

**ESERCIZIO**

La tabella indagine *StudentiFlorida*, disponibile sul sito (in formato pdf e xls), riporta i dati di un'indagine condotta su 60 studenti laureati presso la University of Florida, relativamente alle seguenti variabili:

COLONNA	NOME VARIABILE	DESCRIZIONE
1	id	Codice identificativo dello studente
2	genere	Genere dello studente
3	età	Età (in anni compiuti)
4	punteggioHighSchool	Punteggio ottenuto alla scuola superiore (scala da 1 a 4)
5	punteggioCollege	Punteggio ottenuto al College (scala da 1 a 4)
6	distanzaCampus	Distanza (in miglia) del campus dalla città di provenienza dello studente
7	distanzaAula	Distanza (in miglia) della classe dall'attuale luogo di residenza
8	oreTVsettimana	Numero medio di ore settimanali passate a guardare la TV
9	oreSportSettimana	Numero medio di ore settimanali dedicate allo sport
10	quotidianiPerSettimana	Numero di volte per settimana che si legge un quotidiano
11	AH	Numero di persone conosciute morte per AIDS o sieropositive
12	vegetariano	Regime di alimentazione vegetariano (yes, no)
13	affiliazionePolitica	Affiliazione politica (D = Democratico, R = Repubblicano, I = Indipendente)
14	ideologiaPolitica	Ideologia politica (1 = molto liberale, 2 = liberale, 3 = leggermente liberale, 4 = moderato, 5 = abbastanza conservatore, 6 = conservatore, 7 = molto conservatore)
15	religiosità	Partecipazione ad eventi religiosi (0 = mai, 1 = occasionalmente, 2 = la maggior parte delle settimane, 3 = ogni settimana)
16	opinioneAborto	Opinione circa la legalità dell'aborto nei primi tre mesi di gravidanza (yes, no)
17	azioniDiscriminazione	Impegno in azioni contro le discriminazioni (si, no)
18	vitaMorte	Crede nella vita dopo la morte (si -y, incerto - u, no - n)

Fonte: A. Agresti and B. Finlay, *Statistical Methods for the Social Science*, Prentice Hall, 2008

1) Di seguito sono riportati, per i soli 16 studenti che hanno risposto incerto alla domanda inerente la variabile **vitaMorte**, i valori della variabile **distanzaCampus**:

1200	1300	350	5000	190	1000	420	1200	316	900	180	1100	360	6	80	2000
------	------	-----	------	-----	------	-----	------	-----	-----	-----	------	-----	---	----	------

Su tali valori, calcolare i seguenti indici di forma/rappresentazioni grafiche:

Confronto tra media e mediana	
Indice di Yule-Bowley	

Indice di Fisher	
Boxplot (senza considerare i valori anomali)	
Boxplot (evidenziando eventuali valori anomali)	

2) La seguente tabella riporta la distribuzione di frequenza per la variabile **quotidiani per settimana**:

<b>quotidiani per settimana</b>	<b>ni</b>
0	4
1	6
2	9
3	13
4	4
5	9
6	3
7	9
12	1
14	2
<b>TOT.</b>	<b>60</b>

Su tali valori, calcolare i seguenti indici/rappresentazioni grafiche:

<p>Confronto tra media e mediana</p>	
<p>Indice di Yule-Bowley</p>	
<p>Indice di Fisher</p>	
<p>Boxplot (senza considerare i valori anomali)</p>	

Boxplot (evidenziando eventuali valori anomali)	
---	--

3) La seguente tabella riporta una serie di indici di posizione, variabilità e forma per le due variabili **età** e **quotidiani per settimana**. A partire da tali valori, confrontare le due distribuzioni in termini di variabilità e forma, motivando brevemente le risposte:

Indicatore	Età	Quotidiani per settimana	Quale delle due distribuzioni è più variabile/asimmetrica?
$\mu$	29.17	4.08	
Me	26.5	3	
q <sub>1</sub>	24	2	
q <sub>2</sub>	26.5	3	
q <sub>3</sub>	31	5.25	
$\sigma$	8.40	2.98	
S <sub>Me</sub>	5.1	2.18	
MAD	3.5	2	
$\mu - Me$	+2.67	1.08	
$(q_3 - q_2) - (q_2 - q_1)$	2	1.25	
$\sum_{i=1}^N \frac{(x_i - \mu)^3}{N}$	1578.84	35.86	

4) La seguente tabella riporta la distribuzione congiunta delle variabili **affiliazione politica** e **genere**:

		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico	12	9	21
	indipendente	13	11	24
	repubblicano	6	9	15
TOT.		31	29	60

Calcola la tabella dei profili colonna (distribuzioni in frequenza relativa della variabile affiliazione politica condizionate sulle modalità della variabile colonna e distribuzione marginale):

		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			
TOT.		1	1	1

Dalla sola analisi della tabella dei profili colonna, pensi ci sia un'associazione tra le due variabili in tabella? (motiva brevemente la risposta)

Calcola la tabella delle frequenze teoriche sotto l'ipotesi di indipendenza:

		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			
TOT.				

Calcola la tabella delle differenze tra frequenze teoriche e frequenze osservate:

		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			0
	indipendente			0
	repubblicano			0
TOT.		0	0	0

Calcola gli indici chi-quadro, fi-quadro e v di Cramer:

		GENERE		
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			

$\chi^2$	
$\phi^2$	
$v$	

A partire dagli stessi totali di colonna, prova a ipotizzare una configurazione delle frequenze che si avvicini all'ipotesi di indipendenza assoluta:

		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			
TOT.		31	29	60

Calcola i tre indici di associazione sulla tabella ipotizzata al punto precedente per verificare che essi siano vicini al valore minimo:

FREQUENZE TEORICHE		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			
TOT.		31	29	60

TABELLA DI CALCOLO PER IL $\chi^2$		GENERE		60
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			

$\chi^2$	
$\phi^2$	
$v$	

A partire dagli stessi totali di colonna, prova a ipotizzare una configurazione delle frequenze che si avvicini all'ipotesi di massima connessione:

		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			
TOT.		31	29	60

Calcola i tre indici di associazione sulla tabella ipotizzata al punto precedente per verificare che essi siano vicini al valore massimo:

FREQUENZE TEORICHE		GENERE		TOT.
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			
TOT.		31	29	60

TABELLA DI CALCOLO PER IL $\chi^2$		GENERE		
		donna	uomo	
AFFILIAZIONE POLITICA	democratico			
	indipendente			
	repubblicano			
				60

$\chi^2$	
$\phi^2$	
$\mathbf{v}$	