



**Corso di**

**Statistica**

**Prof. Andrea CICCARELLI**

# Informazioni sul corso - 1

## Obiettivi

**Nell'ambito delle lezioni impartite verranno introdotti i concetti e gli strumenti essenziali della statistica, descrittiva ed inferenziale, con particolare attenzione alla loro utilizzazione nell'analisi dei dati biologici. Lo scopo del corso è quello di fornire agli studenti tutti quegli strumenti di base utili a leggere, sintetizzare, analizzare ed interpretare i fenomeni osservati, con un'ottica di tipo quantitativo. Nella seconda parte del corso, inoltre, verranno introdotti le principali nozioni della statistica inferenziale (in particolare la stima e la verifica d'ipotesi) in previsione di possibili approfondimenti ed applicazioni in campo biologico (e decisionale in genere).**

# Informazioni sul corso - 2

## Statistica descrittiva

- 1) **Concetti introduttivi: la rilevazione dei fenomeni statistici.**
- 2) **Distribuzione di un carattere e rappresentazione: distribuzioni di frequenza; frequenze relative e percentuali; rappresentazioni grafiche.**
- 3) **Sintesi delle distribuzioni - le medie di posizione e analitiche: moda; mediana; media aritmetica.**
- 4) **Sintesi delle distribuzioni – la variabilità: varianza; scarto quadratico medio; indici relativi di variabilità; standardizzazione.**

# Informazioni sul corso - 3

---

## Statistica inferenziale

- 5) **Concetti di base della probabilità.**
- 6) **Cenni sulle variabili casuali: le distribuzioni normale e t di student.**
- 7) **Campionamento e distribuzioni campionarie.**
- 8) **Teoria della stima - stima puntuale e per intervallo: intervallo di confidenza per la media aritmetica; intervallo di confidenza per una proporzione; determinazione della numerosità campionaria.**

# Informazioni sul corso - 4

## SVOLGIMENTO PROVE IN ITINERE ED ESAME FINALE

- I PROVA INTERMEDIA (U.D. 1-2-3)
- II PROVA INTERMEDIA (U.D. 4-5-6)
- ESAME FINALE ORALE

LE PROVE INTERMEDIE SERVONO A VERIFICARE IL NOSTRO GRADO DI PREPARAZIONE!

A PRESCINDERE DA COME SONO ANDATE LE PROVE INTERMEDIE, TUTTI DEVONO FARE L'ORALE!!!!

# Materiali didattici

---

## Libri di testo

**Whitlock M.C., Schluter D., Analisi statistica dei dati biologici, Zanichelli, Bologna, 2010.**

**Oppure**

**Triola M.M., Triola M.F., Statistica per le discipline biosanitarie, Pearson, Milano, 2009.**

# Contatti

---

**Ricevimento:**

**dove:**

nella mia stanza – Fac. Scienze politiche (III LIV.)

**quando:**

Giovedì 11:00-12:30

Per qualsiasi problema, comunicazione, consiglio,  
proposta, ecc ....:

**Andrea Ciccarelli**

**[aciccarelli@unite.it](mailto:aciccarelli@unite.it)**

**(Ma ditemi chi siete (?!) – se volete una risposta ... )**

# La Statistica - 1

---

- ❑ **Statistica vs Matematica**
- ❑ **Approccio diretto, semplice, poco formale**
- ❑ **Capire e non studiare (a memoria ...)**
- ❑ **Necessità di spiegare le “funzioni” della statistica (per capire bene cosa andremo a fare ...)**
- ❑ **Necessità di imparare un linguaggio comune (essenziale non solo per “capirci” ma anche, come vedremo, per utilizzare le opportune metodologie quantitative)**



## La Statistica - 2

---

La statistica deve fornirci le necessarie metodologie quantitative che ci consentano di:

- ❑ **Raccogliere le informazioni**
- ❑ **Descrivere le informazioni**
- ❑ **Sintetizzare**
- ❑ **Fornire indicazioni (tendenze) quantitative (e, dunque, oggettive)**

# Statistica Descrittiva e Inferenziale

## ❖ Statistica Descrittiva

La statistica descrittiva si occupa di descrivere e sintetizzare (con tabelle, grafici, indici sintetici, ecc..) le informazioni (qualitative e quantitative) relative ad un certo gruppo di “soggetti” al fine di far emergere caratteristiche, andamento e le eventuali relazioni che si verificano all’interno del fenomeno analizzato.

## ❖ Statistica Inferenziale

La statistica inferenziale utilizza le informazioni raccolte da un’indagine campionaria e le “manipola” (grazie all’utilizzo di opportuni “modelli”) in modo da poter generalizzare le indicazioni tratte dal gruppo investigato a tutta la popolazione dal quale tale gruppo è stato estratto.

# Popolazione e Campione

## ❖ Popolazione

Per popolazione si intende l'insieme di tutte le unità che compongono un collettivo. Sono esempi di popolazione:

- Tutti i cittadini italiani
- Tutti gli studenti dell'Università di Teramo
- Tutte le imprese del settore dei servizi
- Tutte le palme presenti in Italia (!?!?)
- Tutti i pesci del mare (!?!?)

## ❖ Campione

Una parte della popolazione (Deve avere “certe” caratteristiche ...)

# Indagini censuarie e campionarie - 1

## Indagine censuaria (o Totale)

Effettuata su **TUTTI** i  
componenti della popolazione

## Indagine campionaria

Effettuata solo su **ALCUNI**  
componenti della popolazione

## Pregi e difetti

- ❑ **Eshaustività**
- ❑ **Ricchezza informazioni**
- ❑ **Maggior dettaglio**
  
- ❑ **Costi elevati**
- ❑ **Lunghi tempi elaborazione**
- ❑ **Difficoltà (impossibilità!) a raggiungere tutti**
- ❑ **Qualità informazioni (?)**

- ❑ **Economicità**
- ❑ **Possibilità di ripetere le indagini continuamente nel tempo**
- ❑ **Indagini mirate in base ad esigenze anche improvvise**
- ❑ **(spesso posso fare solo questo!)**
  
- ❑ **Errore campionario**
- ❑ **Minor livello di dettaglio**



## Le Distribuzioni Statistiche – Distribuzione Unitaria

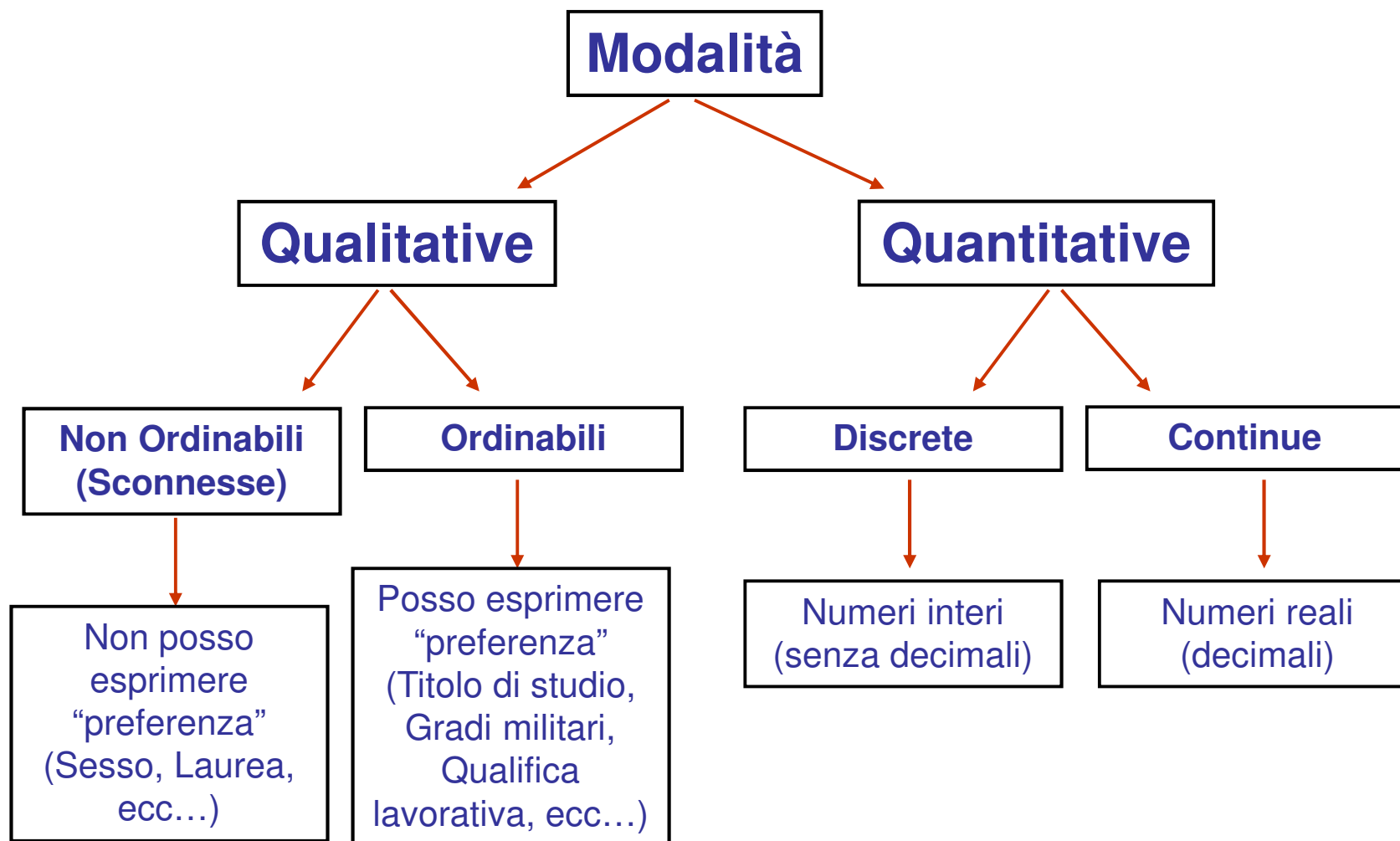
Unità	Sesso	Età	Laurea	Voto	Componenti	Reddito
1	M	24	Economia	109	2	25.100
2	F	25	Economia	110	5	26.500
3	F	24	Matematica	107	3	17.500
4	M	27	Statistica	103	4	35.600
5	F	29	Economia	95	3	76.200
6	F	24	Statistica	107	5	14.700
7	M	25	Matematica	105	4	19.350
8	M	27	Statistica	105	3	22.300
9	M	26	Economia	110	2	18.640
10	F	26	Statistica	110	4	45.020

Unità  
statistiche

Caratteri

Modalità

# Le modalità del carattere



# Le frequenze

- **Absolute ( $n_i$ ):**

rappresentano il numero di volte in cui una determinata modalità del carattere si presenta

- **Relative ( $f_i$ ):**

sono date dal rapporto tra le frequenze assolute e la numerosità totale:  $f_i = n_i/N$

- **Percentuali ( $p_i$ ):**

sono date dal prodotto tra le frequenze relative e 100:  $p_i = f_i \times 100$ .

# Le Distribuzioni Statistiche

---

Le distribuzioni statistiche descrivono il modo in cui uno o più caratteri si manifestano (distribuiscono) in un dato collettivo.

- ❑ **distribuzioni semplici** (un solo carattere)
- ❑ **distribuzioni doppie** (due caratteri)
- ❑ **distribuzioni multiple** (molti caratteri)

Si costruiscono associando ad ogni modalità il numero di volte in cui questa si è presentata



# Distribuzione di Frequenze – Esempio

Unità	Sesso
1	M
2	F
3	F
4	M
5	F
6	F
7	M
8	M
9	M
10	F

$X_i$	$n_i$
Maschi	170
Femmine	130
	300

# Distribuzione di Frequenze – 2

## Esempio

Unità	Laurea	Componenti
1	Economia	2
2	Economia	5
3	Matematica	3
4	Statistica	4
5	Economia	3
6	Statistica	5
7	Matematica	4
8	Statistica	3
9	Economia	2
10	Statistica	4

$$f_i = \frac{n_i}{n}$$

$$N_1 = n_1$$

$$N_2 = N_1 + n_2$$

$$N_3 = N_2 + n_3$$

$X_i$	$n_i$	$f_i$
Economia	120	0,400
Statistica	100	0,333
Matematica	80	0,267
	300	1,000

$X_i$	$n_i$	$N_i$
1	50	50
2	80	130
3	90	220
4	60	280
5	20	300
	300	

## Le Distribuzioni doppie

Unità	Laurea	Sesso
1	Economia	M
2	Economia	F
3	Matematica	F
4	Statistica	M
5	Economia	F
6	Statistica	F
7	Matematica	M
8	Statistica	M
9	Economia	M
10	Statistica	F

	M	F
Economia	75	45
Statistica	35	65
Matematica	60	20
	170	130

# La Rappresentazione Grafica

- ❑ **E' più immediata (intuitiva)**
- ❑ **Facile interpretazione**
- ❑ **La capiscono tutti**

# La Rappresentazione Grafica

---

**Grafico a barre (o a nastri)**

**Grafico a torta**

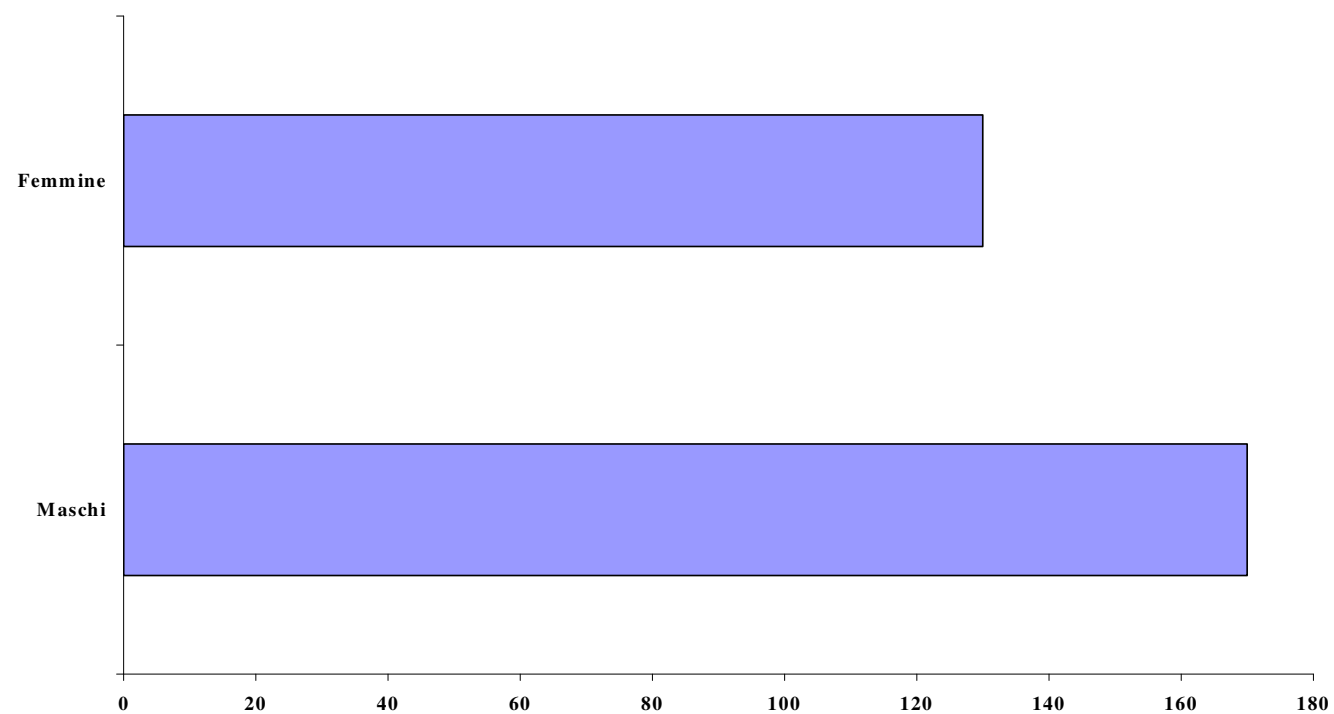
**Istogramma**

## Il grafico a Barre e/o a Nastri

- ❑ Per caratteri qualitativi e quantitativi
- ❑ Generalmente:
  - Qualitativi sconnessi: **Grafico a nastri**
  - Qualitativi ordinati o quantitativi: **Grafico a Barre**
- ❑ Fa capire subito quale modalità ha presentato la più alta frequenza (il rettangolo più “alto”)
- ❑ Base unitaria; altezza proporzionale alla frequenza

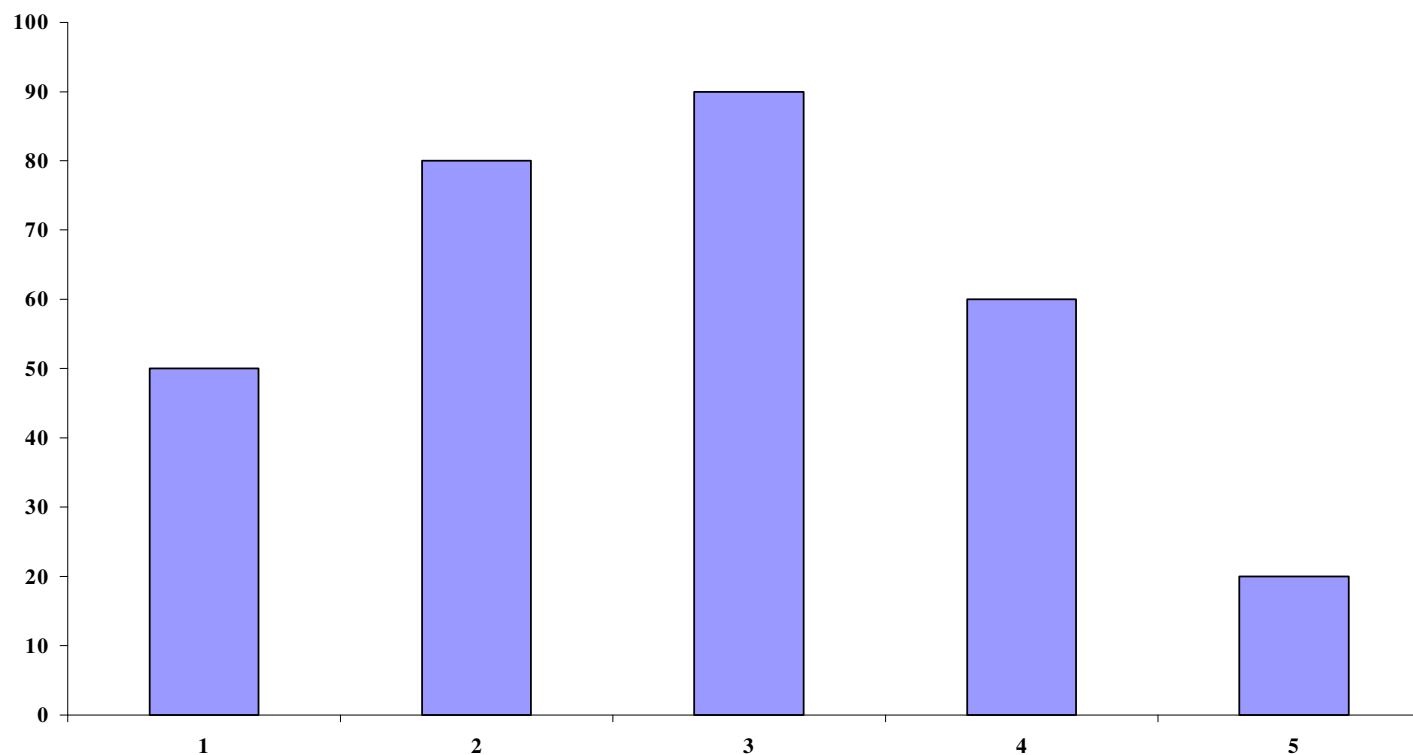
# Il grafico a nastri - esempio

Studenti laureati secondo sesso



# Il grafico a barre - esempio

Studenti laureati secondo il numero dei componenti il nucleo familiare





## Il grafico a Torta

- ❑ **Caratteri qualitativi e quantitativi**
- ❑ **Quando bisogna confrontare una parte con il tutto**
- ❑ **Ogni “fetta” della torta rappresenta l’importanza della modalità che deve descrivere**
- ❑ **L’angolo corrispondente a ciascuna “fetta” rispetta la seguente proporzione:**

$$\frac{p_i}{100} = \frac{g_i}{360} \quad \longrightarrow \quad g_i = \frac{p_i \cdot 360}{100}$$

# Il grafico a Torta - Esempio

**Studenti laureati secondo la tipologia di laurea conseguita**

$X_i$	$n_i$	$p_i$	$g_i$
Economia	120	40,0	144,0
Statistica	100	33,3	120,0
Matematica	80	26,7	96,0
	300	100,0	360,0

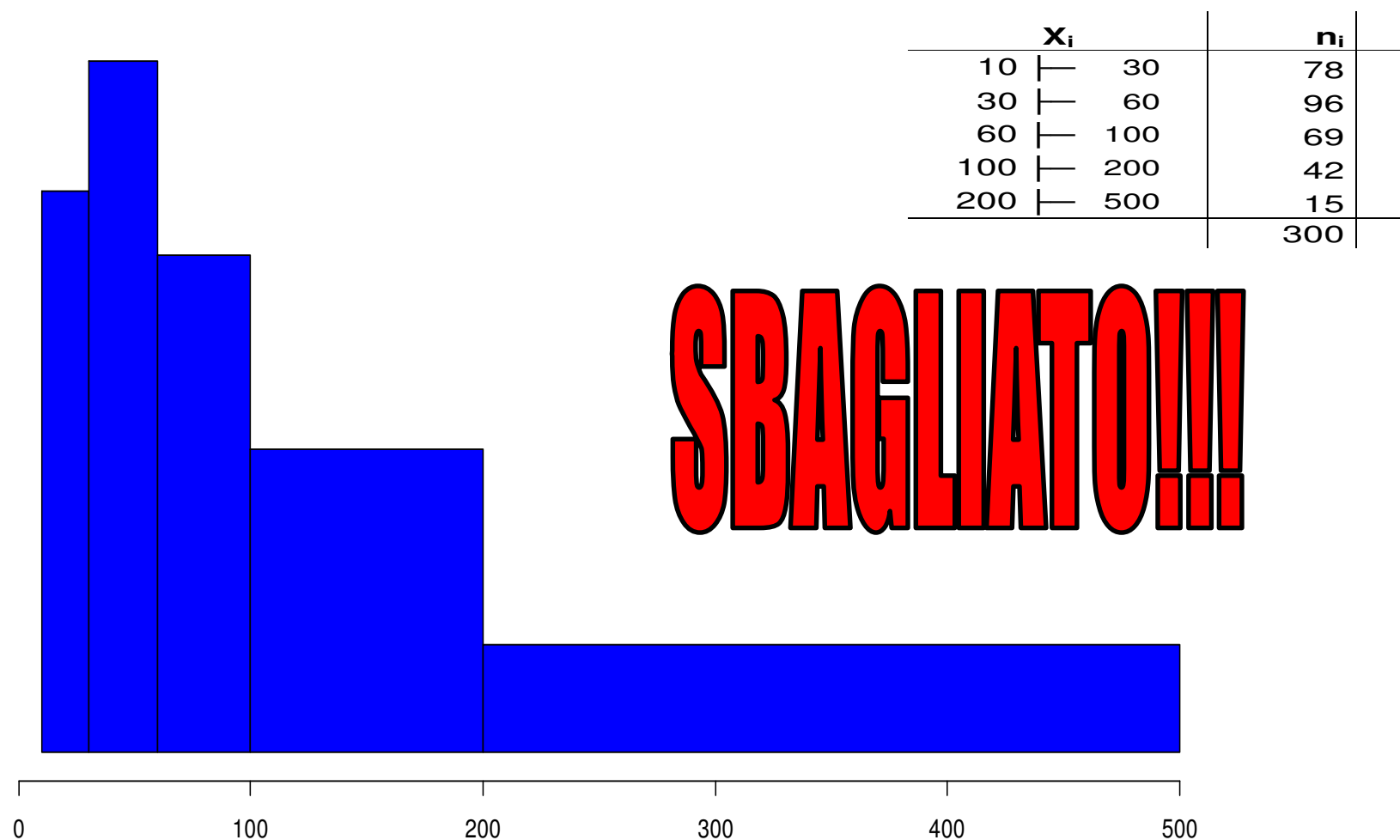
$$p_i = \frac{n_i}{n} \times 100$$

$$g_i = \frac{p_i \cdot 360}{100}$$

# L'Istogramma

- ❑ **Per caratteri quantitativi continui**
- ❑ **Simile al grafico a barre, ma qui le “barre” sono unite, per dare senso di “continuità”**
- ❑ **Base proporzionale all'ampiezza della classe.**

# L'Istogramma - esempio



# L'Istogramma - Costruzione

- ❑ **Bisogna stare attenti all'ampiezza della classe!**
- ❑ **Base proporzionale all'ampiezza della classe. Area proporzionale alla frequenza.**
- ❑ **Altezza proporzionale alla “densità di frequenza”:**

# L'Istogramma – esempio corretto

