

# STATISTICA

*CdL in XXX - Prova del xx/xx/xxxx*

Cognome \_\_\_\_\_ Nome \_\_\_\_\_ Matr \_\_\_\_\_ Firma \_\_\_\_\_

## ESERCIZIO 1

Nella tabella riportata di seguito sono registrati per il 2007 i dati relativi ai primi 7 Paesi UE per numero di ingressi a fini di protezione concessi a rifugiati (in migliaia):

|   | Paese       | Ingressi |
|---|-------------|----------|
| 1 | Germania    | 622      |
| 2 | Lettonia    | 372      |
| 3 | Regno Unito | 310      |
| 4 | Francia     | 183      |
| 5 | Estonia     | 116      |
| 6 | Svezia      | 108      |
| 7 | Paesi Bassi | 96       |

(FONTE: UNHCR)

- 1) Definire il collettivo statistico e la variabile oggetto di studio, specificandone la natura
- 2) Calcolare il numero medio di ingressi e valutare opportunamente la variabilità del fenomeno
- 3) Rappresentare e studiare graficamente la concentrazione
- 4) Misurare il livello di concentrazione e commentare i risultati ottenuti

## ESERCIZIO 2

Il Ministero del Welfare ha pubblicato un bando da 12 milioni di Euro per finanziare progetti su tematiche inerenti l'immigrazione, l'educazione e l'inclusione sociale. Nella tabella di seguito i progetti finanziati sono classificati in base al finanziamento erogato (in migliaia di €):

| Finanziamento | Progetti Finanziati | Ammontare Stanziato |
|---------------|---------------------|---------------------|
| 0 -   100     | 22                  | 1.732.000           |
| 100 -   200   | 10                  | 1.978.000           |
| 200 -   400   | 12                  | 3.896.000           |
| 400 -   1000  | 5                   | 4.394.000           |

- 1) Mediamente qual è stato l'ammontare stanziato per progetto in ogni classe di finanziamento?
- 2) Calcolare per i progetti il finanziamento mediano
- 3) Studiare la variabilità calcolando e commentando il coefficiente di variazione
- 4) Studiare la concentrazione dei finanziamenti erogati e rappresentarla graficamente

# ESERCIZIO 1

|               | N° ingressi |
|---------------|-------------|
| Germania      | 622         |
| Lettonia      | 372         |
| Regno Unito   | 310         |
| Francia       | 183         |
| Estonia       | 116         |
| Svezia        | 108         |
| <b>TOTALE</b> | <b>1807</b> |



il collettivo statistico è dato dai 7 Paesi in tabella, il carattere studiato è invece il numero di ingressi a fini di protezione registrati, di natura quantitativa discreta

$$\bar{x} = 258.14$$

$$\sigma^2 = 31818.41$$

$$\sigma = 178.38$$



avremmo potuto usare anche lo scostamento semplice medio, meno opportuno ma ugualmente valido sarebbe stato lo scostamento semplice mediano per il calcolo del quale però era necessario calcolare la mediana

|               | N° ingressi |
|---------------|-------------|
| Paesi Bassi   | 96          |
| Svezia        | 108         |
| Estonia       | 116         |
| Francia       | 183         |
| Regno Unito   | 310         |
| Lettonia      | 372         |
| Germania      | 622         |
| <b>TOTALE</b> | <b>1807</b> |



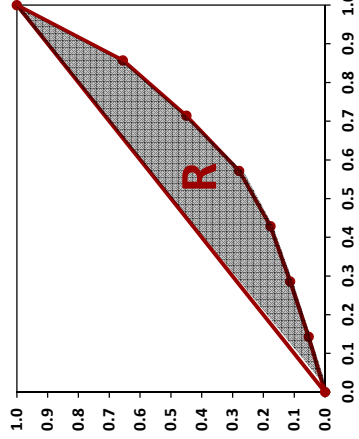
è una distribuzione unitaria: i dati vanno ordinati in senso crescente prima di iniziare lo studio della concentrazione

|   | $P_i$ | $A_i$ | $q_i$ | $P_i - P_{i-1}$ | $q_i + q_{i-1}$ | $(P_i - P_{i-1})(q_i + q_{i-1})$ |
|---|-------|-------|-------|-----------------|-----------------|----------------------------------|
| 1 | 0.143 | 96    | 0.053 | 0.143           | 0.053           | 0.008                            |
| 2 | 0.286 | 204   | 0.113 | 0.143           | 0.166           | 0.024                            |
| 3 | 0.429 | 320   | 0.177 | 0.143           | 0.290           | 0.041                            |
| 4 | 0.571 | 503   | 0.278 | 0.143           | 0.455           | 0.065                            |
| 5 | 0.714 | 813   | 0.450 | 0.143           | 0.728           | 0.104                            |
| 6 | 0.857 | 1185  | 0.656 | 0.143           | 1.106           | 0.158                            |
| 7 | 1.000 | 1807  | 1.000 | 0.143           | 1.656           | 0.237                            |



$$R = 0.364$$

metodo dei trapezi



con il metodo dei trapezi ricaviamo una misura della concentrazione a partire dalla sua rappresentazione, sottraendo dall'area del triangolo equilatero di lato 1 che rappresenta la massima concentrazione le aree del triangolino (in basso a sinistra) e dei 6 trapezi identificabili sotto la spezzata di concentrazione: in questo modo si ottiene l'area di concentrazione, equivalente alla misura data dall'indice R -> il livello della concentrazione osservato nel collettivo di Paesi UE è basso (36.4% della max conc. osservabile)

dalla rappresentazione grafica della concentrazione possiamo già osservare come l'area di concentrazione R sia piccola, e ciò suggerisce che la concentrazione dei rifugiati che hanno ottenuto la protezione umanitaria nei Paesi UE considerati è bassa

## ESERCIZIO 2

1

l'ammontare stanziato rappresenta, per ciascuna classe, l'intensità totale, ottenuta sommando per ogni progetto il finanziamento ottenuto. Per conoscere l'ammontare medio di ogni classe è sufficiente dividere quindi l'ammontare di ciascuna classe per la frequenza della stessa... se avessimo avuto soltanto i valori delle classi e la loro frequenza avremmo ottenuto l'ammontare moltiplicando il valore centrale per la frequenza, ipotizzando convenzionalmente lo stesso finanziamento per ogni progetto: qui è il procedimento contrario, ma con i veri dati osservati!

| Finanziamento | Progetti Finanziati | Ammontare Stanziato |
|---------------|---------------------|---------------------|
| 0 -   100     | 22                  | 1732000             |
| 100 -   200   | 10                  | 1978000             |
| 200 -   400   | 12                  | 3896000             |
| 400 -   1000  | 5                   | 4394000             |
|               | 49                  | 12000000            |

|         |    |             |
|---------|----|-------------|
| 1732000 | 22 | 78727.27 €  |
| 1978000 | 10 | 197800.00 € |
| 3896000 | 12 | 324666.67 € |
| 4394000 | 5  | 878800.00 € |

| X            | n <sub>i</sub> | f <sub>i</sub> | F <sub>i</sub> |
|--------------|----------------|----------------|----------------|
| 0 -   100    | 22             | 0.45           | 0.45           |
| 100 -   200  | 10             | 0.20           | 0.65           |
| 200 -   400  | 12             | 0.25           | 0.90           |
| 400 -   1000 | 5              | 0.10           | 1.00           |
|              | 49             | 1.00           |                |

100 - | 200 rappresenta la classe mediana poiché contiene il progetto che lascia a destra e sinistra, rispetto al finanziamento ottenuto, esattamente il 50% della distribuzione; per poter calcolare la mediana dobbiamo ricorrere a questo punto all'approssimazione lineare

$$Me = 125 \quad (\text{migliaia di euro})$$

poiché l'ammontare stanziato rappresenta l'intensità totale, per calcolare l'ammontare medio basta dividere per la numerosità del collettivo, ottenuta dal totale delle frequenze assolute; a questo punto si calcola la deviazione standard per poter ottenere il coefficiente di variazione della distribuzione dei finanziamenti... il problema è però che abbiamo l'ammontare totale ma non conosciamo i singoli finanziamenti erogati: per tale ragione, pur avendo la vera informazione sull'intensità è necessario utilizzare i val. centrali moltiplicati per le frequenze, in modo da poter calcolare non solo la media ma anche la varianza e la deviazione standard adottando la medesima convenzione (NB. la differenza tra media e mediana è imputabile alla presenza di val. anomali)

$$\bar{x} = 197.96 \quad (\text{migliaia di euro})$$

$$\sigma^2 = 38567.26 \quad (\text{migliaia di euro})$$

$$\sigma = 196.39 \quad (\text{migliaia di euro})$$

$$CV = 0.99$$

il coefficiente di variazione è pari al 99%

2

3

| Finanziamento | Progetti Finanziati | Ammontare Stanziato |
|---------------|---------------------|---------------------|
| 0 -   100     | 22                  | 1732000             |
| 100 -   200   | 10                  | 1978000             |
| 200 -   400   | 12                  | 3896000             |
| 400 -   1000  | 5                   | 4394000             |
|               | 49                  | 12000000            |

| p <sub>i</sub> = F <sub>i</sub> | A <sub>i</sub> | q <sub>i</sub> | p <sub>i</sub> - q <sub>i</sub> |
|---------------------------------|----------------|----------------|---------------------------------|
| 0.45                            | 1732000        | 0.14           | 0.31                            |
| 0.65                            | 3710000        | 0.31           | 0.34                            |
| 0.90                            | 7606000        | 0.63           | 0.27                            |
| 1.00                            | 12000000       | 1.00           | -                               |

abbiamo osservato il 46.5% della max concentrazione osservabile: il livello di concentrazione è medio-basso

$$R = 0.456$$

poiché abbiamo una distribuzione di frequenze (in questo caso in classi) le frazioni relative cumulate di unità statistiche possono essere ricavate a partire dalle frequenze relative cumulate, esprimendo queste lo stesso tipo di informazione; le frazioni relative cumulate di carattere posseduto dalle unità statistiche possono essere calcolate dall'ammontare stanziato, avendo a disposizione non solo quanti progetti appartengono a ciascuna classe ma anche l'ammontare di carattere posseduto: in mancanza di tale informazione avremmo dovuto ricorrere ai val. centrali, sotto l'ipotesi di equidistribuzione nelle diverse classi

