

# DESCRIZIONE DEI DATI - PARTE I

## La statistica

La **statistica** comprende un insieme di metodi per:

- la raccolta
- la descrizione
- l'analisi

di dati relativi a **fenomeni che hanno  
attitudine a variare**

2

Se si rilevano in un gruppo di individui



i valori di altezza o il sesso, ad esempio,  
1.67, 1.74, 1.94, 1.78 ....  
F, F, M, M ....

si può notare che

**i valori misurati variano da  
individuo a individuo**

3

## Perché i valori cambiano da individuo a individuo?

Tra le possibili **fonti di variabilità**, quelle più  
rilevanti sono la:

- ✓ **variabilità strumentale** (legata alla procedura di  
misurazione, agli strumenti e a chi misura);

(ad es. la procedura operativa o lo strumento non sono  
ancora a punto, colui che misura non è  
sufficientemente esperto)

- ✓ **variabilità biologica** (intrinseca).

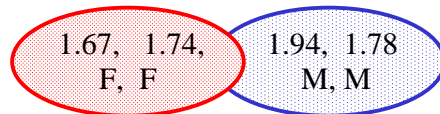
4

## Perché i valori cambiano da individuo a individuo?

*La variabilità strumentale può essere completamente controllata*, agendo sulle modalità di misurazione.

(ad es. ottimizzando la procedura operativa, tarando lo strumento, facendo training al personale)

*La variabilità biologica può essere solo parzialmente limitata*, rendendo più omogeneo l'insieme di soggetti analizzati.



5

## Esempio



Pressione arteriosa del bambino in età scolare

*La variabilità strumentale può essere completamente controllata addestrando il personale che effettua le misurazioni*

(ad es. scelta del bracciale)

*La variabilità biologica può essere parzialmente limitata identificando quei fattori che modificano la pressione arteriosa*

(ad es. età, classe ponderale, familiarità)

6

## Terminologia: universo

L'**universo** (o **popolazione**) consiste della **totalità degli elementi (unità statistiche)** che hanno certe caratteristiche



*Es.:*

☞ Studenti che seguono questa lezione di Statistica Medica

7

## Terminologia: campione

Un **campione** è un sottoinsieme di elementi dell' universo che viene utilizzato per trarre conclusioni sulle caratteristiche dell'universo

campione  universo

Il campione non deve essere selezionato ma deve essere scelto in modo casuale



*Es: Universo:* Studenti che seguono questa lezione di Statistica Medica  
*Campione:* 20 studenti presi a caso tra quelli che seguono questa lezione

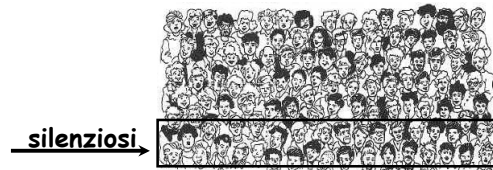
8

## Terminologia: campione

Un **campione** è un sottoinsieme di elementi dell' universo che viene utilizzato per trarre conclusioni sulle caratteristiche dell'universo

campione  $\Rightarrow$  universo

Il campione non deve essere selezionato ma deve essere scelto in modo casuale



*Es: Universo:* Studenti che seguono questa lezione di Statistica Medica  
*Campione:* 20 studenti presi a caso tra quelli che seguono questa lezione

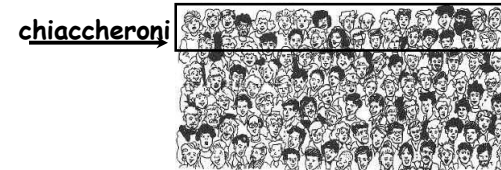
9

## Terminologia: campione

Un **campione** è un sottoinsieme di elementi dell' universo che viene utilizzato per trarre conclusioni sulle caratteristiche dell'universo

campione  $\Rightarrow$  universo

Il campione non deve essere selezionato ma deve essere scelto in modo casuale



*Es: Universo:* Studenti che seguono questa lezione di Statistica Medica  
*Campione:* 20 studenti presi a caso tra quelli che seguono questa lezione

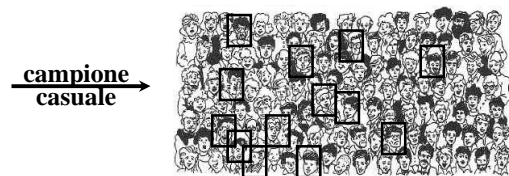
10

## Terminologia: campione

Un **campione** è un sottoinsieme di elementi dell' universo che viene utilizzato per trarre conclusioni sulle caratteristiche dell'universo

campione  $\Rightarrow$  universo

Il campione non deve essere selezionato ma deve essere scelto in modo casuale



*Es: Universo:* Studenti che seguono questa lezione di Statistica Medica  
*Campione:* 20 studenti presi a caso tra quelli che seguono questa lezione

11

## Il campione

Il campione casuale dovrebbe rappresentare una immagine in scala ridotta dell'universo.



campione come  
miniatura  
dell'universo



... ovvero dovrebbe essere **rappresentativo** dell'universo.

Questa è la condizione (non verificabile) di validità del processo di **generalizzazione dei risultati**.

## Il campione



N=100



1) n=5



2) n=25



3) n=75

Quale tra questi tre campioni contiene più informazioni sulla popolazione?

## Terminologia: variabili

Si dice **variabile** una **caratteristica** delle unità statistiche che può assumere una pluralità di valori al variare dell'unità su cui è rilevata

*Es:* Altezza, Sesso, Titolo di Studio, peso alla nascita

Le variabili possono essere:

i) **quantitative**

ii) **qualitative**

vengono indicate con lettere maiuscole scelte, in genere, tra le ultime lettere dell'alfabeto: Y, X, Z

*Es:* Y = Altezza    X = Sesso

14

## Terminologia: dati

I **dati** sono quei **valori numerici** o quelle **modalità**, assumibili da una variabile.

I dati sono rappresentati da lettere minuscole con un indice che distingue le diverse unità fra loro:

*Es:* Y=Altezza     $y_1= 1.67$      $y_2= 1.74$      $y_3= 1.94$      $y_4= 1.78$

X=Sesso     $x_1= F$      $x_2= F$      $x_3=M$      $x_4=M$

15

## Variabili quantitative discrete

Una variabile quantitativa è **discreta** se può assumere come valore un **qualsiasi numero naturale**

*Es.:* - Numero automobili per famiglia  
- Voto esame di statistica  
- Durata dell'allattamento (in mesi)

Le variabili quantitative discrete derivano usualmente da **conteggi**

16

## Variabili quantitative continue

Una variabile quantitativa è **continua**, se può assumere come valore un **qualsiasi numero reale**

*Es.* : Altezza, Peso, Concentrazione di glucosio nel sangue

I valori assunti da una variabile continua dipendono in realtà dal potere di risoluzione dello strumento di misura

*Es.* : Una altezza di 1.78324321.... m, potrebbe essere riportata al cm (1.78) o al mm (1.783) a seconda dell'uso

Le variabili quantitative continue derivano usualmente da **misurazioni**

17

## Variabili qualitative nominali

Una variabile qualitativa è **nominale**, quando ogni possibile ordinamento delle modalità è arbitrario

*Es:* Sesso, Colore degli occhi, tipologia di parto

Etnia pazienti coinvolti in una sperimentazione clinica  
**caucasico - afroamericano - africano - indiano .. etc.**  
=  
**afroamericano - indiano - caucasico - africano .. etc.**

18

## Variabili qualitative ordinali

Una variabile qualitativa è **ordinale**, quando è possibile individuare un ordinamento naturale delle modalità.

*Es.* : Segno zodiacale, Titolo di studio

Misurazione dell'intensità del dolore

**nulla < lieve < moderata < forte**  
**forte > moderata > lieve > nulla**  
**moderata - forte - nulla - lieve**

**NO!!**

19

## Variabili qualitative ordinali

*Es:* Misurazione dell'intensità del dolore

**nulla < lieve < moderata < forte**

Alle modalità si può associare un **codice numerico**:  
(*Es.*: nulla=0, lieve=1, moderata=2, forte=3)  
**che però non ha significato quantitativo:**

- ▶ 2 (dolore moderato) **non** è il doppio di 1 (dolore lieve),  
3 (dolore forte) **non** è il triplo di 1
- ▶ la differenza tra 2 e 1 **non** è uguale a quella tra 3 e 2

20

## Variabili qualitative a due livelli

Vengono chiamate anche dicotomiche (o binarie), segnalano la presenza (o l'assenza) di una caratteristica.

**Es:** Presenza di gravi complicazioni dopo un intervento chirurgico. Le uniche modalita' che questa variabile puo' assumere sono 'SI' , 'NO'

Dalla frequenza di 'SI' si ottiene la frequenza di 'NO' calcolando  $1 - \text{la frequenza di 'SI'}$

**Es:** 0.12 (12%) di 'SI' implica 0.88 (88%) di 'NO'

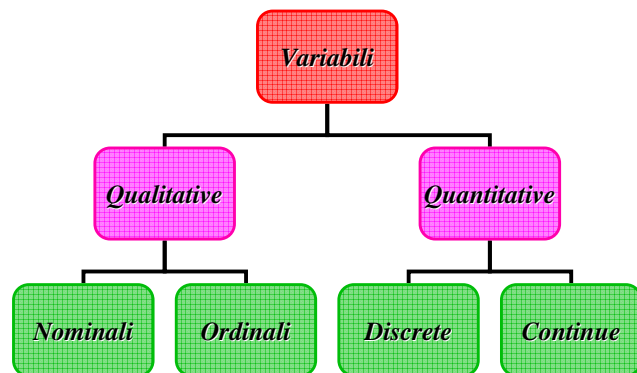
La frequenza di 'SI' si chiama PROPORZIONE

21

Cosa possiamo concludere dai dati relativi al campione di 20 studenti?

Soggetto	Altezza	Sesso	Soggetto	Altezza	Sesso
1	1.76	M	11	1.77	F
2	1.71	F	12	1.69	F
3	1.54	F	13	1.93	M
4	1.82	M	14	1.67	F
5	1.59	F	15	1.72	M
6	1.74	M	16	1.59	F
7	1.95	M	17	1.60	F
8	1.68	M	18	1.81	F
9	1.85	M	19	1.73	F
10	1.74	F	20	1.78	M

## Per riassumere



23

## Esempi

Numero di carie presenti nell'arcata superiore	Quantitativa Discreta
Stato civile	Qualitativa Nominale
Consumo giornaliero di caffeina (mg)	Quantitativa Continua
Consumo giornaliero di caffè della macchinetta (bicchierini)	Quantitativa Discreta
Albumina sierica (g/l)	Quantitativa Continua
Tipologia Epatite	Qualitativa Nominale
Numero di linfonodi metastatici riscontrati alla TAC	Quantitativa Discreta

24

## Distribuzioni di frequenza

Per riassumere i dati si costruiscono le **distribuzioni di frequenza**

possibili valori (modalità) che una variabile può assumere

e

frequenze con cui questi valori si manifestano

25

## Distribuzioni di frequenza

I dati di un'unità per la donazione di sangue mostrano che il numero totale di donatori rispetto ai quattro gruppi sanguigni ammonta a: A 725; B 258; AB 72; e O 1073.

Gruppo sanguigno	f
A	725
B	258
AB	72
O	1073
Totale	n=2128

**f = frequenza assoluta**

numero di volte in cui una certa modalità si manifesta nel campione

258 dei 2128 donatori hanno gruppo sanguigno B

26

## Distribuzioni di frequenza

**p = frequenza relativa**

rapporto tra la frequenza assoluta con cui si manifesta una modalità e la numerosità totale del campione

Gruppo sanguigno	f	f/n	p	p%
A	725	725/2128	0.341	34.1
B	258	258/2128	0.121	12.1
AB	72	72/2128	0.034	3.4
O	1073	1073/2128	0.504	50.4
Totale	n=2128		1.000	100

Il 12% dei donatori ha gruppo sanguigno B

27

## Frequenze assolute e relative

- **frequenze assolute f**
  - ✓ possono assumere valori compresi tra 0 e n (dimensione del campione)
  - ✓ la loro somma è pari a n
- **frequenze relative p**
  - ✓ possono assumere valori compresi tra 0 e 1
  - ✓ la loro somma è pari a 1
- **frequenze relative p%**
  - ✓ possono assumere valori compresi tra 0% e 100%
  - ✓ la loro somma è pari a 100%

28

## Frequenze assolute e relative

Frequenze assolute e relative forniscono le stesse informazioni sulla distribuzione

Tuttavia, le frequenze relative:

- ✓ facilitano la percezione del peso delle modalità;
- ✓ consentono di confrontare la distribuzione di una variabile in campioni di diversa numerosità.

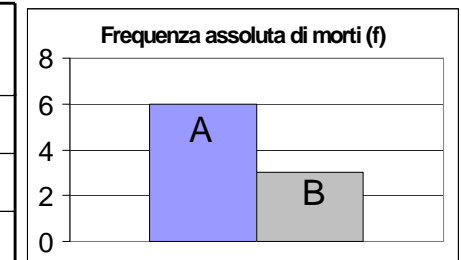
Andrebbero sempre accompagnate dalla numerosità su cui sono state calcolate!

29

## Esempio

Si vuole valutare l'efficacia di un nuovo farmaco (A) sulla mortalità post-infarto (1 mese). Nello studio vengono coinvolti 150 pazienti: 100 sono randomizzati a ricevere il farmaco sperimentale, 50 il trattamento standard (B).

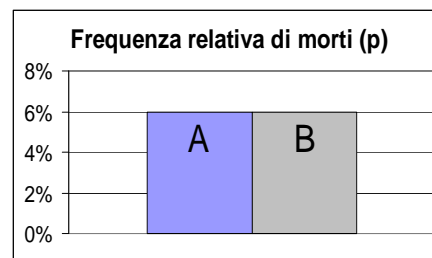
	Trattati con	
	A	B
<b>Morti</b>	6	3
<b>Vivi</b>	94	47
<b>Totale</b>	100	50



## Esempio

Si vuole valutare l'efficacia di un nuovo farmaco (A) sulla mortalità post-infarto (1 mese). Nello studio vengono coinvolti 150 pazienti: 100 sono randomizzati a ricevere il farmaco sperimentale, 50 il trattamento standard (B).

	Trattati con	
	A	B
<b>Morti</b>	6(6%)	3(6%)
<b>Vivi</b>	94(94%)	47(94%)
<b>Totale</b>	100	50



## Attenzione alle informazioni fuorvianti!

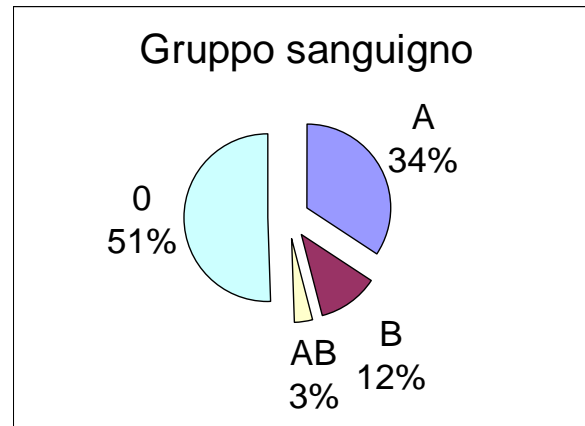
"The antibiotic phosphomycin is advertised as being 100% effective in chronic urinary tract infections."

*L'antibiotico fosfomicin è efficace al 100% nelle infezioni urinarie croniche.*

Lo studio su cui si basa questa informazione ha coinvolto 8 pazienti, dopo aver eliminato i pazienti le cui urine contenevano batteri fosfomicina-resistenti.



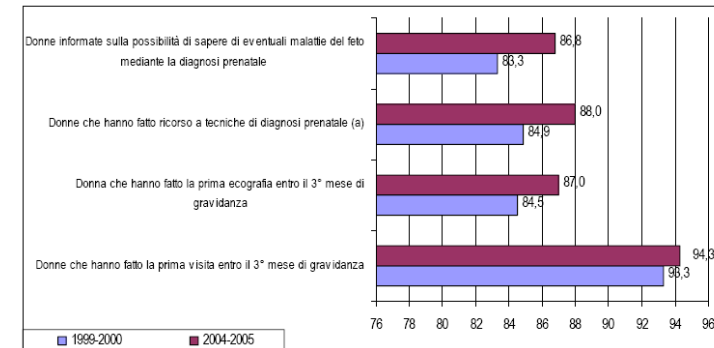
## Grafici per var. qualitative



**Diagramma areolare (o a torta)**

## Esempio: Assistenza in gravidanza

Gráfico 1 Principali indicatori di assistenza in gravidanza. Confronto 2004-2005 (dati provvisori) con 1999-2000 (per 100 donne con le stesse caratteristiche)



(a) Le tecniche di diagnosi prenatale rilevate sono dosaggio alfa fetoproteina, prelievo villi coriali, amniocentesi, ecografia morfologica fetale, tri-test.

Istituto nazionale di statistica  
Gravidanza, parto, allattamento al seno  
2004 - 2005

**Diagramma a barre orizzontali**

34

## Variabili quantitative discrete

Successione delle **frequenze** che corrispondono ai **valori** assunti da una **variabile quantitativa discreta**.

*Numero di morti causate da incidenti stradali rilevate da 14 reparti di emergenza in una regione durante un week-end.*

X	frequenze semplici		frequenze cumulate	
	assolute f	relative p	assolute F	relative P
0	7	0.500	7	0.500
1	3	0.214	10	0.714
2	2	0.143	12	0.857
3	1	0.071	13	0.929
4	1	0.071	14	1.000

## Frequenze cumulate

X	frequenze semplici		frequenze cumulate	
	assolute f	relative p	assolute F	relative P
0	7	0.500	7	0.500
1	3	0.214	7+3=10	0.714
2	2	0.143	7+3+2=12	0.857
3	1	0.071	7+3+2+1=13	0.929
4	1	0.071	7+3+2+1+1=14	1.000

In 12 dei 14 reparti di emergenza (pari al 86% del totale) sono state riscontrate 2 o meno morti causate da incidenti stradali

$$0.875 = 0.5 + 0.214 + 0.143 = 12/14$$

36

## Frequenze cumulate assolute e relative

- frequenze cumulate assolute F
  - ✓ La prima frequenza cumulata è pari alla prima frequenza assoluta.
  - ✓ L'ultima frequenza cumulata è pari alla numerosità campionaria.
- frequenze cumulate relative P
  - ✓ La prima frequenza cumulata relativa è pari alla prima frequenza relativa.
  - ✓ L'ultima frequenza cumulata relativa è pari ad uno.

37

## Grafici - Var. quantitative discrete

Morti causate da incidenti stradali

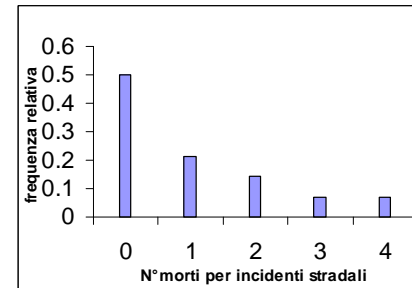


Diagramma ad aghi (frequenze relative)

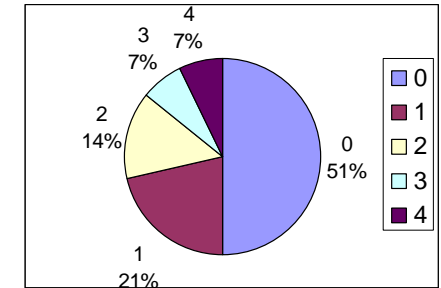


Diagramma a torta (frequenze relative)

38

## Distribuzioni di frequenza : il caso di variabili continue

In un'indagine condotta da un gruppo di neonatologi si sono rilevati i valori della lunghezza supina (cm) in un campione di 60 neonati. Le misurazioni, eseguite con l'infantometro Harpenden, sono riportate di seguito.

51.0	46.5	48.7	54.5	46.0	51.2	55.0	50.2	44.5	56.3
49.4	47.8	50.0	48.2	52.2	51.1	50.2	53.4	49.2	46.5
49.0	49.7	52.9	48.9	47.0	54.7	50.3	47.4	50.5	51.5
52.5	44.4	50.8	51.2	50.8	52.3	47.7	50.5	49.5	50.9
51.5	49.8	46.2	49.5	50.0	48.2	48.5	51.7	52.9	51.6
51.8	53.0	48.9	54.0	52.5	50.8	53.8	49.5	50.5	52.7

39

## Possiamo migliorare un po' la situazione ...

44.4	48.2	49.5	50.5	51.5	52.9
44.5	48.2	49.5	50.5	51.5	52.9
46.0	48.5	49.7	50.8	51.6	53.0
46.2	48.7	49.8	50.8	51.7	53.4
46.5	48.9	50.0	50.8	51.8	53.8
46.5	48.9	50.0	50.9	52.2	54.0
47.0	49.0	50.2	51.0	52.3	54.5
47.4	49.2	50.2	51.1	52.5	54.7
47.7	49.4	50.3	51.2	52.5	55.0
47.8	49.5	50.5	51.2	52.7	56.3

## Distribuzioni di frequenza : il caso di variabili continue

La **distribuzione di frequenza** di una **variabile continua** si rappresenta in modo analogo a quella degli altri tipi di variabili, ma....

in questo caso, la frequenza non è riferita ad un singolo valore, ma ad **intervalli (o classi)** di valori.

41

## Distribuzioni di frequenza : il caso di variabili continue

*Lunghezza supina (cm) in un campione di 60 neonati.*

Estremi di classe	Valore centrale	Freq. semplici		Freq.cumulate	
		f	p%	F	P%
44.25 - 45.75	45.0				
45.75 - 47.25	46.5				
47.25 - 48.75	48.0				
48.75 - 50.25	49.5				
50.25 - 51.75	51.0				
51.75 - 53.25	52.5				
53.25 - 54.75	54.0				
54.75 - 56.25	55.5				
56.25 - 57.75	57.0				

9 classi di uguale ampiezza (1.50cm)

42

## Distribuzioni di frequenza : il caso di variabili continue

*Lunghezza supina (cm) in un campione di 60 neonati.*

Estremi di classe	Valore centrale	Freq. semplici		Freq.cumulate	
		f	p%	F	P%
44.25 - 45.75	45.0	2	3.3	2	3.3
45.75 - 47.25	46.5	5	8.3	7	11.7
47.25 - 48.75	48.0	7	11.7	14	23.3
48.75 - 50.25	49.5	14	23.3	28	46.7
50.25 - 51.75	51.0	16	26.7	44	73.3
51.75 - 53.25	52.5	9	15.0	53	88.3
53.25 - 54.75	54.0	5	8.3	58	96.7
54.75 - 56.25	55.5	1	1.7	59	98.3
56.25 - 57.75	57.0	1	1.7	60	100.0

5 dei 60 neonati hanno una lunghezza supina compresa fra 45.75 e 47.25

43

## Gli estremi di classe

**[44.25-45.75)** o **44.25 - 45.75**  
 classe chiusa a sinistra e aperta a destra  
 estremo sn incluso

**(44.25-45.75]** o **44.25 + 45.75**  
 classe chiusa a destra e aperta a sinistra  
 estremo dx incluso

**[44.25-45.75]** o **44.25 - 45.75**  
 classe chiusa a sinistra e a destra  
 estremo sn e dx inclusi

**(44.25-45.75)** o **44.25 - 45.75**  
 classe aperta a sinistra e a destra  
 estremo sn e dx esclusi

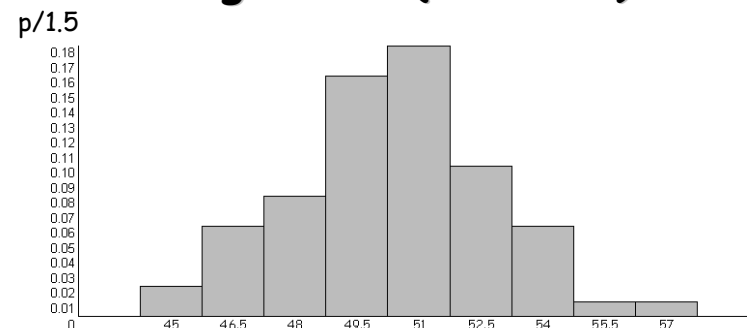
44

## Le classi

- ✓ La scelta del **numero** di classi e degli **estremi** è arbitraria. Entrambi vengono determinati in base a criteri di convenienza.
- ✓ Il **numero** di classi può oscillare e dipende dalla numerosità dei dati.
- ✓ Scegliere **estremi** che siano clinicamente/biologicamente **significativi** o naturali e, preferibilmente, di **uguale ampiezza**.  
NO: 44.137 - 45.541      SI: 44.00 - 45.50
- ✓ Le classi debbono essere mutuamente esclusive (fate attenzione agli estremi!!).

45

## Istogramma (corretto)



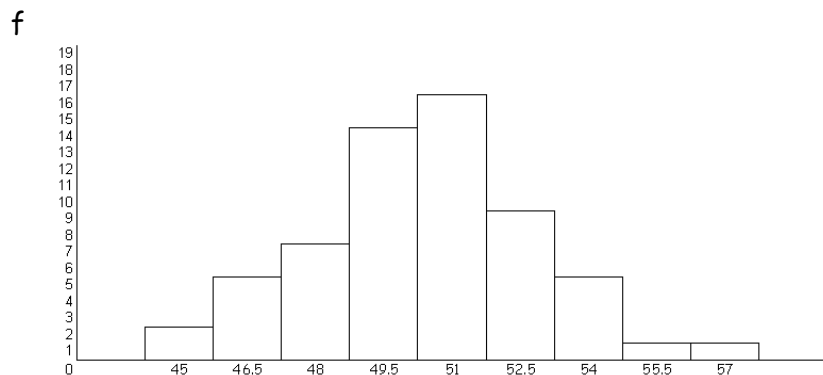
Ciascun rettangolo ha :

- per base l'ampiezza della classe
- per altezza la frequenza relativa della classe diviso l'ampiezza (densità di frequenza)
- un'area pari alla frequenza relativa

Globalmente i rettangoli ricoprono un'area unitaria

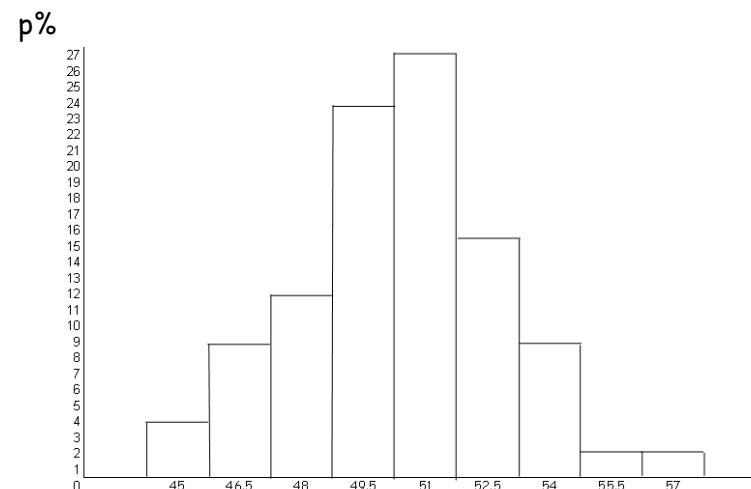
46

## Diagramma a barre (erroneamente chiamato istogramma)



47

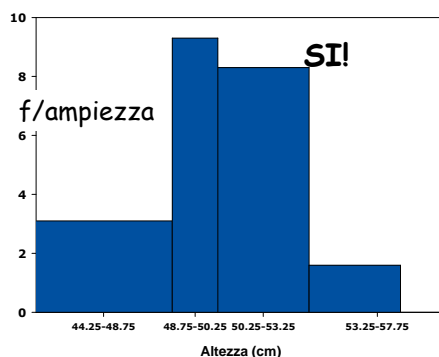
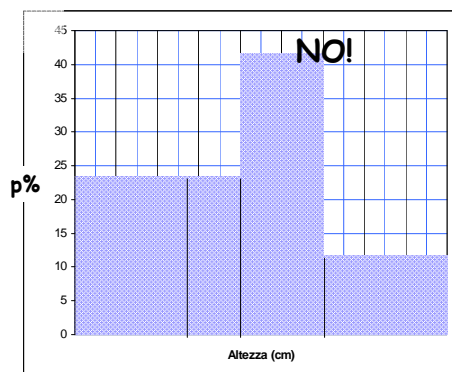
## Diagramma a barre (erroneamente chiamato istogramma)



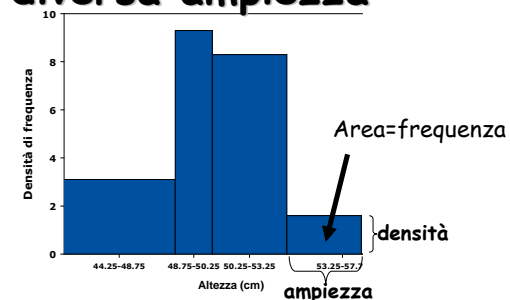
48

## Classi di diversa ampiezza

Estremi di classe	Ampiezza di classe	freq. semplici		Densità freq.	
		f	p%	f/amp	p%/amp
(44.25, 48.75]	4.5	14	23.3	3.1	5.2
(48.75, 50.25]	1.5	14	23.3	9.3	15.5
(50.25, 53.25]	3	25	41.7	8.3	13.9
(53.25, 57.75]	4.5	7	11.7	1.6	2.6

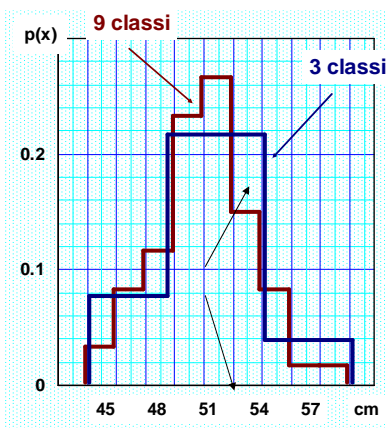


## Classi di diversa ampiezza

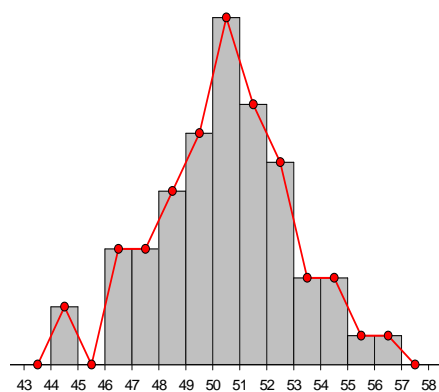


- ✓ Ogni istogramma (rettangolo) rappresenta una classe:  
**base** = ampiezza della classe  
**altezza** = densità di frequenza
- ✓ L'area di ogni rettangolo è pari alla frequenza assoluta (o relativa) della classe su cui insiste.
- ✓ L'area totale deve essere pari a  $n$  o  $1$ , a seconda del tipo di frequenze raffigurate.

## Ampiezza delle classi

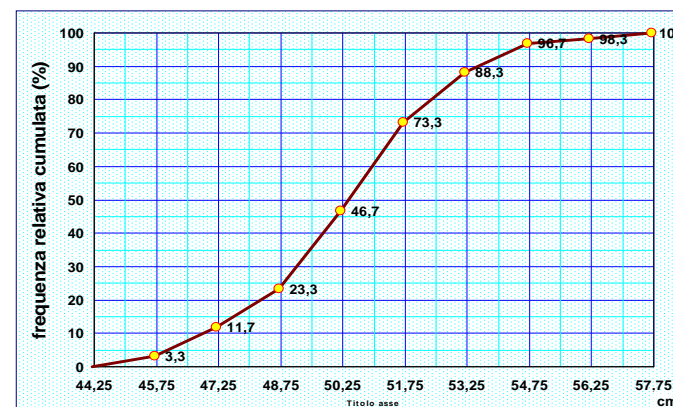


Al diminuire del numero di classi si perdono i dettagli sulla distribuzione.



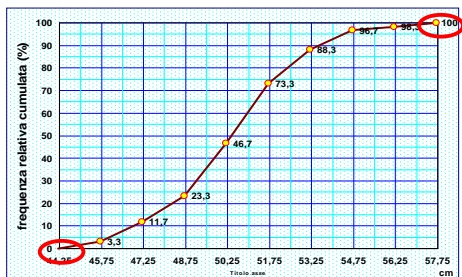
All'aumentare del numero di classi si guadagnano dettagli sulla distribuzione (ma sino ad un certo punto!!)

## Grafico delle frequenze cumulate



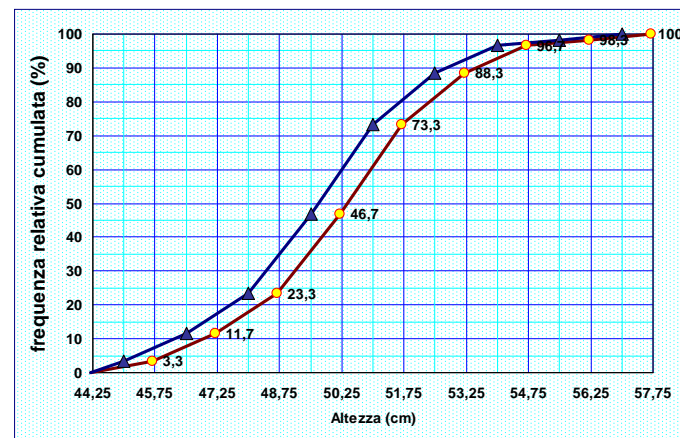
**Ogiva di Galton**

## Grafici per var. continue



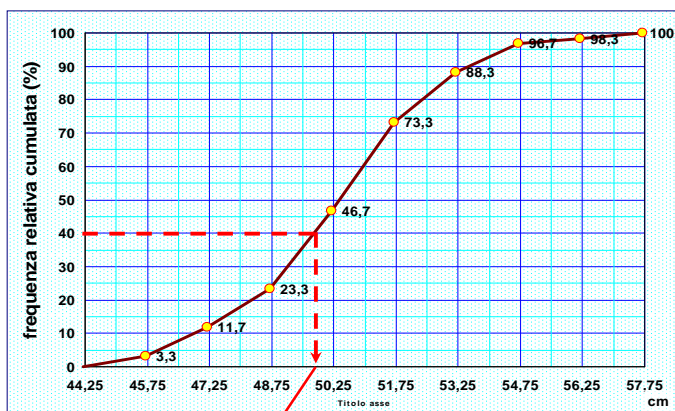
- ✓ La spezzata parte da 0 e termina a 1 o 100%.
- ✓ La spezzata si ottiene congiungendo con dei segmenti i due punti che hanno per coordinate:  
[estr inf, freq cum prec] ● — ● [estr sup, freq cum]
- ✓ Si assume che la distribuzione dei dati nelle classi sia uniforme (interpolazione lineare)

## Grafici per var. continue



Se si congiungessero i valori centrali si otterrebbe una rappresentazione scorretta.

## Grafici per var. continue



Qual è il valore di altezza sotto il quale trovo il 40% dei neonati?

~ 49,75 cm

## Descrizione di una variabile in più popolazioni

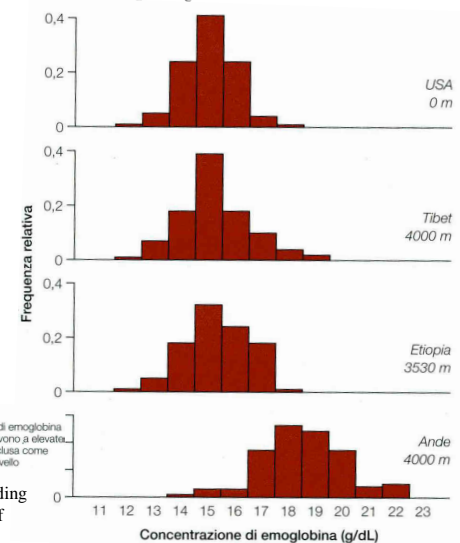
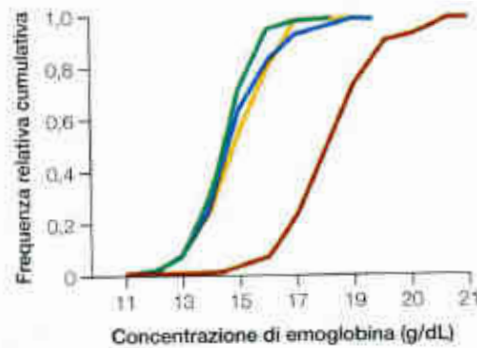


Figura 2.4-1  
Istogrammi che mostrano la concentrazione di emoglobina in maschi di popolazioni umane viventi in altitudini elevate, altitudini in tre differenti parti del mondo. È inclusa come controllo una quarta popolazione che vive a livello del mare (USA).

da Beal et al. 2002. Proceeding of the National Academy of Science 99:17215-17218.

## Descrizione di una variabile in più popolazioni



■ USA 0 m      ■ Tibet 4000 m  
■ Etiopia 3530 m      ■ Ande 4000 m

**Figura 2.4-2**  
 Distribuzioni di frequenza cumulative della concentrazione di emoglobina in maschi umani che vivono ad altitudini elevate in Etiopia, in Tibet e sulle Ande. È inclusa come controllo una quarta popolazione che vive a livello del mare negli Stati Uniti. (Da Beall et al., 2002; ridisegnato.)

57

## Descrizione di una variabile in più popolazioni

### Distribuzione di frequenza a doppia entrata

	Livello ematico di emoglobina (Hb, g/dl)					Totale
	12 (11.5,12.5]	13 (12.5,13.5]	14 (13.5,14.5]	15 (14.5,15.5]	16 (15.5,16.5]	
donne	18	65	14	2	1	<b>100</b>
uomini	2	40	71	58	29	<b>200</b>
<b>Totale</b>	<b>20</b>	<b>105</b>	<b>85</b>	<b>60</b>	<b>30</b>	<b>300</b>

Quale proporzione di soggetti ha livello di Hb > di 14.5 g/dl ?

Quale proporzione di donne ha livello di Hb > di 14.5 g/dl ?

58

## Definizione di Percentile

Il **percentile**  $x_p$  ( $0 \leq p \leq 1$ ) della distribuzione di una variabile continua è quel valore della variabile che soddisfa queste condizioni

- il  $p\%$  delle osservazioni assume valori  $\leq$  di  $x_p$ ,
- l'  $(1-p)\%$  delle osservazioni assume valori  $>$  di  $x_p$

I percentili sono utili per:

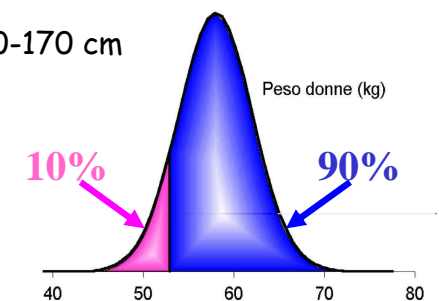
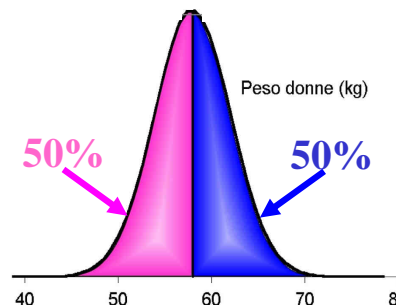
- Descrivere una distribuzione
- Identificare range di normalità
- Classificare il valore di un soggetto rispetto alla distribuzione del fenomeno

59

## Percentili da un istogramma

Peso delle donne di altezza 160-170 cm

$$p = 0.10 \quad x_{0.10} = 53$$



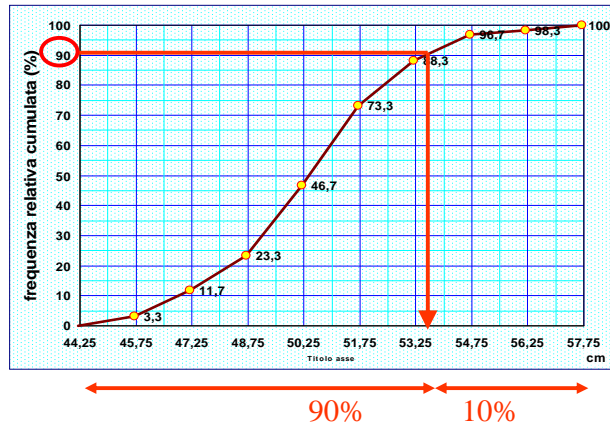
$$p = 0.50 \quad x_{0.50} = 58$$

60

# Percentili a partire dalle frequenze relative cumulate

Lunghezza dei bambini

Es.  $p=0.90$



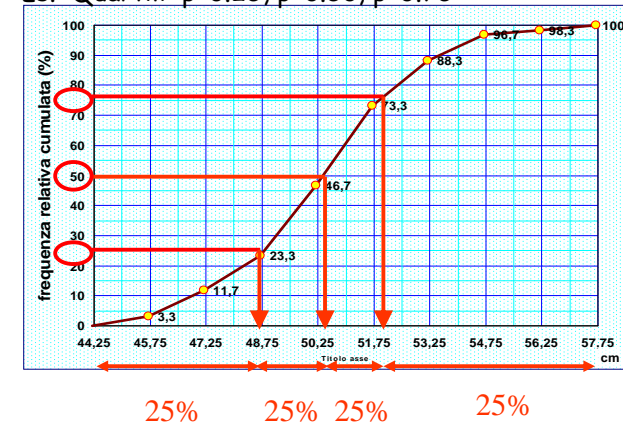
61

# Percentili particolari: quartili

Quartili: suddividono i dati in quattro parti uguali (25%)

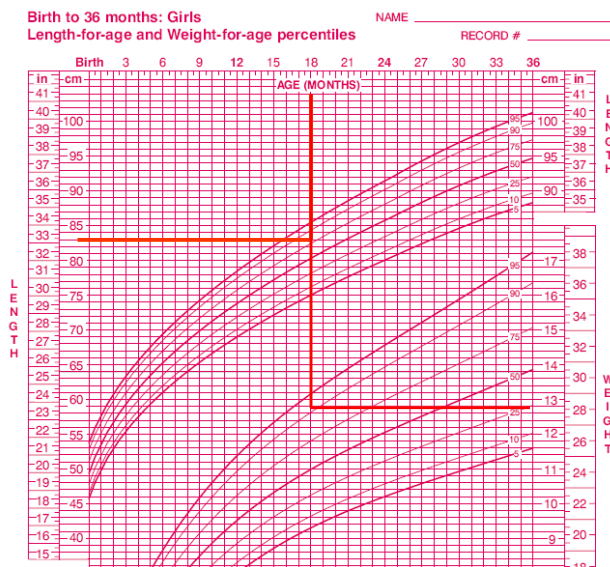
Lunghezza dei bambini

Es. Quartili:  $p=0.25$ ,  $p=0.50$ ,  $p=0.75$



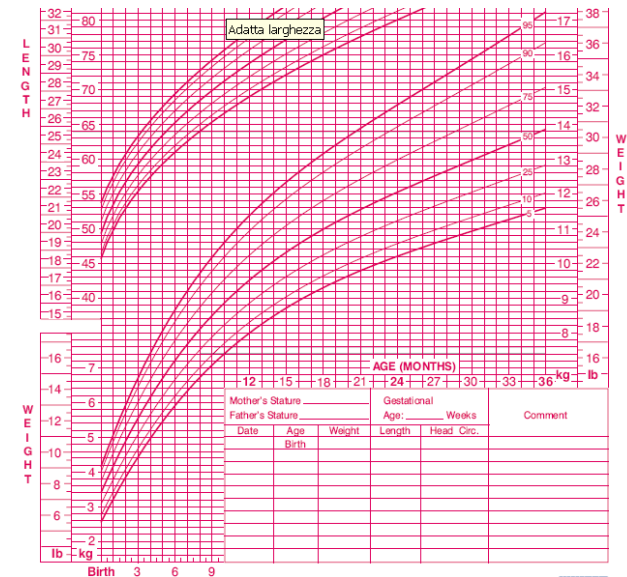
62

# Curve Percentile - lunghezza e peso



63

# Curve Percentile - peso neonate



64



## Esercizio per lo studente

Glicemia (mg/dl) in 500 soggetti anziani  
Raggruppamento in 5 classi di uguale ampiezza

<i>Estremi di classe</i>	<i>valore centrale</i>	<i>freq. semplici</i>		<i>freq. cumulate</i>	
		f	p%	F	P%
65-75	<b>70</b>	75	15	75	15
75-85	<b>80</b>	100	20	175	35
85-95	<b>90</b>	150	30	225	65
95-105	<b>100</b>	125	25	450	90
105-115	<b>110</b>	50	10	500	100

- Rappresentare graficamente il fenomeno mediante un istogramma
- Accorpare le ultime due classi e costruire il relativo istogramma