

Statistica - Esercizi

9 aprile 2020

Vero o Falso?

- a. Un campione con $n > 30$ è rappresentativo dell'intera popolazione
- b. L'insieme degli studenti iscritti all'Università di Trieste è un campione rappresentativo di tutta la popolazione degli studenti universitari in Italia
- c. Se si osservano due variabili qualitative nominali X e Y su un collettivo di n unità statistiche posso ottenere la distribuzione di frequenza della variabile $X|Y = y$
- d. In una distribuzione in classi le frequenze relative diviso l'ampiezza di ogni classe ci restituiscono la densità di frequenza
- e. La moda è il picco di una distribuzione
- f. La media non può mai essere negativa
- g. Mediana, primo e terzo quartile non possono mai coincidere
- h. Il valore della mediana è sempre uguale a 0.5
- i. La mediana può essere calcolata per tutti i tipi di variabili

Esercizio 1 Per ognuna dei seguenti punti identificare la popolazione di riferimento dell'indagine e determinare se le variabili siano qualitative (nominali o ordinali) oppure quantitative (discrete o continue).

- a. Numero di telefonate giornaliere effettuate da una famiglia residente in Campania
- b. Ampiezza della casa di proprietà dei cittadini del comune di Trieste
- c. Spesa giornaliera in euro dei clienti di un centro commerciale
- d. Presenza di connessione e Internet all'interno delle aule dell'Università di Trieste
- e. Intenzioni di voto degli elettori alle prossime politiche in Italia
- f. Grado di soddisfazione (su scala da 1 a 5) sui servizi agli studenti dell'Università di Trieste
- g. Titolo di studio dei dipendenti della regione Friuli Venezia Giulia

Esercizio 2 Un'azienda produttrice di cibo per gatti sta pianificando di condurre un'indagine di mercato (via questionario) sulle abitudini di acquisto di cibo in scatola per animali domestici delle famiglie italiane. Tra le domande che l'azienda pensava di porre ci sono: 1) dove viene acquistato il cibo per animali principalmente?; 2) Quanti gatti possiedono?; 3) numero di confezioni di cibo per gatti acquistate alla settimana; 4) spesa media settimanale in euro per cibo per gatti.

- si descriva la popolazione di riferimento
- si elenchino alcune possibili modalità per le variabili che l'azienda vuole includere nel questionario
- Considerando le domande del questionario dell'azienda, quali variabili sono qualitative?
- Quali quantitative? Quali sono discrete?
- ...e per quali variabili sarebbe possibile costruire una distribuzione di frequenza in classi?

Esercizio 3 I valori di una variabile quantitativa continua variano da 11.2 a 98.6.

- Se i valori fossero raggruppati in 9 classi, si indichino i limiti inferiore e superiore di ciascuna classe
- Qale ampiezza per l'intervallo di ciascuna classe avete scelto?

Esercizio 4 I punteggi al test GMAT (fatto negli USA per verificare la preparazione matematica degli studenti che varia da 200 a 800) su un campione di 50 candidati ad un master in statistica per le scienze sociali ha indicato che nessuno di questi ha ottenuto meno di 450 punti. La distribuzione di frequenza in classi è stata costruita considerando i seguenti intervalli: da 450 a 499 (in cui si sono collocati 2 candidati); da 500 a 549 (18 candidati); e così via fino all'ultima classe da 700 a 749. Quali sono le percentuali dei:

- candidati con punteggio inferiore a 500?
- tra 500 e 549?
- inferiore a 550?
- inferiore a 750?

Esercizio 5 Una variabile qualitativa nominale rilevata su di una popolazione assume la seguente distribuzione di frequenza

Variabile	Frequenza	Freq. Relativa	%
A	11	0.22	...
B	30	...	...
C	...	...	...

- Completare la tabella.
- Costruire un diagramma a barre.
- Costruire un diagramma a torta.

Esercizio 6 *Costo del lavoro*

Un'indagine su di un campione di 3756 piccole imprese aveva lo scopo di rilevare il costo del lavoro sostenuto da queste. Abbiamo a disposizione la seguente tabella per la costruzione di un istogramma.

Classi di costo lavoro orario (euro)	Densità di frequenza relativa
(0, 12]	0.0125
(12, 30]	0.02
(30, 50]	0.02
(50, 100]	?

- Quali e quante sono le unità statistiche?
- Completare la tabella calcolando la frequenza assoluta delle imprese in ciascuna classe
- Calcolare i valori centrali di ciascuna classe
- Rappresentare i dati con un opportuno grafico

Esercizio 7 *(Dall'esame del 9 giugno 2015)*

Nell'Anno Accademico 2013/2014, ad un campione di studenti dell'ultimo anno dei corsi di laurea di Scienze Politiche (SP) e di Economia (EC) dell'università di Trieste è stato chiesto quanti esami avessero ancora da sostenere prima della laurea (da 1 a 5). Di seguito si riporta la tabella (di contingenza) tra corso di laurea e numero di esami da sostenere.

	1	2	3	4	5
SP	14	6	4	4	2
EC	9	5	6	4	3

- Utilizzare un opportuno grafico per rappresentare la variabile numero di esami ancora da sostenere.
- Calcolare mediana e quartili del numero di esami ancora da sostenere.
- Si determinino le distribuzioni di frequenza relativa della variabile numero di esami ancora da sostenere condizionata al corso di laurea.

Esercizio 8 I seguenti dati rappresentano la durata in ore di un campione di 40 lampadine prodotte da due aziende: Azienda A ed Azienda B

Azienda A					Azienda B				
684	697	720	723	821	819	836	888	897	903
831	835	848	852	852	907	912	918	942	943
859	860	868	870	876	952	959	962	986	992
893	899	905	909	911	994	1004	1005	1007	1015
922	924	926	926	938	1016	1018	1020	1022	1034
939	943	946	954	971	1038	1072	1077	1077	1082
972	977	984	1005	1014	1096	1100	1113	1113	1116
1016	1041	1052	1080	1093	1153	1154	1174	1188	1230

- Al fine di confrontare le 2 aziende costruire la distribuzione di frequenza assoluta e relativa per ciascuna azienda (servono le classi o no?). Suggerimento: si parta da 650 per entrambe le aziende.

-
- b. Costruire i grafici opportuni (usando le frequenze relative... perchè?)
 - c. Dalla visualizzazione dei grafici quale azienda sembra produca le migliori lampadine? si motivi la scelta
 - d. Si usi la distribuzione delle frequenze cumulate (relative o assolute?) per rispondere alla stessa domanda del punto d.