

FAC-SIMILE DI QUESITI CHE VERRANNO PROPOSTI NELLA PROVA SCRITTA

1) Si osservi la tabella che segue:

COMUNE DI FERRARA - FAMIGLIE PER NUMERO DI COMPONENTI E TIPOLOGIA AL 31/12/2001 (Fonte: Servizio Statistica Comune di Ferrara)

TIPOLOGIA FAMIGLIARE	NUMERO COMPONENTI								TOTALE
	1	2	3	4	5	6	7	8 o +	
Persone sole	17.736	0	0	0	0	0	0	0	17.736
Coppie coniugate	0	11.656	0	0	0	0	0	0	11.656
Coppie coniugate+figli	0	0	10.850	4.883	523	77	14	6	16.353
Coppie coniugate+figli+altri	0	0	0	682	465	145	23	7	1.322
Coppie coniugate+altri	0	0	546	58	9	2	1	0	616
Un genitore+figli	0	4.055	924	107	12	2	0	0	5.100
Un genitore+figli+altri	0	0	588	464	185	36	7	7	1.287
Coppie conviventi	0	1.429	0	0	0	0	0	0	1.429
Coppie conviventi+figli	0	0	703	192	26	6	5	0	932
Coppie conviventi+figli+altri	0	0	0	24	6	3	0	2	35
Coppie conviventi+altri	0	0	19	5	2	2	0	0	28
Intestatario+genitori (anche con altri)	0	289	103	21	4	1	0	0	418
Altra tipologia	0	886	142	50	19	3	1	0	1.101
TOTALE	17.736	18.315	13.875	6.486	1.251	277	51	22	58.013

Qual è l'unità statistica ?

- a) ciascun componente della famiglia;
- b) la tipologia familiare;
- c) la famiglia.

RISPOSTA ESATTA: c).

Commento

Lo scopo di questo quesito è valutare, da parte dello studente, la capacità di leggere correttamente una tabella a doppia entrata e, al tempo stesso, valutare se lo studente ha ben chiaro il concetto di "unità statistica". Per il resto, si tratta di un mero quesito "di attenzione": basta infatti leggere il titolo della tabella ("FAMIGLIE PER NUMERO DI COMPONENTI E TIPOLOGIA AL 31/12/2001") per essere informati del fatto che le unità statistiche conteggiate nella tabella sono le famiglie. Le risposte non corrette (rispettivamente, "ciascun componente della famiglia" e "tipologia familiare") alludono invece, più o meno direttamente, alle due variabili alle quali si fa riferimento nella tabella, cioè la "tipologia familiare" e il "numero di componenti della famiglia".

2) Se vogliamo *destagionalizzare* una serie storica, l'utilizzo di una media mobile *ponderata* comporta che:

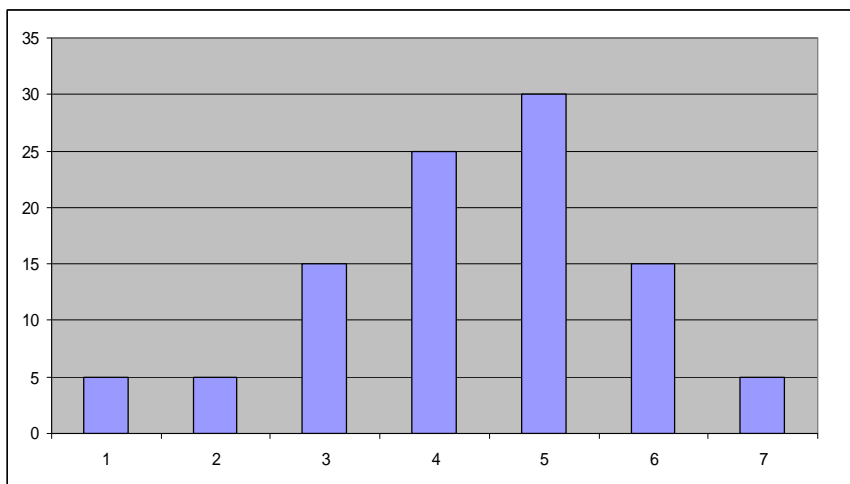
- a) i dati risultano *meno modificati* rispetto al caso in cui avessimo utilizzato una media mobile semplice;
- b) i dati risultano *più modificati* rispetto al caso in cui avessimo utilizzato una media mobile semplice;
- c) non c'è una regola generale: il fatto che i dati siano più o meno modificati dipende da come è fatta la serie storica.

RISPOSTA ESATTA: a).

Commento

Lo scopo di questo quesito è valutare, da parte dello studente, la conoscenza della tecnica delle medie mobili per la destagionalizzazione di una serie storica. Per quanto riguarda le medie mobili, l'applicazione di una qualsivoglia struttura di ponderazione ha lo scopo di rendere meno forte, nel risultato finale, la modificazione della serie storica originaria: pertanto, l'utilizzo di una media mobile ponderata comporta che, rispetto a quanto accadrebbe con una media mobile semplice, i dati siano *meno modificati*. Per quanto riguarda le risposte non corrette: b) è ovviamente il contrario della risposta corretta, mentre la c) afferma, in modo erroneo, che la maggiore o minore modificazione della serie storica dipenderebbe dalla conformazione della serie storica stessa, mentre in realtà dipende solo dall'aver applicato o meno una ponderazione alla media mobile.

3) Con riferimento a una certa settimana dell'anno 2011, 100 pazienti di un reparto ospedaliero sono stati classificati secondo il numero di giornate di degenza durante quella stessa settimana (una "giornata di degenza" prevede che il paziente abbia pernottato nell'ospedale). Il numero di giornate di degenza varia tra 1 e 7. La distribuzione dei 100 pazienti per numero di giornate di degenza è rappresentata dal seguente grafico.



Con riferimento alla distribuzione rappresentata dal grafico, una sola delle seguenti affermazioni è vera. Quale?

- a) La moda è pari a 4, mentre la mediana è pari a 5;
- b) La moda è pari a 5, mentre la mediana è pari a 4;
- c) Moda e mediana coincidono, e sono entrambe pari a 4.

RISPOSTA ESATTA: b).

Commento

Con questo quesito, si vuole valutare la conoscenza, da parte dello studente, dei concetti di moda e di mediana di una distribuzione statistica. Il fatto di avere indicato 100 pazienti fa sì che le frequenze assolute presentate nel grafico siano immediatamente leggibili come se fossero valori percentuali. Per quanto riguarda la moda, il grafico esprime in modo chiaro ed evidente che il valore più frequente della distribuzione, e quindi la moda, è il valore “5”. Per quanto riguarda la mediana, la semplice somma, facilmente effettuabile mentalmente, delle modalità “1”, “2”, “3” e “4” è pari a $5+5+15+25=50$. Pertanto, in corrispondenza del valore “4” è stato accumulato il 50% della distribuzione, e pertanto il valore “4” è il valore della mediana richiesto dal quesito. Quindi, la risposta corretta è la b). La risposta non corretta a), semplicemente, inverte tra loro i due risultati corretti, mentre la risposta c) allude, in modo erroneo, al fatto che mediana e moda coinciderebbero (mentre, come abbiamo visto, ciò non è vero), e sarebbero entrambe pari al valore “4” (mentre, come abbiamo visto, quest’ultimo è il valore della mediana, ma non della moda, che è invece “5”).

4) Supponiamo di dover effettuare un test di ipotesi, a due code, utilizzando la distribuzione normale standardizzata (Z). Supponiamo poi che il valore-soglia prescelto per la statistica-test sia pari a $+1,96$ per la parte positiva della distribuzione, e a $-1,96$ per la parte negativa. Come noto, questo valore della Z corrisponde, per un test a due code, a un livello di significatività $\alpha=0,05$ (valore che andrebbe dimezzato nel caso del test a una sola coda). Come si caratterizzerà, in questo caso, la “regione di rifiuto” del test?

- a) Sarà compresa tra 0 e $+1,96$;
- b) Sarà compresa tra $-1,96$ e $+1,96$;
- c) Si troverà alla destra del valore $+1,96$ fino a $+\infty$, e alla sinistra del valore $-1,96$ fino a $-\infty$.

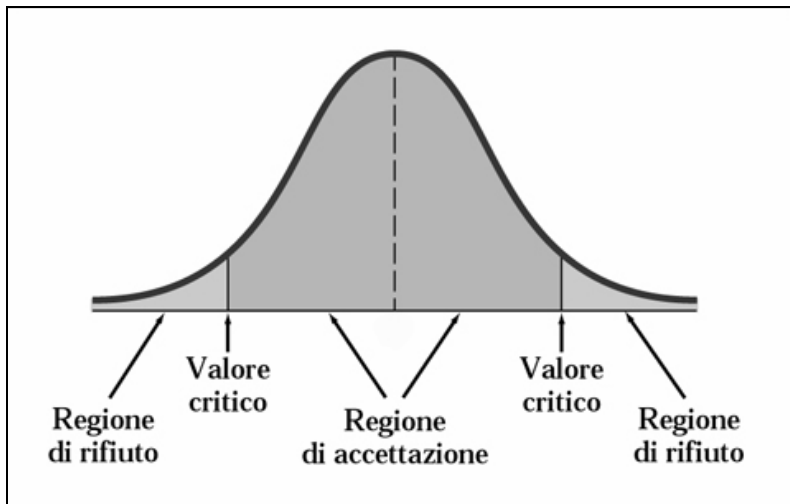
RISPOSTA ESATTA: c).

Commento

Con questo quesito, si vuole valutare, da parte dello studente, la comprensione della logica generale dei test di ipotesi, e in particolare la comprensione del concetto di “regione di rifiuto” dell’ipotesi nulla, e del concetto di “test a una oppure due code”. Il quesito, per essere reso immediatamente comprensibile in termini pratici, fa riferimento a un test che fa uso della distribuzione normale standardizzata (Z), per la quale sia stato già prescelto un valore soglia della statistica-test pari a $+1,96$, per la parte positiva della distribuzione, e a $-1,96$ per la parte negativa.

Il quesito è stato volutamente facilitato, fornendo anche spiegazioni sul livello di significatività (per un test a due code) corrispondente al valore soglia prescelto.

Come si può facilmente vedere osservando la figura sottostante, la “regione di rifiuto” di un test di ipotesi è data da quel sottoinsieme del supporto della distribuzione (in questo caso la normale standardizzata, Z) in corrispondenza del quale la *somma delle aree sottese dalla curva* è sicuramente *più piccola* del livello di significatività prescelto (in questo caso, **0,05**). Pertanto, la risposta corretta al quesito è la c), cioè: poiché stiamo parlando di un test a due code, la regione di rifiuto è composta da due parti separate: quella *inferiore* al valore negativo $-1,96$ (fino a meno infinito) e *superiore* al valore positivo $+1,96$ (fino a più infinito).



Per quanto riguarda le due risposte non corrette: ovviamente, la risposta b) fa riferimento all'esatto contrario di ciò che era richiesto dal quesito, cioè alla “regione di accettazione” dell'ipotesi nulla, che è appunto compresa tra $-1,96$ e $+1,96$. La risposta a) è totalmente priva di senso, perché fa riferimento all'eventuale “regione di accettazione” (e non di rifiuto) per un test a una sola coda (positiva), che sarebbe effettivamente compresa tra 0 e $+1,96$, visto che un test a una sola coda (positiva) richiede l'uso della sola parte positiva della distribuzione di riferimento. Tuttavia, come abbiamo visto, il quesito che è stato presentato faceva riferimento a un *test a due code*.