STATISTICA DESCRITTIVA

metodi di esposizione e sintesi dei dati

dati numerici

numero di valori non è troppo elevato

Sia v uno dei valori allora f si dice frequenza assoluta di v quel numero che indica quanti fra i dati sono uguali a v

es. 30 studenti

$v = 60$ kg

$f_1 = 5$ studenti che pesano 60 kg

$v = 55$ kg

$f_2 = 8$

campione fatto da n dati $(x_1, \dots, x_n)$

a ciascuno di essi associo una frequenza $(f_1, \dots, f_n)$

dati campione	frequenze assolute
x_1	f_1
x_2	f_2
⋮	⋮
x_n	f_n

- frequenza relativa $\frac{f}{n}$

$n \rightarrow$ ampiezza del campione (numerosità)

dati campione	frequenze relative
x_1	f_1/n
x_2	f_2/n
⋮	⋮
x_n	f_n/n

elisabetta.chiodaroli@unipi.it

Ricevimento: Di Cosmo, Elisabetta Chiodaroli
martedì pomeriggio 15-17

"Probabilità e statistica per l'ingegneria e la scienza" Sheldon, Ross

▲ → questo simbolo indica tutto ciò che il professore ha già chiesto all'orale

STATISTICA

interpretazione dati con ricorso di inferenza su un insieme di oggetti detto popolazione

(insieme vasto e completo)

es. materiali prodotti da un'azienda

popolazione italiana

consideriamo un sottogruppo della popolazione detto campione

es. 1000 materiali fra quelli prodotti è rappresentativo il campione della popolazione?

Si veglia il campione a caso

scelta a caso → campione è rappresentativo della popolazione

"statistica" deriva dalla parola stato

raccolta dati di interesse per lo stato

nata nel Rinascimento a Firenze

pesi	frequenze
50	3
55	8
60	5
70	4
58	10

ampiezza campione = 30



f. relative

3/10

4/15

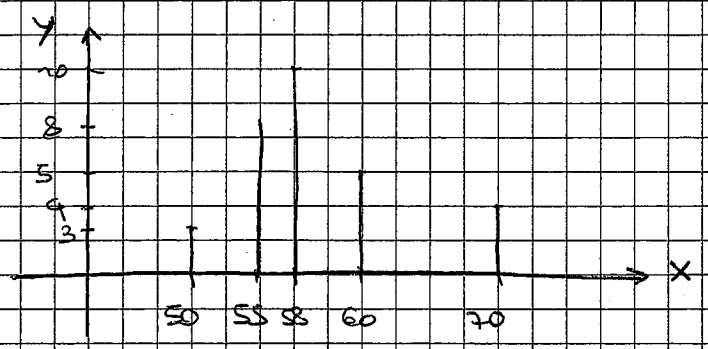
1/6

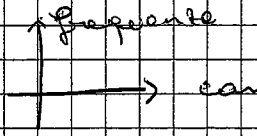
2/15

1/3

frequenza relativa = $\frac{\text{frequenza assoluta}}{\text{ampiezza campione}}$

- grafico a trattini







numero di valori NON è troppo numeroso (altrimenti non vale)

attrumenti

studio ore durate di 1000 notebooks

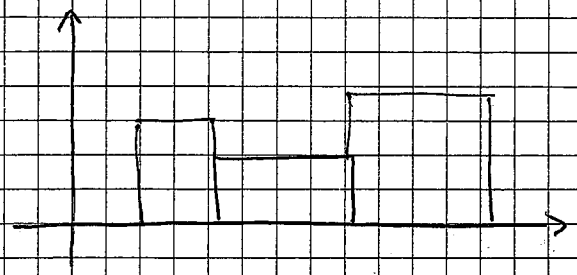
code d'ipotesi

allora suddivido i valori in classi

- [un anno, due anni) I classe
- [due anni, tre anni) II classe

- [tre anni, quattro anni)

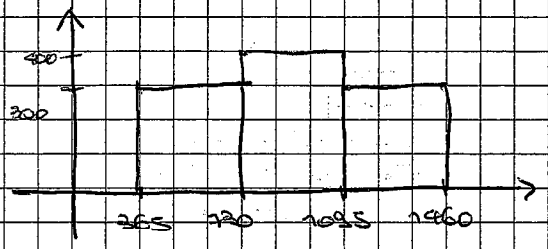
grafico a colonne → ISTOGRAMMA



[365, 730) → 300 notebooks

[730, 1095) → 400 notebooks

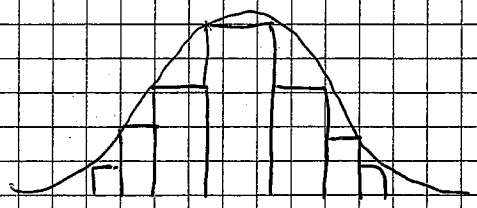
[1095, 1460) → 300 notebooks



master COPY

numero ottimale i 10 classi

- ISTOGRAMMA forma a CAMPANA



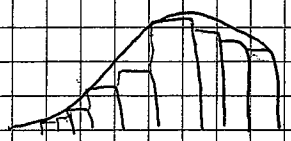
Allora e' istogramma si dice NORMALE

- forma ASIMMETRICA (coda più lunga a destra)

l'istogramma si dice RIGHT SKEWED



coda più lunga a sinistra
LEFT SKEWED



ISTOGRAMMA con 2 picchi si dice BIMODALE



master copy

• Grandezze statistiche
(calcolabili dai dati)

- MEDIA CAMPIONARIA

$(x_1, \dots, x_n)$ campione $= x$

* $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ media aritmetica

(valore medio dei dati raccolti)

- VARIANZA CAMPIONARIA VAR(x)

$(x_i - \bar{x})^2$ (distanza x_i - valor medio)^2

(scarto quadratico)

* $VAR(x) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ $VAR(x) \geq 0$

- VARIANZA EMPIRICA

* $VAR_e(x) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$