Esercitazione 1

CdL in Scienze Politiche

Statistica Sociale

1 Esercizio 1

La tabella sottostante riporta la distribuzione del numero di furti in Sardegna secondo la provincia di residenza:

Provincia	Furti
Cagliari-Hinterland	50
Sud Sardegna	300
Oristano	200
Nuoro	350
Sassari	100

Risolvi i punti seguenti:

1. calcola le frequenze relative
2. calcola le frequenze relative cumulate
3. calcola le frequenze percentuali
4. calcola le frequenze percentuali cumulate
5. quale tipologia di grafico è appropriata per rappresentare la distribuzione dei furti tra le province?
6. rappresenta graficamente la distribuzione.

1.1 Soluzioni

Provincia	Frequenze Assolute	Relative	Rel. cumulate
Cagliari-Hinterland	50	0.05	0.05
Sud Sardegna	300	0.3	0.35
Oristano	200	0.2	0.55
Nuoro	350	0.35	0.9
Sassari	100	0.1	1
Totale	1000		

		Frequenze relative		Frequenze percentuali	
Provincia	Furti	Relative	Cumulate	Percentuali	Cumulate
Cagliari-Hinterland	50	0.05	0.05	5%	5%
Sud Sardegna	300	0.30	0.35	30%	35%
Oristano	200	0.20	0.55	20%	55%
Nuoro	350	0.35	0.90	35%	90%
Sassari	100	0.10	1	10%	100%
Totale	1000	1		100%	

Dato che si tratta di una **variabile categoriale nominale** potremmo utilizzare principalmente due tipologie di grafico: il **grafico a torta** e il **grafico a barre**. Costruiamo il **grafico a barre**:

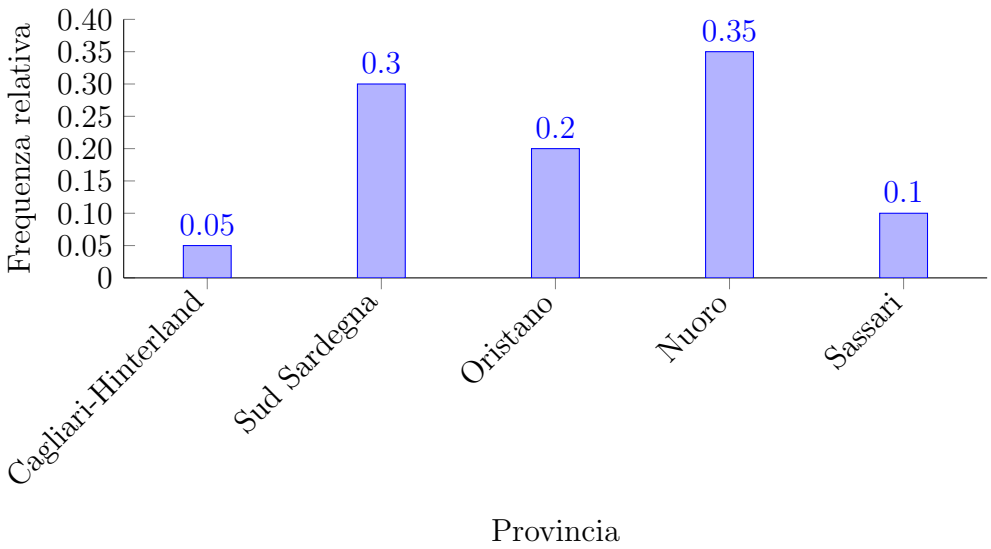

Figura 1: Distribuzione dei furti per provincia secondo le frequenze relative.

2 Esercizio 2

Si considerino i giorni di assenza da scuola di 7 studenti:

$$11, 5, 3, 5, 14, 7, 4$$

Calcola:

1. la media
2. la deviazione standard
3. la mediana

2.1 Soluzioni

- la media campionaria è pari a $\bar{y} == 7$.
- la deviazione standard campionaria $s = 4.04$
- la mediana è pari a 5

3 Esercizio 3

I risultati di un'indagine sul numero di case per famiglia sono riportati nella tabella seguente:

n. case	Famiglie
0	5
1	15
2	15
3	25

Calcola:

1. Calcola la media
2. Calcola la mediana
3. Calcola la deviazione standard
4. Calcola il minimo
5. Calcola il massimo

3.1 Soluzioni

1. la media è pari a $\bar{y} = 2$
2. la mediana è pari a 2
3. la deviazione standard è pari a $s = 1.008$
4. il minimo è 0
5. il massimo è 3

4 Esercizio 4

Si considerino i voti conseguiti da 11 studenti delle scuole superiori al diploma:

$$82, 80, 84, 93, 63, 62, 91, 72, 89, 67, 96$$

Risolvi i punti seguenti:

1. quale tipologia di grafico è appropriata per rappresentare la distribuzione dei dati?
2. rappresenta graficamente la distribuzione.
3. calcola la mediana
4. calcola lo scarto interquartile
5. calcola l'estremo inferiore del range o campo di variazione
6. calcola l'estremo superiore del range o campo di variazione
7. costruire il boxplot

4.1 Soluzioni

I voti costituiscono una variabile quantitativa. La loro distribuzione può essere rappresentata mediante un **istogramma**, dopo aver raggruppato i valori in classi, oppure tramite un **diagramma ramo-foglia**.

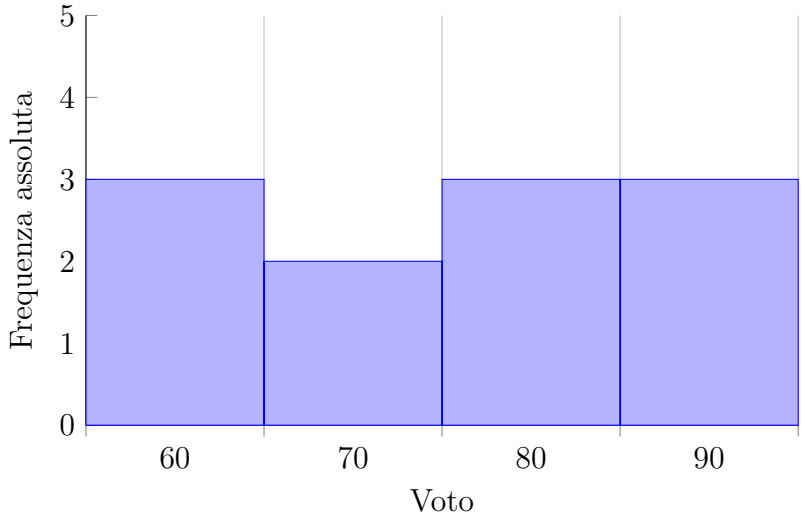

Figura 2: Distribuzione dei voti conseguiti al diploma.

Ramo	**Foglie**			
6	2	3	7	
7	2			
8	0	2	4	9
9	1	3	6	

Figura 3: Diagramma ramo-foglia dei voti conseguiti al diploma.

La mediana è pari a 82. Il primo e il terzo quartili sono pari, rispettivamente, a 67 e 91. Lo **scarto** o **range** interquartile è pari a $IQR = 24$. L'**estremo inferiore del campo di variazione** è pari al minimo (62), mentre l'**estremo superiore** è pari al massimo (96).

Prima di costruire il grafico, verifichiamo la presenza di eventuali valori anomali. I limiti inferiore e superiore sono calcolati mediante lo scarto interquartile:

$$L_{\text{inf}} = Q_1 - 1.5 \times IQR = 67 - 1.5 \times 24 = 67 - 36 = 31, \tag{1}$$

$$L_{\text{sup}} = Q_3 + 1.5 \times IQR = 91 + 1.5 \times 24 = 91 + 36 = 127. \tag{2}$$

Sono considerati potenzialmente anomali i valori inferiori a 31 o superiori a 127. Poiché tutti i voti osservati sono compresi tra 62 e 96, non sono presenti valori anomali.

Il box si estende dal primo quartile $Q_1 = 67$ al terzo quartile $Q_3 = 91$. La linea interna al box indica la mediana, pari a 82. In assenza di valori anomali, i baffi si estendono fino al valore minimo, pari a 62, e al valore massimo, pari a 96.

Il boxplot mostra che il 50% centrale dei voti è compreso tra 67 e 91 e che la mediana è pari a 82. Il voto minimo è 62, mentre il voto massimo è 96. Non sono presenti valori anomali.

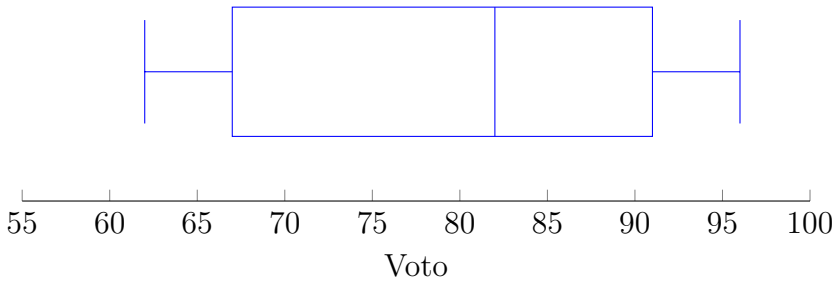

Figura 4: Boxplot dei voti conseguiti al diploma.