

LEZIONI DI STATISTICA SANITARIA

Dott. **SIMONE ACCORDINI**

Lezione n.4

- *Distribuzioni di frequenza
per una variabile quantitativa*



Sezione di Epidemiologia & Statistica Medica
Università degli Studi di Verona

**PESO, STATURA e SESSO delle MATRICOLE di MEDICINA
dell'UNIVERSITA' di VERONA nell'A.A. 95/96**

| PESO Kg | STAT. cm | SESSO | PESO Kg | STAT. cm | SESSO | PESO Kg | STAT. cm | SESSO |
|------------|-------------|-------|------------|-------------|-------|------------|-------------|-------|
| 56 | 159 | F | 77 | 192 | M | 51 | 171 | F |
| 66 | 169 | F | 60 | 173 | F | 48 | 156 | F |
| 50 | 160 | F | 78 | 182 | M | 55 | 167 | F |
| 53 | 170 | F | 52 | 167 | F | 60 | 177 | M |
| 54 | 168 | F | 47.5 | 164 | F | 58 | 170 | F |
| 53 | 161 | F | 64 | 166 | F | 67 | 167 | F |
| 63 | 172 | M | 52 | 160 | F | 50 | 172 | F |
| 53 | 170 | F | 72 | 184 | M | 58 | 169 | F |
| 62 | 161 | F | 48 | 169 | F | 77 | 179 | M |
| 56 | 163 | F | 66 | 170 | M | 52 | 162 | M |
| 50 | 160 | F | 55 | 172 | F | 49 | 160 | F |
| 52 | 170 | F | 67 | 177 | M | 49 | 165 | F |
| 58 | 173 | F | 66 | 170 | M | 62 | 178 | M |
| 52 | 167 | F | 50 | 160 | F | 68 | 174 | M |
| 73 | 178 | M | 51 | 167 | F | 75 | 181 | M |
| 57 | 166 | F | 95 | 195 | M | 48 | 167 | F |
| 52 | 165 | F | 58 | 160 | F | 53 | 160 | F |
| 56 | 171 | F | 67 | 178 | F | 49 | 167 | F |
| 67 | 175 | M | 67 | 175 | M | 52 | 165 | F |
| 63 | 182 | F | 60 | 160 | F | 55 | 155 | F |
| 55 | 169 | F | 56 | 165 | F | 84 | 188 | M |
| 58 | 165 | F | 50 | 165 | F | 56 | 170 | F |
| 55 | 175 | M | 52 | 170 | F | 60 | 171 | F |
| 66 | 176 | M | 58 | 172 | F | 52 | 176 | M |
| 55 | 164 | F | 60 | 170 | F | 62 | 180 | F |
| 47 | 160 | F | 54 | 166 | F | | | |
| 47 | 155 | F | 60 | 165 | F | | | |
| 63 | 169 | M | 74 | 172 | M | | | |
| 61 | 177 | F | 53 | 173 | F | | | |
| 53 | 170 | F | 72 | 183 | M | | | |
| 55 | 168 | M | 52 | 168 | F | | | |
| 53 | 162 | F | 51 | 164 | F | | | |
| 62 | 162 | F | 81 | 176 | M | | | |
| 45 | 160 | F | 50 | 160 | F | | | |
| 57 | 167 | F | 51 | 171 | F | | | |
| 45 | 158 | F | 64 | 180 | F | | | |
| 53 | 168 | F | 82 | 183 | M | | | |
| 50 | 160 | F | 47 | 156 | F | | | |
| 55 | 162 | F | 70 | 175 | M | | | |
| 70 | 177 | M | 58 | 168 | F | | | |
| 64 | 178 | F | 59 | 173 | F | | | |
| 52 | 164 | F | 68 | 165 | F | | | |
| 75 | 175 | M | 63 | 177 | F | | | |
| 75 | 178 | M | 50 | 159 | F | | | |
| 70 | 165 | F | 65 | 150 | F | | | |
| 58 | 167 | F | 60 | 170 | F | | | |
| 45 | 160 | F | 51 | 167 | F | | | |
| 50 | 167 | F | 75 | 182 | M | | | |
| 56 | 156 | F | 62 | 170 | M | | | |
| 59 | 165 | F | 85 | 174 | M | | | |

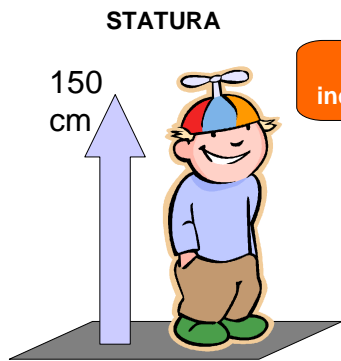
dati
individuali

seriazione
statistica delle
variabili peso,
statura e sesso
(n = 125)



DISTRIBUZIONE DI FREQUENZA DI UNA VARIABILE QUANTITATIVA

| valore x_i | relativa (%) | | rel. cumul. (%) $P_i(\%)$ |
|-----------------|-------------------|--------------|------------------------------|
| | assoluta n_i | $p_i(\%)$ | |
| 150 | 1 | .8 | .8 |
| 155 | 2 | 1.6 | 2.4 |
| 156 | 3 | 2.4 | 4.8 |
| 158 | 1 | .8 | 5.6 |
| 159 | 2 | 1.6 | 7.2 |
| 160 | 13 | 10.4 | 17.6 |
| 161 | 2 | 1.6 | 19.2 |
| 162 | 4 | 3.2 | 22.4 |
| 163 | 1 | .8 | 23.2 |
| 164 | 4 | 3.2 | 26.4 |
| 165 | 10 | 8.0 | 34.4 |
| 166 | 3 | 2.4 | 36.8 |
| 167 | 11 | 8.8 | 45.6 |
| 168 | 5 | 4.0 | 49.6 |
| 169 | 5 | 4.0 | 53.6 |
| 170 | 12 | 9.6 | 63.2 |
| 171 | 4 | 3.2 | 66.4 |
| 172 | 5 | 4.0 | 70.4 |
| 173 | 4 | 3.2 | 73.6 |
| 174 | 2 | 1.6 | 75.2 |
| 175 | 5 | 4.0 | 79.2 |
| 176 | 3 | 2.4 | 81.6 |
| 177 | 5 | 4.0 | 85.6 |
| 178 | 5 | 4.0 | 89.6 |
| 179 | 1 | .8 | 90.4 |
| 180 | 2 | 1.6 | 92.0 |
| 181 | 1 | .8 | 92.8 |
| 182 | 3 | 2.4 | 95.2 |
| 183 | 2 | 1.6 | 96.8 |
| 184 | 1 | .8 | 97.6 |
| 188 | 1 | .8 | 98.4 |
| 192 | 1 | .8 | 99.2 |
| 195 | 1 | .8 | 100.0 |
| TOTALE | 125 | 100.0 | |



Costruiamo gli intervalli di classe:

- Trovo il valore minimo e il valore massimo → min = 150 cm
max = 195 cm
- Calcolo il campo di variazione (range):
 $X_{\max} - X_{\min}$ → r = 45 cm
- Stabilisco il numero degli intervalli → K = 9
- Calcolo l'ampiezza degli intervalli:
 $\delta_i = \text{range} / K$ → $\delta_i = 45/9 = 5$ cm
- Costruisco gli intervalli di classe (esclusivi ed esaustivi)
- Conto il numero di individui (frequenza) per ogni classe

DISTRIBUZIONE DI FREQUENZA DI UNA VARIABILE QUANTITATIVA DISCRETIZZATA IN INTERVALLI DI CLASSE

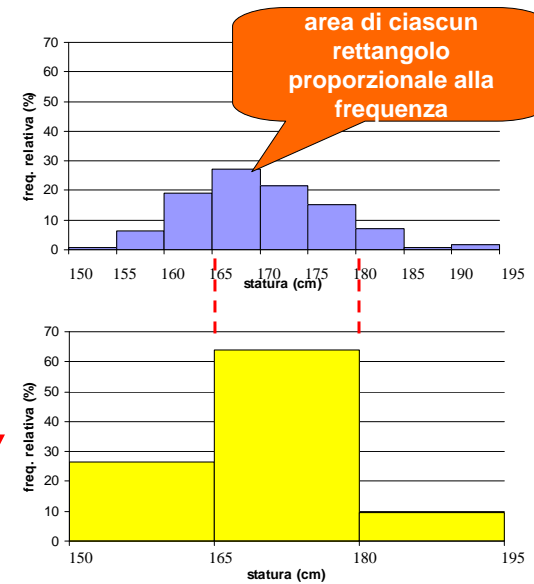
dati raggruppati in intervalli di classe

| CLASSE | n_i | | P_i (%) | |
|-----------|-----------|-----------------------------|--------------------|----------------|
| | FREQUENZA | | FREQUENZA CUMULATA | |
| | ASSOLUTA | RELATIVA % | ASSOLUTA | RELATIVA % |
| [150-155) | 1 | $(1/125) \cdot 100 = 0.8$ | 1 | 0.8 |
| [155-160) | 8 | $(8/125) \cdot 100 = 6.4$ | 1+8= 9 | 0.8+ 6.4= 7.2 |
| [160-165) | 24 | $(24/125) \cdot 100 = 19.2$ | 1+8+24=33 | 7.2+19.2=26.4 |
| [165-170) | 34 | $(34/125) \cdot 100 = 27.2$ | 1+8+24+34=67 | 26.4+27.2=53.6 |
| [170-175) | 27 | 21.6 | 94 | 75.2 |
| [175-180) | 19 | 15.2 | 113 | 90.4 |
| [180-185) | 9 | 7.2 | 122 | 97.6 |
| [185-190) | 1 | 0.8 | 123 | 98.4 |
| [190-195] | 2 | 1.6 | 125 | 100.0 |
| TOTALE | 125 | 100 | | |

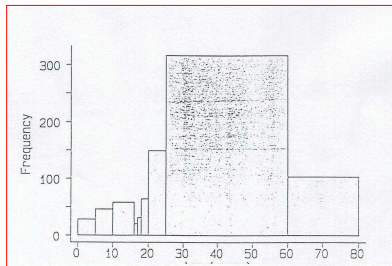


RAPPRESENTAZIONE GRAFICA DELLA DISTRIBUZIONE DI FREQUENZA DI UNA VARIABILE QUANTITATIVA: ISTOGRAMMA

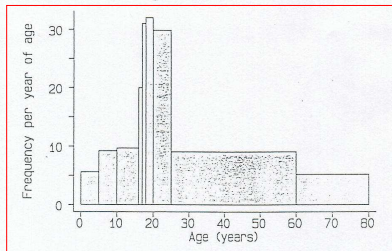
perdita di informazione al diminuire del numero di intervalli



Esempio: Vittime di incidenti stradali nel London Borough of Harrow nel 1985.



scorretto



corretto

| ETA' | FREQUENZA |
|---------------|------------|
| 0-4 | 28 |
| 5-9 | 46 |
| 10-15 | 58 |
| 16 | 20 |
| 17 | 31 |
| 18-19 | 64 |
| 20-24 | 149 |
| 25-59 | 316 |
| 60+ | 103 |
| TOTALE | 815 |

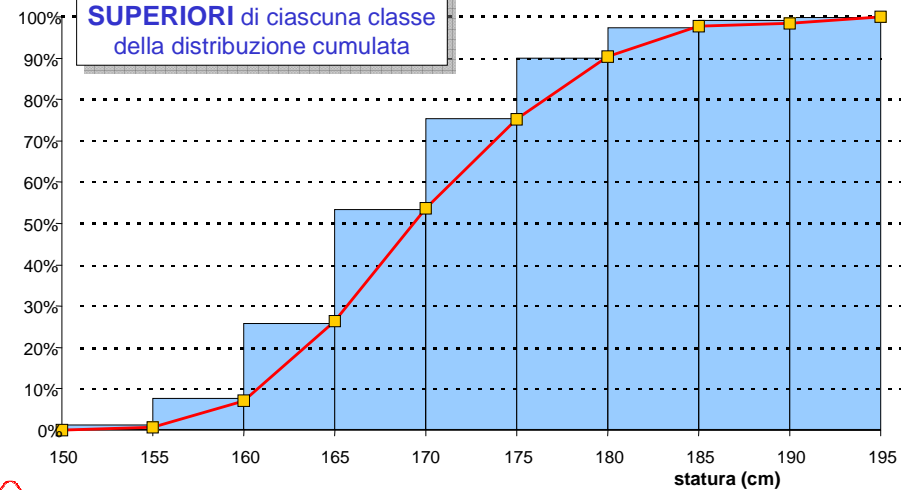
$\frac{316}{815} = 38.8\%$

$\frac{316}{35} = 9.0$
base altezza



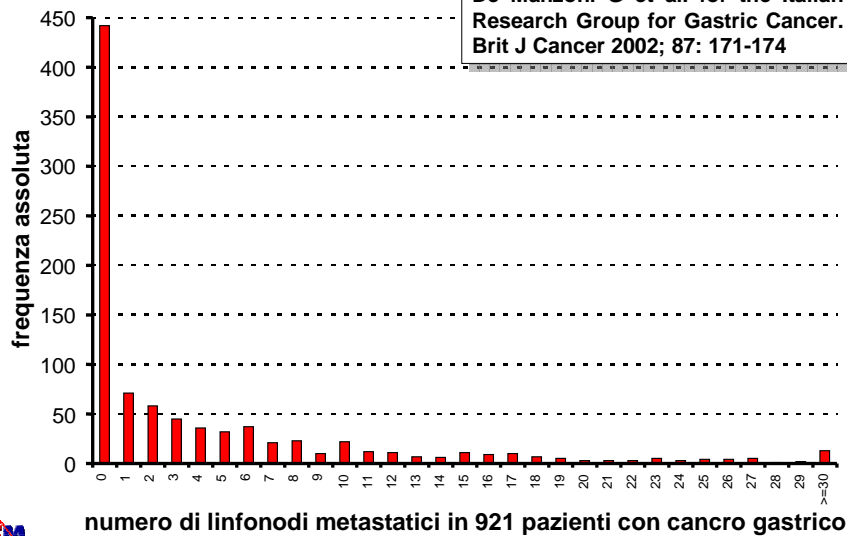
RAPPRESENTAZIONE GRAFICA DELLA DISTRIBUZIONE DI FREQUENZA CUMULATA DI UNA VARIABILE QUANTITATIVA: CURVA AD OGIVA

L'**OGIVA** si ottiene unendo i punti corrispondenti ai **LIMITI SUPERIORI** di ciascuna classe della distribuzione cumulata



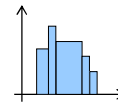
RAPPRESENTAZIONE GRAFICA DELLA DISTRIBUZIONE DI FREQUENZA DI UNA VARIABILE QUANTITATIVA DISCRETA: DIAGRAMMA A BARRE

De Manzoni G et al. for the Italian Research Group for Gastric Cancer. Brit J Cancer 2002; 87: 171-174



GRAFICI IN SINTESI

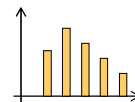
ISTOGRAMMA



variabile quantitativa continua

1. rettangoli adiacenti
2. le basi dei rettangoli possono essere diverse
3. frequenza sempre proporzionale all'**AREA** dei rettangoli (o all'altezza con basi uguali)

DIAGRAMMA A BARRE



variabile quantitativa discreta

variabile qualitativa ordinale

1. barre separate (per evidenziare la **non continuità dei valori** / la **distinzione tra le modalità**)
2. le basi delle barre sono tutte di uguale ampiezza
3. frequenza proporzionale alla **ALTEZZA** delle barre

