

Statistica Sociale

Determina se le seguenti osservazioni sono vere o false

1. Nel calcolo della varianza, gli scarti più elevati (tra i valori osservati e la media) acquistano molta più importanza rispetto agli scarti più piccoli
2. Un indice di variabilità deve assumere il suo valore minimo se e solo se tutte le modalità osservate sono uguali a zero
3. L'indice χ^2 è negativo se tra i due caratteri prevale la discordanza.
4. L'indice χ^2 assume il valore 0 solo se le variabili sono statisticamente indipendenti.
5. C'è dipendenza perfetta tra due caratteri quando ad ogni modalità del carattere Y corrisponde una sola modalità del carattere X.

- Con riferimento alla seguente distribuzione (incompleta) di un collettivo di individui secondo il sesso (Y) ed il comune di residenza (X)

		Y	
		M	F
X	A		5
	B	4	20
	C	8	

Completare la tabella sotto l'ipotesi di indipendenza

- Nella seguente tabella è data la distribuzione congiunta secondo il sesso, l'abitudine al fumo e l'età di un collettivo di 536 individui.

Sesso	Abitudine al fumo	Età			Totale
		14-30	30-50	50-90	
Maschi	Fumatore	20	35	30	85
	Ex-fumatore	4	8	8	20
	Non fumatore	40	45	75	160
	Totale	64	88	113	265
Femmine	Fumatore	15	30	15	60
	Ex-fumatore	3	7	9	19
	Non fumatore	50	70	72	192
	Totale	68	107	96	271

- a) Calcolare le distribuzioni relative condizionate dell'età rispetto alle modalità "Fumatore", "ex-fumatore" e "fumatore" per gli individui di sesso maschile
- b) Calcolare le medie e varianze condizionate dell'età rispetto al carattere abitudine al fumo per gli individui di sesso maschile
- c) Costruire la tabella doppia di frequenze per abitudine al fumo ed età rispetto all'intero collettivo
- d) Per il carattere età della tabella al punto c) calcolare le distribuzioni relative condizionate e le medie condizionate

- In una città sono stati osservati giornalmente la condizione metereologica e il livello di traffico automobilistico per un periodo di tempo di un anno. Si è potuta costruire così la seguente tabella doppia:

		Livello di traffico			
		Basso	Medio	Alto	Totale
Tempo	Sereno	84	26	11	121
	Variabile	29	98	29	156
	Pioggia	7	26	55	88
	Totale	120	150	95	365

Verificare la presenza di associazione.

- La seguente tabella riporta il numero di ore di studio per la preparazione dell'esame di statistica e la votazione finale per un gruppo di 6 studenti

Ore di studio	Voto statistica
25	28
20	23
29	19
38	30
15	20
30	26

Rappresentare i due caratteri in un grafico a dispersione (scatterplot) e commentare il tipo di relazione osservata tra le due variabili

- Sono stati rilevati i pesi (in kg) di 10 maschi (M) e 10 femmine (F) di una particolare specie di pesce

M	1,2	3,0	5,2	4,0	3,5	4,3	3,3	4,8	3,8	3,2
F	1,3	2,2	1,5	2,3	1,8	1,7	2,1	2,0	1,9	2,1

- Calcolare per i maschi e per le femmine la deviazione standard, la distanza interquartile ed il campo di variazione
 - Costruire due boxplot e commentarli
 - Verificare attraverso i boxplot la presenza di valori anomali
- Per un campione casuale di studenti è stato rilevato lo sport preferito. La tabella seguente riporta queste informazioni con riferimento al sesso dei telespettatori.

Sport preferito	Maschi	Femmine
Calcio	182	64
Basket	60	44
Pallavolo	96	86

- Rappresentare graficamente le distribuzioni dello sport preferito condizionatamente al sesso;
- Quanti sarebbero i maschi che indicano il calcio come sport preferito se vi fosse indipendenza tra sesso e preferenza indicata?

- Le seguenti affermazioni sono corrette? Motivare la risposta.
 - a) la correlazione fra n.ro di anni d'istruzione e reddito calcolato su 100 lavoratori dipendenti corrisponde a $r = 0.18$ euro.
 - b) c'è una forte correlazione fra genere dei lavoratori in Italia e il loro reddito.
 - c) è stata trovata un'alta correlazione ($r = 1.09$) fra i voti all'esame di statistica e i voti all'esame di matematica degli studenti della Facoltà di Economia.