8	2
[30; 60)	19	31	29	12
[60; 200)	7	33	27	7

- Si rappresenti la spezzata di regressione e si commenti opportunamente il grafico ottenuto;
- si calcoli il rapporto di correlazione η^2 e si commenti il risultato ottenuto.

Es. D) Si indichi se ognuna delle seguenti affermazioni è vera o falsa

N.	Domanda	V	F
1	Una distribuzione doppia di frequenza associa a ogni coppia di modalità dei due caratteri (x_i, y_j) la corrispondente frequenza congiunta n_{ij}		
2	Sommando le frequenze congiunte in una colonna di una tabella a doppia entrata si ottiene una frequenza marginale		
3	Da una distribuzione doppia di frequenza di due caratteri X e Y è possibile ottenere sia la distribuzione marginale di X sia quella di Y		
4	La somma delle frequenze n_{ij} della distribuzione condizionata di Y dato $X = x_i$ è sempre inferiore alla frequenza marginale $n_{i\cdot}$ della modalità x_i		
5	Dividendo una frequenza congiunta n_{ij} per la corrispondente frequenza marginale $n_{i\cdot}$ si ottiene una frequenza condizionata relativa		
6	Due caratteri qualitativi X e Y sono statisticamente indipendenti se e solo se ogni frequenza congiunta n_{ij} è pari a $N/(n_{i\cdot}n_{\cdot j})$ dove $n_{i\cdot}$ e $n_{\cdot j}$ indicano le frequenze marginali dei due caratteri e N indica la frequenza totale		
7	Non può mai accadere che un carattere dipenda perfettamente da un altro carattere		
8	Può accadere che tutte le contingenze $c_{ij} = n_{ij} - \hat{n}_{ij}$ di una distribuzione doppia di frequenza siano pari a 0		
9	La somma di tutte le contingenze $c_{ij} = n_{ij} - \hat{n}_{ij}$ di una distribuzione doppia di frequenza è sempre pari a 0		
10	Tutte le contingenze relative di una distribuzione doppia di frequenza, c_{ij} / \hat{n}_{ij} , sono sempre comprese tra 0 e 1		
11	L'indice di connessione χ^2 non può mai essere negativo		
12	Il massimo dell'indice di connessione χ^2 è pari a $\min[(r-1), (c-1)]$, dove r è il numero di modalità del carattere X e c è il numero di modalità del carattere Y		
13	L'indice di connessione di Cramér, C , è sempre compreso tra 0 e 1		
14	L'indice di dipendenza in media η^2 è calcolabile per qualsiasi distribuzione doppia		
15	La media condizionata di Y dato $X = x_i$ è sempre calcolabile se X è un carattere quantitativo		
16	Il carattere Y è indipendente in media da X se e solo se $\bar{y}_1 = \bar{y}_2 = \dots = \bar{y}_r$ dove \bar{y}_i è la media condizionata di Y dato $X = x_i$		
17	La devianza spiegata D_S , utilizzata nell'ambito della misurazione della dipendenza in media, è sempre minore o uguale della devianza totale D_Y		
18	Quando si ha massima dipendenza in media, la devianza spiegata D_S è uguale a quella totale D_Y		
19	La dipendenza in media implica la dipendenza statistica		
20	L'indipendenza in media implica l'indipendenza statistica		