

1. Se il rapporto di concentrazione è pari a 0...

- o ... la distribuzione del carattere è uniforme
- o ... la curva di Lorenz giace quasi interamente sull'asse delle ascisse
- o ... allora anche la mediana = 0
- o ... allora anche lo scarto interquartile = 0

la risposta giusta era la 4: entrambe gli indici misurano la variabilità. Il rapporto di concentrazione pari a 0 indica che il carattere è equidistribuito (non uniformemente distribuito) allora, e solo allora  $R=0$ .

Il termine "uniformemente distribuito" si usa per le variabili casuali, quando la probabilità è costante per tutti i valori della v.c. (discreta) o la densità di probabilità è costante per tutto l'intervallo in cui la v.c. è definita (continua)

2. Se anziché misurare in metri misurassi la lunghezza di 15 pezzi di stoffa in cm, otterrei un valore della varianza...

- o ... 100 volte più grande
- o ... 100 volte più piccolo
- o ... 10000 volte più grande
- o ... 1500 volte più piccolo

per avere una lunghezza espressa in cm, devo moltiplicare per 100 la lunghezza espressa in metri siccome nella varianza si elevano al quadrato gli scarti dalla media, avendo scarti 100 volte maggiori, avremo scarti quadrati 10mila volte maggiori la risposta giusta è pertanto la 3

3. Abbiamo una popolazione di 16 donne e 30 uomini. Delle 16 donne, 8 votano verde e 8 giallo. I 30 uomini votano tutti giallo. Allora l'indice chi-quadro...

- o È pari a 0
  - o È pari a 1
  - o È certamente minore di 1 e certamente maggiore di 1
  - o È certamente maggiore di 0
- |   | G  | V | G  | V        |          |
|---|----|---|----|----------|----------|
| D | 8  | 8 | 16 | 2,059497 | 9,782609 |
| U | 30 | 0 | 30 | 1,098398 | 5,217391 |
|   | 38 | 8 | 46 |          | 18,15789 |

Qui la risposta giusta era la quarta. Anche senza calcolare l'indice (che vale 18) sappiamo che è pari a 0 solo se i due caratteri sono indipendenti.

Ma dato che il 100% degli uomini vota giallo e le donne votano 50% e 50%, sesso e voto hanno un'associazione molto evidente.

Quindi chi-quadro non è 0. Non è limitato superiormente (dipende anche dal numero di osservazioni). L'indice derivato da chi-quadro che è compreso tra 0 e 1 è l'indice V di Cramer

4. Dalle tavole, trovo che per la normale standardizzata, ho dedotto che l'area a dx di  $z=-1,2$  è 0,8849. Quale valore ho letto per fare questa affermazione?

- o 0,1151
- o proprio 0,8849
- o ho sbagliato qualcosa: l'affermazione non può essere vera!
- o 0,2302

sulla tavola della normale standardizzata sono riportate le aree della n.st. a sx di un certo valore z positivo.

Per i valori negativi di z si sfrutta la simmetria della n.c.

Così, se voglio conoscere l'area a dx di -1,2 devo cercare l'area a sx di +1,2 (le due aree sono uguali per la simmetria della normale). Sulla tavola dunque, in corrispondenza di  $z=+1,2$  leggo proprio il numero 0,8849

la risposta giusta dunque era la 2

5. Un carattere X, è distribuito uniformemente, con valori compresi tra 100 e 200 (uniforme continua). Allora...

- o ... la probabilità di estrarre un valore x uguale alla media è  $1/n$
- o ... la probabilità di estrarre un valore x tra 110 e 120 è 0,1
- o ... la probabilità di estrarre un valore x uguale alla media è maggiore della probabilità di estrarre un valore  $x = 107$
- o ... lo scarto interquartile è facilmente calcolabile: è 0,5

Per un carattere distribuito uniformemente, la prob di estrarre X in un certo intervallo è uguale al rapporto della lunghezza dell'intervallo e l'intervallo complessivo

la p di estrarre un valore in un intervallo di larghezza 10, su un intervallo di lunghezza totale 100 è dunque 10%. Quindi la risposta 2 è corretta.

se il carattere è continuo, la p di estrarre un qualsiasi valore (estatto) è 0 (quindi la 1 e la 3 sbagliate)

per la variabile descritta  $Q1=125$  e  $Q2=175$  quindi lo scarto interquartile è 50.

6. Abbiamo adottato come scala di equivalenza "0,7+0,3(n-1)", dove n è il numero di componenti. Per avere lo stesso livello di benessere, quante risorse economiche in più ha bisogno una famiglia di 4 persone rispetto a una di 3?

- o Il 30% in più
- o Circa il 23% in più
- o Circa il 20% in più
- o Circa il 19% in più

i coefficienti di equivalenza per 3 e 4 sono, rispettivamente, 1,3 e 1,6. Quindi, dato che  $1,6/1,3=1,2308$ , la risposta giusta è la 2

7. Calcolare 75° percentile del carattere "età degli ospiti di un albergo" (in anni compiuti) sapendo che: 85 turisti hanno da 0 a 18 anni, 65 turisti tra 18 e 45 e 50 turisti tra 45 e 80 (ipotizzare che all'interno di ogni classe la densità di frequenza sia costante).

- o 27
- o 150
- o 45
- o 48,35

su un totale di  $85+65+50=200$  turisti, il 75%, ovvero 150, hanno un'età fino a 45 anni, Quindi la risposta giusta è la 3

8. Per i tempi da 1 a 4 sono stati calcolati numeri indici a base mobile sempre pari a 1,02 (tutti e 4 gli indici a base mobile sono uguali quindi). Allora...

- o Anche il numero indice del tempo 4 con base tempo 0 è pari a 1,02
- o Il numero indice del tempo 4 con base tempo 0 è circa 1,082
- o Il numero indice del tempo 4 con base tempo 1 è circa 1,082
- o Il numero indice del tempo 4 a base tempo 1 è pari a 1,00496

Per la circolarità, o transitività,  ${}_0I_4 = {}_0I_1 \cdot {}_1I_2 \cdot {}_2I_3 \cdot {}_3I_4$

$$1,02 \times 1,02 \times 1,02 \times 1,02 = 1,082$$

quindi la risposta giusta è la 2

9. I prezzi e le quantità di quattro beni, A, B, C e D sono, al tempo 0 e al tempo t, indicati nella seguente tabella.

Calcolare il numero indice sintetico di Paasche tra 0 e t

- o Vale circa 1,079
- o Vale circa 1,083
- o Vale circa 1,0645
- o Non si può calcolare esattamente, ma è certamente inferiore a 1

la sommatoria di  $p_t q_t$  è 121469. La sommatoria di  $p_0 q_t$  è 112590,7 quindi l'indice di Paasche è 1,0789. La risposta giusta è dunque la 1

	prezzi			quantità		
	a	b	c	a	b	c
0	243,83	2,20	2013,83	168	1894	24
t	277,94	2,23	2094,45	185	2297	31

10. Se un carattere quantitativo X è indipendente in media da un carattere Y, allora...

- o L'indice V di Cramer risulta pari a 0
- o L'indice V di Cramer potrebbe anche essere 1
- o L'indice V di Cramer risulta pari a 1
- o L'indice V di Cramer non è calcolabile

L'enunciato ci dice che il rapporto di correlazione è 0, ovvero che le medie di X sono le stesse per ciascuna modalità di Y

Ma questo non ci dà informazioni sull'indice V di Cramer (viceversa se l'indice V di Cramer fosse 0 sarebbe 0 anche il rapporto di correlazione)

Se ad esempio X è numero di figli e Y è la nazionalità di 100 donne, con 2 modalità "italiana" o "non italiana", nel caso della seguente distribuzione di frequenza

	ita	non ita
1 figlio	0	10
2 figli	80	0
3 figli	0	10

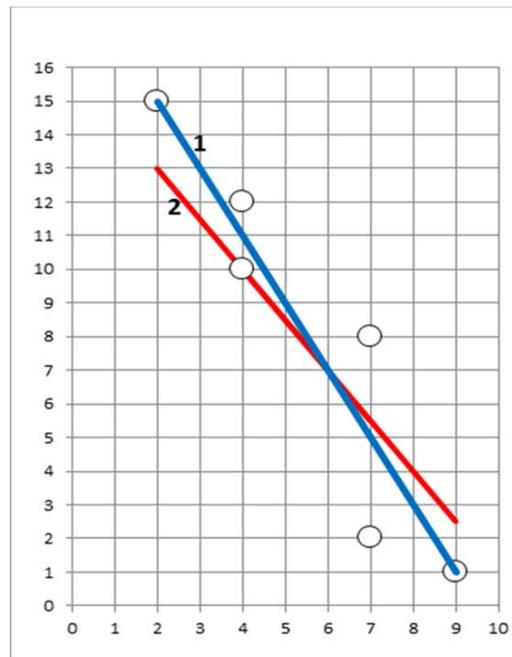
avremmo che Y dipende perfettamente da X (e quindi V di Cramer =1), ma la media del numero di figli è la stessa per le "ita" e le "non ita" (quindi  $r_{corr}=0$ )

La risposta corretta era pertanto la 3

Anche la risposta 4 è stata giudicata comunque corretta, nel senso che "se so che il rapporto di correlazione=0 non posso dire niente sul V di Cramer"

**Regressione** Sono state misurate le variabili X (in ascissa) e Y su 6 osservazioni. Sappiamo che una delle due rette disegnate è la retta dei minimi quadrati

1. E' la "1" o la "2"?
2. Calcolare i coefficienti a e b della retta dei m. q.
3. Quanto vale  $R^2$



Dal grafico si può dedurre che:  
 la retta 1 ha equazione  $y=19-2x$   
 la retta 2 ha equazione  $y=16-1,5x$   
 quindi:

x	y	yretta1	scartiq1	yretta2	scartiq2
2	15	15	0	13	4
4	10	11	1	10	0
4	12	11	1	10	4
7	2	5	9	5,5	12
7	8	5	9	5,5	6
9	1	1	0	2,5	2

20                      29

Dunque la retta 1 è quella dei m.q.  
 e la devianza residua è 20  
 dato che la devianza degli y (o devianza totale) è 154  
 la devianza di regressione è 134  
 e quindi  $r$ -quadrato è  $134/154$   
 0,87013

