

3^a FASE di UN'INDAGINE STAT.

Esempio n. 1

TAB. II.3.1. Distribuzione della popolazione residente attiva (in migliaia di unità) per settore di attività (Censimento generale della popolazione italiana, 1961 e 1971)

m_i	1961		1971	
	f_i (1)	p_i (2)	f_i (1)	p_i (2)
Agricoltura, foreste, caccia e pesca	5.693	0,291	3.242	0,172
Industrie	7.963	0,406	8.350	0,443
Altre attività	5.936	0,303	7.238	0,385
	19.592	1	18.830	1

Esempio n. 2

Esercizio 3 (8/10/1993)

Quattrocento (400) studenti della provincia di Verona in possesso di abbonamento ferroviario o AMT sono stati classificati secondo il tipo di scuola (i non abbonati non sono stati considerati).

Abbonamento	Scuola frequentata		
	Ist. Professionale	Ist. Tecnico	Liceo
F.S.	0.12	0.17	0.12
A.M.T.	0.05	0.04	0.18
Entrambi	0.13	0.09	0.10

Determinare se esiste dipendenza tra i due caratteri esaminati, misurandone il grado in caso affermativo.

CANAZZI 99 Esempio n. 3

3) Le frequenze congiunte di due mutabili statistiche sono riportate nella seguente tabella

	L	M	N
A	70	30	0
B	20	60	20
C	10	90	100

calcolare:

- la probabilità che si verifichi l'evento A oppure B (l'uno o l'altro);
- la probabilità dell'evento C sapendo che si è verificato l'evento M ovvero $P(C/M)$;
- la probabilità che si verifichi l'evento A oppure L (l'uno o l'altro);
- la probabilità che si verifichino insieme l'evento B e M (l'uno e l'altro);
- mostrare come il risultato conseguito al punto d) si può ottenere applicando i teoremi del calcolo delle probabilità.

Esempio n. 4
Esercizio 1.

Tabella 1. Laureati nella Facoltà di Economia e Commercio, divisi secondo il punteggio e l'indirizzo seguito, nell'A.A. 89-90.

Indirizzo	Voto di Laurea			
	≤95	96-105	106-110	lode
Aziendale	2	8	4	1
Economico	1	4	1	0
Giuridico	4	7	2	1
Tradizionale	6	2	3	2

- a) Dare la rappresentazione grafica della distribuzione dei laureati classificati per indirizzo di studi seguito.
 b) Calcolare la mediana, la media, la varianza ed il coefficiente di variazione dei voti (la lode vale 115).
 c) Spiegare (in non più di 5 righe) se è più opportuno usare, come indice di localizzazione, la mediana oppure la media aritmetica, tenuto conto che i dati a disposizione sono quelli della tabella.

Esempio n. 5 STATISTICA I - 7 GENNAIO 1999

1) Gli abbonati X ad un giornale negli ultimi anni Z sono stati i seguenti (in migliaia)

Anno z_i	1990	1991	1992	1993	1994	1995	1996	1997	1998
x_i	10	14	18	20	22	25	30	32	35

Sui valori di X determinare:

- a) gli indici di variazione a base fissa 1990;
 b) gli indici di variazione a base mobile;
 c) la media annua degli indici a base mobile, giustificando la scelta.

Esempio n. 6 27/06/1994

2) Un prodotto di serie viene commercializzato in confezioni di 6 unità. Un controllo statistico della qualità prevede l'estrazione casuale senza reinserimento di 500 confezioni dall'intera produzione costituita di un milione di confezioni. La classificazione del campione secondo il numero di pezzi difettosi (x) trovati in ogni confezione ha dato i seguenti risultati

x	0	1	2	4	5	6
$f(x)$	215	190	45	30	15	5

- a) Rappresentare graficamente la variabile;
 b) stimare la frazione di pezzi difettosi presente nell'intera produzione;
 c) verificare l'ipotesi, al livello di significatività dell'1%, che la distribuzione di x provenga da una popolazione ove la frazione di difettosi sia del 10%.

Esempio n.7

3) Data la seguente tabella di frequenze a doppia entrata

25/01/1999

3

x	y		
	5	7	9
1	10	20	10
2	20	60	20
3	10	20	30

- si verifichi quale valore assume il coefficiente di correlazione lineare r nell'ipotesi di indipendenza fra X e Y ;
- rispettando le distribuzioni marginali, costruire la tabella nell'ipotesi di massima dipendenza;
- calcolare il coefficiente r sulla tabella di cui in b).
- giustificare i risultati di r ottenuti.

Es.3. Sulla seguente distribuzione congiunta di frequenze

23/09/99

x	y		
	0	2	4
1	40	10	0
3	20	60	20
5	0	20	30

- determinare i parametri della retta di regressione;
- calcolare il rapporto di determinazione come rapporto fra devianze;
- giudicare della significatività di r al livello del 4%.

Esempio n.8

Esercizio 2 27/10/1993

Mille candidati che partecipano ad un test devono dattiloscivere una relazione. Per ogni candidato viene rilevato il numero di errori Y ed il tempo di esecuzione X (in minuti). I dati ottenuti sono riportati nella seguente tabella.

Y	X		
	5 - 15	15 - 20	20 - 30
0	0	80	20
1 - 5	50	420	30
6 - 10	50	200	150

- Calcolare e interpretare il coefficiente di correlazione.
- Verificare la significatività ($\alpha=1\%$) del coefficiente trovato.
- Determinare i parametri della retta interpolante $y = a + bx$.

Esempio n.9

ESAME DI STATISTICA I - 18 dicembre 1996

1. Sui seguenti valori

$x_{i-1} < x \leq x_i$	0 - 1	1 - 5	5 - 7	7 - 8	8 - 10
$f(x)$	10	100	40	30	20

200

- dare una rappresentazione grafica della distribuzione di ~~probabilità~~ **frequenze**;
- calcolare media aritmetica, moda, varianza, mediana e 20-esimo percentile;
- giudicare la simmetria.

Esempio n.10 STATISTICA I - 17 febbraio 1999

Es.1. Da un'indagine statistica i lettori di giornali per età sono risultati i seguenti (le classi d'età sono chiuse a destra)

età	6-10	10-20	20-50	50-70	70-85
numero	20	200	420	210	150

Calcolare:

- a) l'età mediana e quella media (aritmetica);
- b) la classe d'età modale interpretandone il significato;
- c) rappresentare graficamente la distribuzione delle età;
- d) il coefficiente di variazione;
- e) esprimere un giudizio sulla simmetria della distribuzione usando i soli indicatori trovati al punto a).

STATISTICA I - 31 maggio 1999

1) Il numero di decessi $f(x)$ per età x è indicato nella seguente tabella

$a < x \leq b$	0 - 1	1 - 3	3 - 7	7 - 13	13 - 15
$f(x)$	250	50	400	900	400

Calcolare:

- a) l'età media (aritmetica), la mediana e la moda;
- b) il coefficiente di variazione;
- c) rappresentare graficamente la distribuzione di probabilità;
- d) indicare in quale età la mortalità annua è più elevata.

STATISTICA I - 28 giugno 1999

1) Il numero di batterie esaurite $f(x)$ rilevate in funzione della durata in mesi x è indicato nella tabella seguente

x	0-2	2-8	8-12	12-20
$f(x)$	200	800	600	400

Della durata x calcolare:

- a) i centri di ordine zero, uno e due indicando il significato di ciascuno di essi;
- b) rappresentare graficamente la distribuzione di frequenze;
- c) il periodo in cui la possibilità di esaurimento mensile appare più elevata indicando se il risultato ottenuto rientra fra gli indici statistici noti.

COMMENTARE BREVEMENTE LE PROCEDURE E MOSTRARE COME SI SONO OTTENUTI I RISULTATI

19 GIUGNO
2006

Esercizio 3. In un controllo statistico della produzione 25 pezzi di serie estratti a caso sono stati sottoposti a due prove, una relativa alla durata (A) e l'altra relativa al livello qualitativo (B). I risultati ottenuti in ciascuna prova prevedevano i giudizi p =positivo; n =negativo e sono stati i seguenti:

Prova	Risultati																									
A	n	n	p	p	n	p	p	p	p	n	n	n	p	p	p	n	n	n	p	n	p	p	n	n	n	
B	n	n	p	n	n	p	p	n	p	n	p	n	p	p	p	n	n	n	p	n	n	n	n	n	p	n

a) si costruisca la distribuzione congiunta (tabella a doppia entrata)

dei risultati delle due prove, la distribuzione marginale di A e quella di B, facendo riferimento per tutte alle frequenze assolute;

25 SETTEMBRE
2006

Esercizio 4. Un gruppo di aspiranti a coprire un dato incarico è stato sottoposto ad un test per rilevare la conoscenza della *Statistica* e dell'*Economia Aziendale*. I risultati, negativo (n) o positivo (p) sono stati i seguenti:

Prova	Risultati																								
Statistica	p	p	n	n	n	p	p	p	p	n	n	n	p	p	p	n	n	n	p	n	n	n	p	n	
Ec.Aziend.	n	n	n	p	n	p	p	n	p	n	p	n	p	p	p	n	n	n	p	n	n	n	p	n	

a) Costruire una tabella a doppia entrata che riporti le frequenze dei risultati.

8 GENNAIO
2007

Esercizio 3. Il tempo libero in ore disponibile in una giornata feriale da parte di un campione rappresentativo di 200 veronesi ultra sessantenni (80 maschi e 120 femmine) è risultato il seguente (la distribuzione per sesso è percentuale):

Ore tempo libero	0-2	2-4	4-6	6-8	8-10	totale	Numero
maschi	2,50	15,00	31,25	27,50	23,75	100,0	80
femmine	5,00	12,50	50,00	25,00	7,50	100,0	120

Determinare:

a) la distribuzione di frequenze assolute e percentuali delle ore di tempo libero per l'intero gruppo di anziani;

24 SETTEMBRE
2007

Esercizio 4. Nella seguente tabella sono state raccolte le percentuali di *disaccordo* oppure di *accordo* nei confronti della segregazione razziale (S) ottenute da un campione di cittadini americani bianchi suddivisi secondo il livello di istruzione (I).

Segregazione razziale (S)	% per Istruzione (I)	
	fino alla media superiore	Università
Disaccordo	68,5	89,5
D'accordo	31,5	10,5
Totale	100	100
n. risposte	1200	800

- a) Dopo aver costruito la tabella a doppia entrata delle frequenze assolute, verificare se esiste una connessione statisticamente significativa tra le due mutabili (livello di significatività 5%);
- b) calcolare e commentare il coefficiente di contingenza C.

22 GENNAIO
2007

Esercizio 1. Data la seguente tabella di frequenze

SESSO	TEST HIV	FASCIA D'ETA'		
		GIOVANE	ADULTO	ANZIANO
maschio	positivo	25	56	11
	negativo	22	76	26
femmina	positivo	38	59	8
	negativo	15	26	38

- a) fornire le distribuzioni univariate di frequenze e di probabilità di sesso, test HIV e fascia d'età;
- b) costruire la tabella di frequenze a doppia entrata per sesso e test HIV;