

Compiti tematici dai capitoli 2,3,4

a cura di Giovanni M. Marchetti

2016 ver. 0.8

1.

In un'indagine recente, i rispondenti sono stati classificati rispetto al sesso, lo stato civile e l'area geografica di residenza. I dati sono sintetizzati nella seguente tabella (nella quale M è usato per indicare il sesso maschile e F per indicare il sesso femminile):

Area	Single		Sposato		Totale
	M	F	M	F	
Nord-est	12	17	22	10	61
Nord-ovest	31	26	7	23	87
Centro	45	33	52	38	168
Sud	34	19	18	13	84
Totale	122	95	99	84	400

Qual è la proporzione di rispondenti non sposati?

- A) 0.543
- B) 0.510
- C) 0.620
- D) 0.305

Soluzione. Interpreta la proporzione come una probabilità. Casi possibili: 400 rispondenti. Casi favorevoli: single M o single femmina = $122 + 95 = 217$. Quindi la proporzione è $217/400 = 0.5425$ che arrotondato a 3 cifre è 0.543.

2.

Dati due eventi A e B , se almeno uno di loro è necessariamente vero si dice che A e B sono collettivamente esaustivi. Vero o Falso?

Soluzione. Due eventi sono esaustivi se la loro unione è l'intero spazio campionario. Ossia se l'evento "si verifica A oppure si verifica B " è l'evento certo. Quindi è Vero.

3.

In una recente indagine sulla fiducia dei consumatori, 160 rispondenti sono stati classificati in base al loro livello di fiducia e al loro titolo di studio:

Fiducia	Titolo di studio		
	Diploma	Laurea	Master
Bassa	13	17	15
Media	27	22	13
Alta	32	14	7

Supponiamo di estrarre a caso un consumatore: gli eventi "ha un master" e "ha un alto livello di fiducia" sono statisticamente indipendenti?

- A) No.
- B) Forse.

- C) Sì.
- D) Non ci sono sufficienti informazioni per rispondere.

Soluzione. Sono indipendenti se

$$P(\text{Master} \cap \text{Alta}) = P(\text{Master})P(\text{Alta})$$

Ora risulta che i casi possibili sono 160, mentre

$$P(\text{Master} \cap \text{Alta}) = 7/160, P(\text{Master}) = \frac{15 + 13 + 7}{160} = 35/160, P(\text{Alta}) = \frac{32 + 14 + 7}{160} = 53/160$$

Quindi poiché $7/160 \neq (35/160)(53/160)$ i due eventi non sono indipendenti.

4.

La regola empirica si applica a qualsiasi distribuzione, indipendentemente dalla sua forma, come guida per interpretare la distribuzione. Vero o Falso?

Rileggi la definizione: la regola empirica si applica se la popolazione è simile a una variabile normale (cioè Gaussiana). Falso.

5.

Se in una certa prova si ha che $P(A) = 0.7$ può accadere che $P(A \cup B) = 0.5$?

- A) Solo se A e B sono eventi incompatibili
- B) Solo se $B = \emptyset$
- C) SI è possibile
- D) NO mai

Soluzione.

- A) Falsa: per esempio perché se B è disgiunto da A , $P(A \cup B) = P(A) + P(B) \geq 0.7$.
- B) Falsa: se $B = \emptyset$, $P(A \cup B) = P(A) = 0.7$
- C) Falsa: vedi D).
- D) Vera: L'evento $A \cup B$ contiene sempre A e quindi si verifica sempre se si verifica A . Quindi $P(A \cup B) \geq P(A)$ e quindi non può mai succedere che $P(A \cup B) = 0.5 < P(A)$.

6.

Quale delle seguenti affermazioni descrive uno svantaggio nell'utilizzo del campo di variazione come misura di variabilità o di dispersione?

- A) Non è misurato nella stessa unità di misura dei dati.
- B) È molto influenzato dai valori anomali.
- C) Produce una misura della variabilità troppo elevata.
- D) Si può calcolare solo per variabili continue.

Soluzione.

- A) Falso: è $x_{\max} - x_{\min}$ e quindi è misurato nella stessa unità di misura dei dati.
- B) Vero: perché ovviamente basta un solo valore o molto grande o molto piccolo per farlo cambiare a differenza degli altri indici di variabilità.
- C) Falso: è una frase senza senso.
- D) Falso: anche la varianza si può calcolare solo per variabili continue e non per questo è uno svantaggio.

7.

La probabilità dell'intersezione di due eventi non può essere superiore alla somma delle loro probabilità. Vero o Falso?

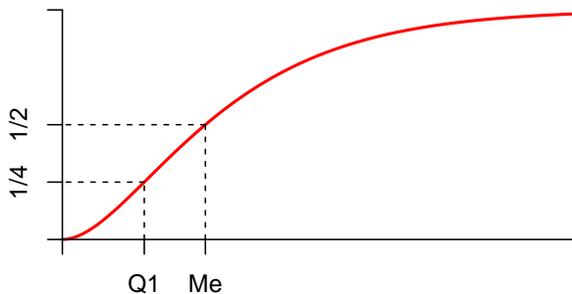
Soluzione. $P(A \cap B) = P(A) + P(B) - P(A \cup B)$ quindi $P(A \cap B) \leq P(A) + P(B)$. Vero.

8.

Il primo quartile è sempre minore o uguale alla mediana,

- A) Solo se la media è maggiore di zero
- B) Solo se la moda è maggiore della mediana
- C) SI è vero
- D) NO è falso

Soluzione. Per definizione $P(X \leq Q_1) = 0.25$ mentre $P(X \leq Me) = 0.5$, come si vede dalla funzione di ripartizione.



Quindi è chiaro che Q_1 non può essere più grande di Me altrimenti la probabilità di essere inferiore alla mediana sarebbe minore di 0.25.

9.

Quale delle seguenti affermazioni è vera?

- A) La covarianza è sempre maggiore del coefficiente di correlazione.
- B) La covarianza può essere uguale al coefficiente di correlazione.
- C) Né la covarianza né il coefficiente di correlazione possono essere pari a zero.
- D) Il coefficiente di correlazione è sempre più grande della covarianza.

Soluzione.

- A) Falso. La covarianza cambia al variare dell'unità di misura e quindi si può rendere più piccola del coefficiente di correlazione.
- B) Vero. Perché il coefficiente di correlazione è la covarianza fra le due variabili standardizzate.
- C) Falso. Evidentemente altrimenti non esisterebbe il concetto di incorrelazione.
- D) Falso. La covarianza cambia al variare dell'unità di misura e quindi si può rendere più grande del coefficiente di correlazione.

10.

Il coefficiente di correlazione lineare è sempre compreso nell'intervallo $[0, 1]$. Vero o Falso?

Soluzione. Falso. È sempre compreso in $[-1, 1]$.

11.

Il diagramma a barre è comunemente utilizzato per descrivere dati qualitativi. Vero o Falso?

Soluzione. Vero.

12.

Quale delle seguenti affermazioni è sempre vera per qualsiasi coppia di eventi A e B definiti in uno spazio degli eventi elementari S ?

- A) Se l'unione degli eventi A e B è l'insieme vuoto, allora sia A che B sono insiemi vuoti.
- B) Se gli eventi A e B sono collettivamente esaustivi, allora $A \cup B = \emptyset$.
- C) Se l'intersezione degli eventi A e B è l'insieme vuoto, allora A e B sono collettivamente esaustivi.
- D) Se gli eventi A e B sono mutuamente esclusivi e collettivamente esaustivi, allora l'unione di A e B non è necessariamente uguale a S .

Soluzione. A) Vera: $A \cup B = \emptyset$ non può avvenire se uno dei due eventi contiene degli elementi. B) Falso: la definizione è $A \cup B = S$. C) Falso: $A \cap B = \emptyset$ implica che gli eventi sono disgiunti non esaustivi. D) Falso: in questo caso gli eventi sono una partizione di S e quindi la loro unione è necessariamente uguale a S .

13.

Supponiamo di lanciare due dadi e si consideri la somma dei punti dei due dadi. Sia A l'evento "si osserva un numero pari" e B l'evento "si osserva un numero maggiore di 7". Cos'è $\bar{A} \cup \bar{B}$?

- A) $\{2, 3, 4, 5, 6, 7\}$
- B) $\{2, 4, 6\}$
- C) $\{2, 3, 4, 5, 6, 7, 9, 11\}$
- D) $\{3, 5, 7\}$

Soluzione. $A = \{2, 4, 6, 8, 10, 12\}$ e quindi $\bar{A} = \{3, 5, 7, 9, 11\}$. Nota che non può avvenire che la somma di due dadi sia 1!

$B = \{8, 9, 10, 11, 12\}$ quindi $\bar{B} = \{2, 3, 4, 5, 6, 7\}$. Conclusione: $\bar{A} \cup \bar{B} = \{2, 3, 4, 5, 6, 7, 9, 11\}$.

14.

Se due insiemi di dati hanno lo stesso campo di variazione, allora:

- A) I due insiemi hanno la stessa varianza.
- B) La distanza tra il valore più piccolo e il valore più grande nei due insiemi ha lo stesso valore.
- C) I due insiemi hanno lo stesso scarto interquartile.
- D) il valore più piccolo e il valore più grande nei due insiemi sono gli stessi.

Soluzione. A) Falso in generale. Per esempio $x_1 = 0, x_2 = 1$ ha campo di variazione 1. Anche $y_1 = 1, y_2 = 1.5, y_3 = 2$ ha campo di variazione 1. Tuttavia $\text{var}(x) = 1/2$ e $\text{var}(y) = 1/4$. B) Vero per definizione. C) Falso come per A). D) Falso: non è necessario che $x_{\max} = y_{\max}$ e $x_{\min} = y_{\min}$ per ottenere lo stesso campo di variazione.

15.

La differenza interquartile corrisponde al campo di variazione del 50% centrale dei dati. Vero o Falso?

Soluzione. Vero: la differenza interquartile è $Q_3 - Q_1$ e

$$P(Q_1 \leq X \leq Q_3) = 0.5$$

Quindi è il campo di variazione del 50% centrale dei dati.

16.

In una ricerca di marketing, ai consumatori è stato dato uno fra 4 marche di detersivo per lavastoviglie ed è stato chiesto di usarlo per un mese. Allo scadere del periodo è stato chiesto un giudizio sul detersivo utilizzato in termini di qualità complessiva. I risultati sono i seguenti.

Marca	Giudizio				Totale
	Mediocre	Medio	Discreto	Buono	
A	5	17	11	10	43
B	14	26	8	18	66
C	10	23	11	17	61
D	11	19	7	5	42
Totale	40	85	37	50	212

Qual è la proporzione di consumatori che ha valutato il proprio detersivo discreto o buono?

- A) 0.39
- B) 0.49
- C) 0.29
- D) 0.41

Soluzione.

Casi possibili 212. Casi favorevoli: $37 + 50 = 87$. Quindi la proporzione è $87/212 = 0.410$.

17.

Fra i consumatori che hanno valutato il prodotto A, qual è la proporzione che lo ha giudicato mediocre?

- A) 0.189
- B) 0.116
- C) 0.125
- D) 0.024

Soluzione. Casi possibili: 43. Casi favorevoli: 5. Quindi la proporzione è $0.1163 \simeq 0.116$.

18.

Una tavola con due righe, A_1 e A_2 , e due colonne, B_1 and B_2 , riporta le seguenti probabilità congiunte:

$$P(A_1 \cap B_1) = 0.10, P(A_1 \cap B_2) = 0.30, P(A_2 \cap B_1) = 0.05, P(A_2 \cap B_2) = 0.55.$$

Quanto vale $P(B_1)$?

- A) 0.15
- B) 0.60
- C) 0.40
- D) 0.85

Soluzione. La tavola è

	B_1	B_2
A_1	0.10	1.30
A_2	0.05	0.55

La $P(B_1)$ è il totale della prima colonna cioè $P(B_1) = P(A_1 \cap B_1) + P(A_2 \cap B_1) = 0.10 + 0.05 = 0.15$.

19.

L'operazione di standardizzazione di una variabile statistica consiste nel dividere gli scarti dalla media aritmetica per la varianza. Vero o Falso?

Soluzione. Falso: $z = (x - \mu)/\sigma$ NON $z = (x - \mu)/\sigma^2$.

20.

Per tutti gli individui di un collettivo viene rilevato il codice di avviamento postale del comune di residenza. Quale delle seguenti definizioni è appropriata per questi dati?

- A) Dati quantitativi
- B) Dati qualitativi
- C) Serie storica
- D) Dati numerici

Soluzione

- A) Falso: non si possono eseguire operazioni algebriche con un senso. Solo verificare se una altro codice è uguale o diverso.
- B) Vero.
- C) Falso. Non ci sono riferimenti temporali.
- D) Falso: vedi A).

21.

Se $P(A) = 0.25$ e $P(B) = 0.75$, allora gli eventi A e B sono collettivamente esaustivi. Vero o Falso?

Soluzione. Falso. Esempio: in un mazzo di carte sia A "pesco una carta di cuori" e B "pesco una carta non di picche". A e B non sono esaustivi.

22.

Si considerino due investimenti possibili aventi lo stesso tasso di rendimento atteso. Negli ultimi mesi l'investimento A ha avuto un prezzo medio di chiusura di 14.00 e una deviazione standard di 4.00. L'investimento B un prezzo medio di chiusura di 58.00 e una deviazione standard di 15.00. Il valore di mercato dell'investimento A ha una variabilità relativa maggiore di quello dell'investimento B . Vero o Falso?

Soluzione. Per valutare la variabilità relativa si usa il coefficiente di variazione $CV = \sigma/\mu$. Quindi

$$CV_A = 4/14 = 0.286, \quad CV_B = 15/58 = 0.259$$

Quindi $CV_A > CV_B$. Vero.

23.

Considera la seguente distribuzione di frequenza.

Modalità	Frequenza	Freq. % Cumulate
fino a 584.0	1	4
584.0 - 1774.4	*	64
1774.4 - 2964.8	4	80
2964.8 - 4155.2	3	92
4155.2 - 5345.6	1	96
più di 5345.6	1	100

Qual è la frequenza mancante nella posizione occupata dall'asterisco?

- A) 15
- B) 16
- C) 3
- D) 25

Soluzione.

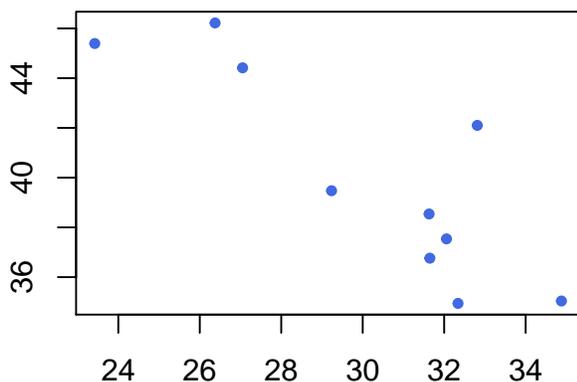
La frequenza relativa della classe $(584.0, 1774.4]$ è $0.64 - 0.04 = 0.6$. La frequenza assoluta della stessa classe è x e la frequenza totale è $10 + x$. Quindi

$$0.6 = x/(10 + x)$$

da cui si ottiene l'equazione $0.6(10 + x) = x$ che risolta dà $x = 15$.

24.

Sia dato lo scatter seguente.



Quale tra i valori seguenti potrebbe essere il coefficiente di correlazione per le due variabili?

- A) -0.8
- B) 0.0
- C) -0.3
- D) -1 .

Soluzione. È chiaro che il coefficiente di correlazione deve essere negativo. Ovviamente si scartano la B) (incorrelazione) e la D) (perfetto allineamento). Tra A) e C) si sceglie A) perché l'associazione lineare è più vicina a -1 che a 0 . Qui ci vuole un po' d'occhio.

25.

La varianza di $X + c$, dove c è una costante, è pari alla varianza di X più la costante. Vero o Falso?

Soluzione. Falso: $\text{var}(X + c) = \text{var}(X)$.