

# LEZIONI DI EPIDEMIOLOGIA

*Dott. SIMONE ACCORDINI*

*Lezione n. 1*

- *Introduzione all'epidemiologia*
- *Outcome*
- *Parametro di occorrenza*



*Sezione di Epidemiologia & Statistica Medica  
Università degli Studi di Verona*

## STATISTICA

**insieme di procedure finalizzate al trattamento di informazioni relative a fenomeni collettivi, che si manifestano con determinazioni tipicamente non costanti**

**Oggetto della STATISTICA sono quei fenomeni che variano all'interno di un collettivo di riferimento, la POPOLAZIONE STATISTICA, costituito da UNITA' STATISTICHE o elementari.**



## POPOLAZIONE STATISTICA

qualsiasi insieme di persone, animali, vegetali o cose da cui possono essere raccolte le informazioni

oggetto di interesse dell'indagine:  
insieme di entità sulle cui caratteristiche vogliamo trarre conclusioni

## UNITA' STATISTICA

singola persona, animale, vegetale o cosa appartenente alla popolazione sulla quale viene effettuata la misurazione di uno o più fenomeni

oggetto della raccolta dei dati,  
detentore dell'informazione che vogliamo rilevare e analizzare

Esempio: indagine sulla salute respiratoria nella popolazione adulta italiana (20-44 anni) nel 1998-2000 (Italian Study on Asthma in Young Adults - ISAYA)

POPOLAZIONE  
STATISTICA



*tutti gli adulti di età 20-44 anni residenti in Italia nel periodo dello studio*

UNITÀ  
STATISTICA



*singolo adulto di età 20-44 residente in Italia nel 1998-2000*



## POPOLAZIONI DI INTERESSE PER LA STATISTICA APPLICATA ALLA MEDICINA:

- **insiemi di esseri umani** (residenti in una certa area; soggetti sani, malati oppure deceduti);
- **insiemi di unità amministrative** (reparti, ospedali, comuni);
- ...

esempio (popolazione di residenti - sani o malati):

*indagine ISAYA → adulti di età 20-44 anni residenti in Italia nel 1998-2000 selezionati indipendentemente dallo stato di salute*

esempio (popolazione di soggetti sani):

*sperimentazione sull'effetto del fluoro nel prevenire l'insorgenza di carie nei bambini → bambini sani (senza carie)*

esempio (popolazione di soggetti malati):

*sperimentazione sull'effetto di un chemioterapico per la cura di una particolare patologia tumorale → soggetti che presentano la patologia*



Molte ricerche vengono programmate con lo scopo di pervenire a conclusioni generali, valide per tutte le unità statistiche della popolazione, sfruttando i risultati ottenuti da un **numero ridotto di osservazioni**



## CAMPIONE STATISTICO

sottoinsieme di unità statistiche appartenenti alla popolazione che vengono selezionate per l'analisi (sono quelle realmente studiate)

GENERALIZZAZIONE  
DELLE  
CONCLUSIONI

NB: il campione deve essere **rappresentativo** (con le stesse caratteristiche della popolazione dalla quale è stato estratto)



CAMPIONAMENTO  
CASUALE



### ESEMPIO DI INDAGINE CAMPIONARIA: ISAYA

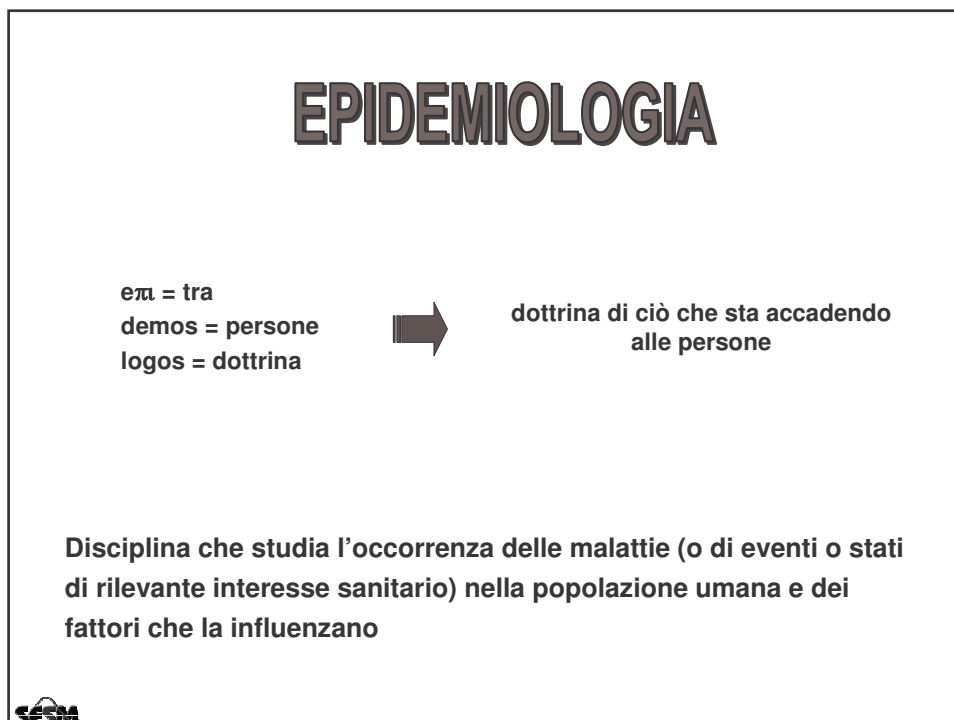
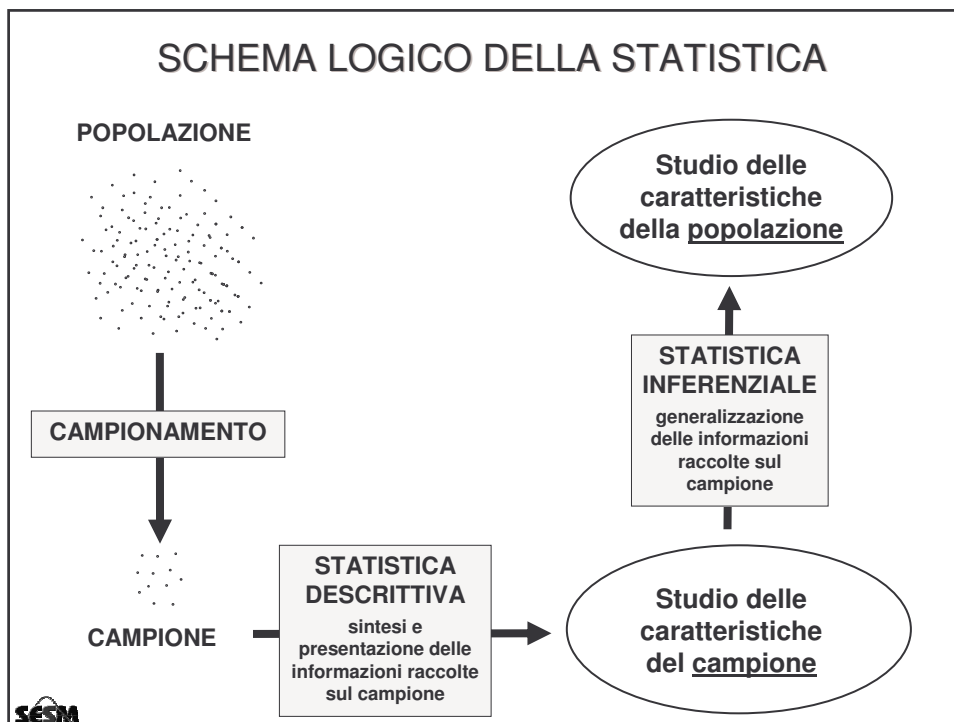
3000 soggetti adulti di età 20-44 anni estratti **casualmente** dalle liste dei residenti in ciascuna delle 9 città coinvolte nell'indagine

### ESEMPIO DI INDAGINE **NON** CAMPIONARIA: censimento

vengono raccolte informazioni da **tutti i residenti** nel territorio italiano nell'anno del censimento (**non** su di un campione)

→ numerosità e composizione demografica della popolazione residente italiana





## Perché l'epidemiologia?

- La distribuzione delle malattie nella popolazione umana può assumere andamenti caratteristici (leggi delle malattie)
- La distribuzione delle malattie nella comunità umana può essere predetta
- Le caratteristiche della distribuzione di una malattia possono suggerire le cause delle malattie e possono condurre a misure per prevenire e controllare le malattie



## TRADIZIONALE CLASSIFICAZIONE DELL'EPIDEMIOLOGIA

### EPIDEMIOLOGIA DESCRITTIVA

Studio della distribuzione di una malattia nella popolazione e dei principali fattori che ne determinano le variazioni.

- Chi si ammala?
- Dove ci si ammala?
- Quando ci si ammala?

### EPIDEMIOLOGIA ANALITICA

Valutazione di specifiche ipotesi sulla relazione tra un fattore e una malattia.

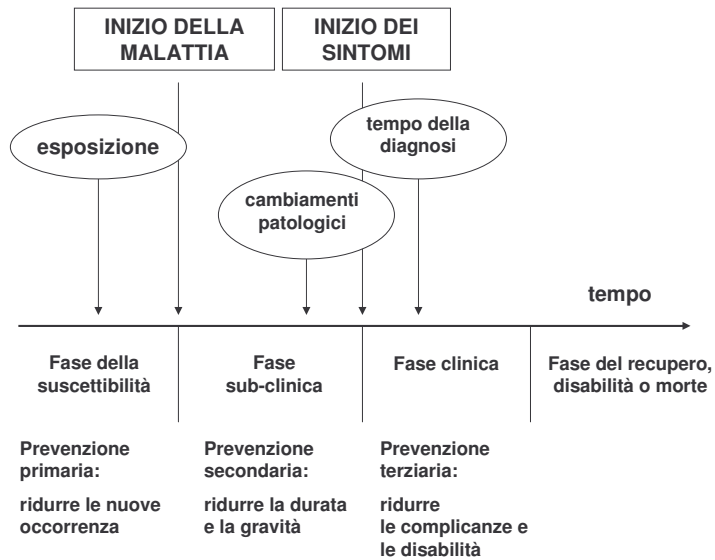
- Come varia la malattia in funzione di differenti tipi di esposizione?
- Come varia la malattia in funzione di differenti suscettibilità individuali?

### EPIDEMIOLOGIA VALUTATIVA

Studio dei cambiamenti nella distribuzione di una malattia dopo l'applicazione di misure di prevenzione.

- Come varia la malattia dopo l'introduzione della misura di prevenzione?

## La storia naturale della malattia



**Outcome:** stato o evento di cui si misura l'occorrenza in una popolazione

→ sempre misurato a livello individuale e rappresenta il risultato di qualche processo fisio/patologico

**Parametro**

**di occorrenza (P):** misura che riassume la frequenza o l'intensità con cui compare l'outcome nella popolazione

Esempi:

**OUTCOME**

**PARAMETRO DI OCCORRENZA**

*presenza di asma*

→ *prevalenza di asma*

*morte per cancro*

→ *rischio di morte per cancro*  
*tasso di mortalità per cancro*

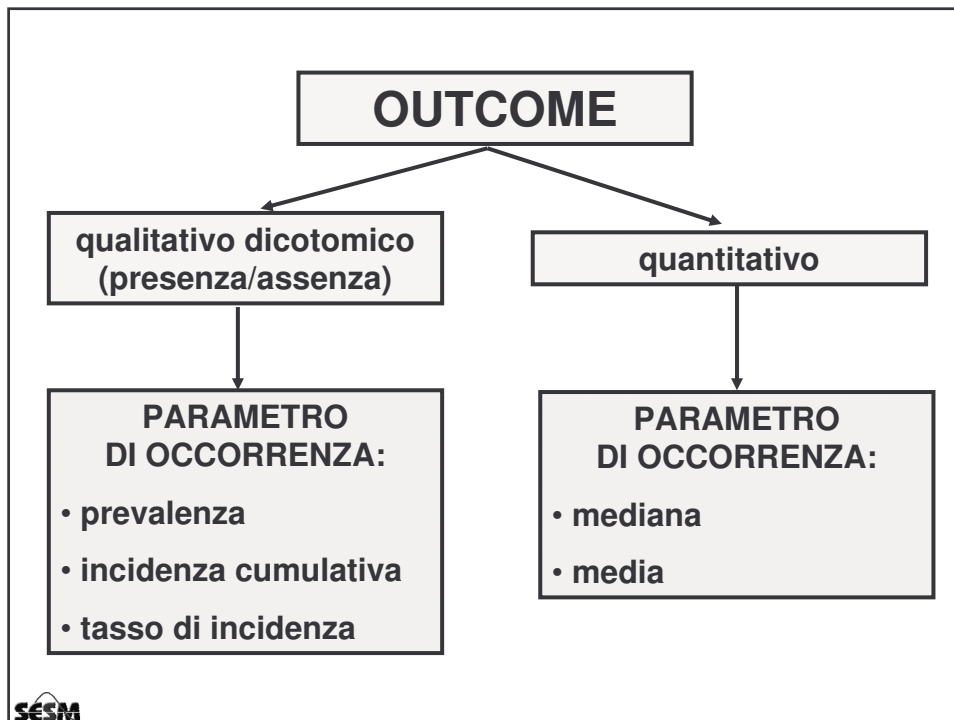
*livello di glicemia*

→ *valore medio della glicemia*

*tempo di sopravvivenza*

→ *mediana del tempo di sopravvivenza*

**SES**



**SES**

## OUTCOME QUALITATIVO DICOTOMICO

In generale, i fenomeni dicotomici ( $X$ ) hanno **DISTRIBUZIONE BERNOULLIANA**  
( $\rightarrow$  MODELLO TEORICO)

Presenza della caratteristica ( $X = 1$ ):

$$\text{Prob}(X = 1) = \pi$$

Assenza della caratteristica ( $X = 0$ ):

$$\text{Prob}(X = 0) = 1 - \pi$$

**CARATTERISTICA  
DELLA POPOLAZIONE:**

$\pi$  = probabilità (rischio)  
nella popolazione

**Esempio:**

*presenza di asma (STATO)  $\rightarrow$  probabilità di avere l'asma*

*morte per cancro (EVENTO)  $\rightarrow$  probabilità di morire per cancro*

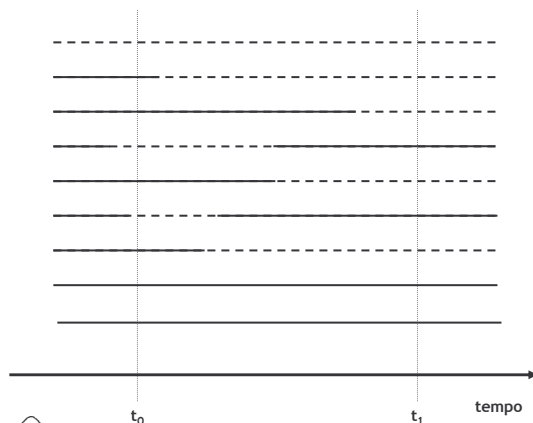


**PREVALENZA (P):** proporzione della popolazione affetta da una particolare malattia in un dato istante o periodo di tempo  
 $\rightarrow$  probabilità di avere l'outcome (stato)

$$P = \frac{r}{n}$$

$n$  = numerosità della popolazione al tempo  $t_0$  o nel periodo tra  $t_0$  e  $t_1$

$r$  = numero di casi al tempo  $t_0$  o nel periodo tra  $t_0$  e  $t_1$



In  $t_0$ :

numero di soggetti = 9

numero di casi = 3

$$\Rightarrow P = 3/9 = 0.33$$

Tra  $t_0$  e  $t_1$ :

numero di soggetti = 9

numero di casi = 7

$$\Rightarrow P = 7/9 = 0.78$$





**Esempio [de Marco, et al. *Clin Exp Allergy* 2002; 32: 1405-12]:**

*In uno studio condotto in Italia nel 1998-2000 (indagine ISAYA), 706 soggetti di età 20-44 anni hanno riportato di avere avuto attacchi di asma negli ultimi 12 mesi ad un questionario di screening. I soggetti che hanno fornito una risposta valida alla domanda sugli attacchi di asma sono stati 18804.*

$$\text{Pr} = \frac{706}{18804} = 0.0375 \Rightarrow \text{Pr} = 3.75 \%$$



**Esercizio:**

**1/1/1999: 4 casi di una certa malattia presenti in una popolazione di 100 soggetti**

**1/1/1999 - 1/1/2003: 3 soggetti sono guariti**

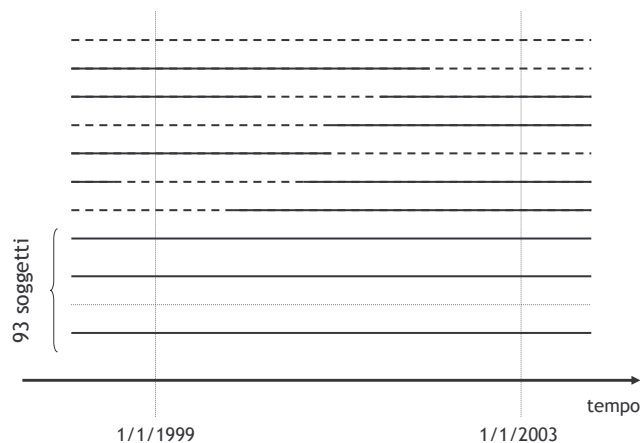
**1 soggetto sano ha sviluppato la malattia ed è guarito**

**2 soggetti sani hanno sviluppato la malattia ma non sono guariti**

**A) Qual è la prevalenza all'1/1/1999?**

**B) Qual è la prevalenza all'1/1/2003?**

**C) Qual è la prevalenza nel periodo 1/1/1999 - 1/1/2003?**

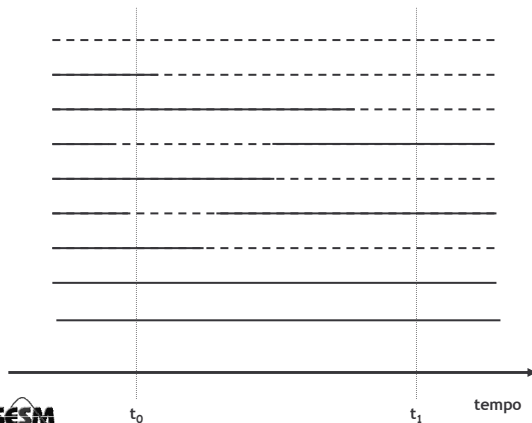


**INCIDENZA CUMULATIVA (CI):** proporzione della popolazione libera da una particolare malattia (in un dato momento) che ha sviluppato la malattia durante un periodo di tempo successivo  
 → probabilità di sperimentare l'outcome (**evento**)

$$CI = \frac{r}{n} \text{ tra } t_0 \text{ e } t_1$$

n = popolazione totale a rischio in  $t_0$

r = numero di nuovi casi nel periodo tra  $t_0$  e  $t_1$



In  $t_0$ :  
 numero di soggetti a rischio = 6

Tra  $t_0$  e  $t_1$ :  
 numero di nuovi casi = 4

⇒  $CI = 4/6 = 0.67$  tra  $t_0$  e  $t_1$

**Esempio:**

*In uno studio sulla relazione tra contraccettivi orali (CO) e sviluppo di batteriuria, 2390 donne tra i 16 e 45 anni, libere da malattia, sono state seguite per 3 anni. Di queste, 486 usavano CO all'inizio del 1973. Tra il 1973 e il 1976, 27 di queste svilupparono la malattia.*

$$CI = \frac{27}{486} = 0.056 \Rightarrow CI = 5.6\%$$

probabilità che una donna (di età 16-45 anni) utilizzatrice di CO sviluppi un'infezione urinaria in un periodo di tre anni

**NB:** 5.6% in 3 anni ≠ 5.6% in 3 mesi ≠ 5.6% in 10 anni

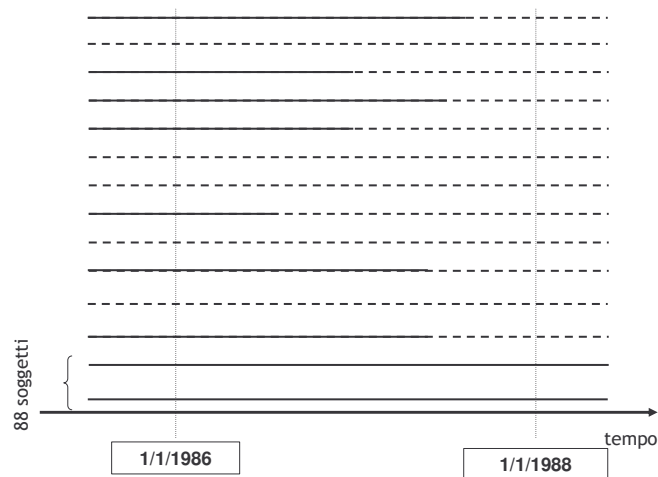
**Esercizio:**

*1/1/1986: 5 casi di angina presenti in una popolazione di 100 soggetti*

*1/1/1986-1/1/1988: 7 nuovi casi di angina*

**A) Qual è la prevalenza di angina nei 2 anni?**

**B) Qual è l'incidenza cumulativa nei 2 anni?**



**SES**

**MA ...**

- i soggetti possono entrare nello studio in momenti diversi
- alcuni soggetti vengono persi al follow-up (→ drop-out)

**POICHE' ...**

- un soggetto è effettivamente a rischio solo fino a quando non sviluppa la malattia

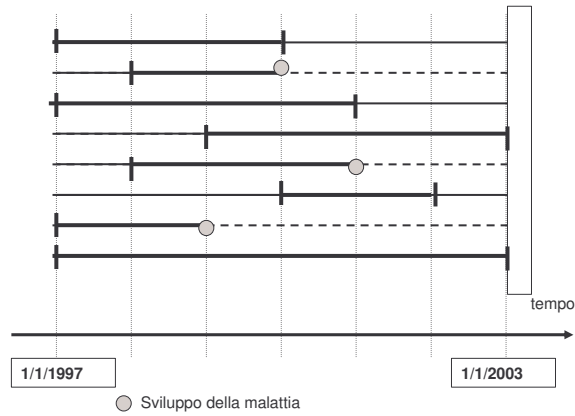


**PERSONE-TEMPO:**

somma di tutti i tempi di osservazione dei soggetti a rischio

**SES**

Esempio:



Persone-tempo = 3+2+4+4+3+2+6 = 26 persone-anno

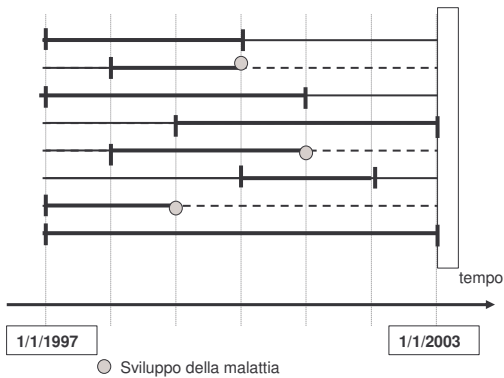


**TASSO DI INCIDENZA (I):** misura della velocità media di comparsa dei nuovi eventi di malattia in una popolazione (non è una probabilità)

$$I = \frac{r}{PT} \text{ tempo}^{-1}$$

PT = persone-tempo

r = numero di nuovi casi nel periodo tra  $t_0$  e  $t_1$



Tra  $t_0$  e  $t_1$ :

PT = 26 persone-anno

numero di nuovi casi = 3

$I = 3/26$  persone-anno

= 0.115 casi per anno<sup>-1</sup>



Il tasso:

- è espresso come numero di casi per tempo<sup>-1</sup>
- viene generalmente moltiplicato per una costante di convenienza (100; 1000; ...)

$$I = \frac{3}{26} * 1000 = 115 \text{ (x 1000 anni}^{-1}\text{)}$$

115 casi ogni 1000 persone osservate per un anno

SESIM

- l'unità di tempo è arbitraria: il tasso può essere espresso in giorni<sup>-1</sup>, mesi<sup>-1</sup>, anni<sup>-1</sup>, ...

$$I = \frac{3}{26 \text{ anni}} * 1000 = 115 \text{ (x 1000 anni}^{-1}\text{)}$$

115 casi ogni 1000 persone osservate per un anno

$$I = \frac{3}{312 \text{ mesi}} * 1000 = 9.6 \text{ (x 1000 mesi}^{-1}\text{)}$$

9.6 casi ogni 1000 persone osservate per un mese

SESIM

Esempio:

*In uno studio sulla relazione tra l'utilizzo di ormoni nella post-menopausa e l'insorgenza di CHD, si sono evidenziati 90 nuovi casi tra 32317 donne in menopausa, per un periodo totale di follow-up di 105786 persone-anno.*

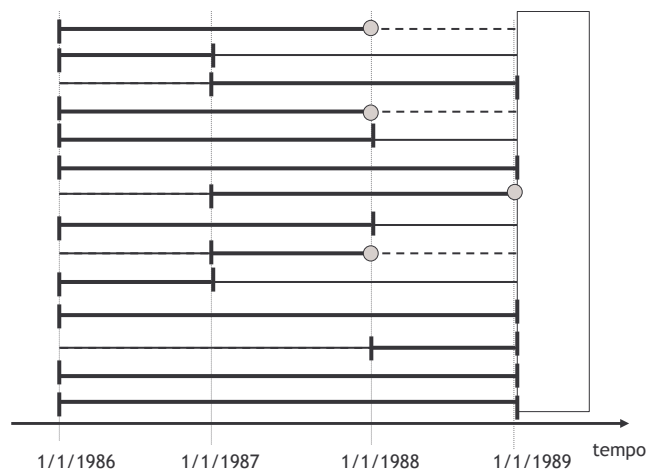
$$I = \frac{90}{105786} = 0.0085 \text{ x anni}^{-1}$$

0.85 casi per 100 persone all'anno  
*oppure*  
8.5 casi per 1000 persone all'anno  
*oppure*  
85 casi per 10000 persone all'anno

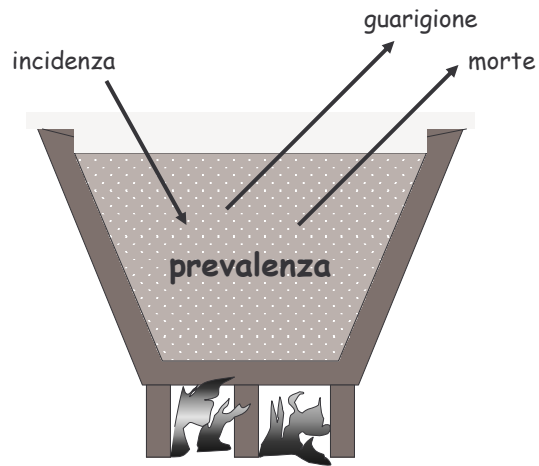


Esercizio:

*Calcolate il tasso di incidenza sulla base dell'esperienza di malattia dei 14 soggetti tra l'1/1/1986 e l'1/1/1989 rappresentata nel seguente grafico.*



### Relazione fra incidenza e prevalenza

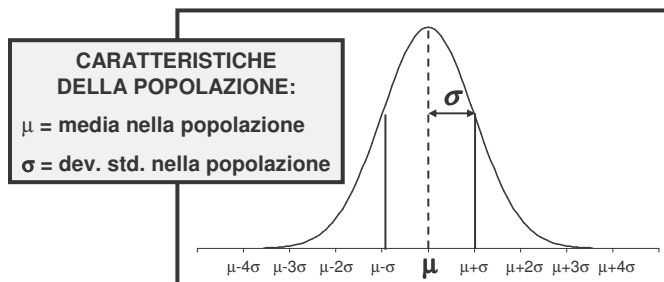


$$Pr = I * \text{durata media della malattia}$$



### OUTCOME QUANTITATIVO (DISTRIBUZIONE SIMMETRICA)

Molte variabili biologiche ( $X$ ) hanno **DISTRIBUZIONE NORMALE** o approssimativamente normale ( $\rightarrow$  MODELLO TEORICO)

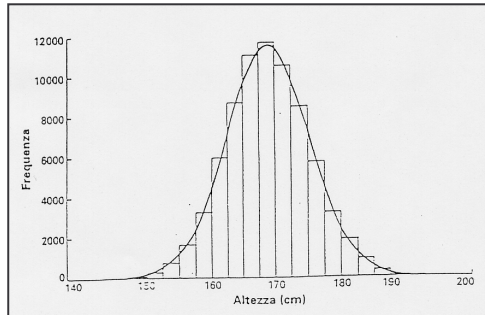


$$\text{Prob}(\mu-\sigma < X < \mu+\sigma) = 0.68$$



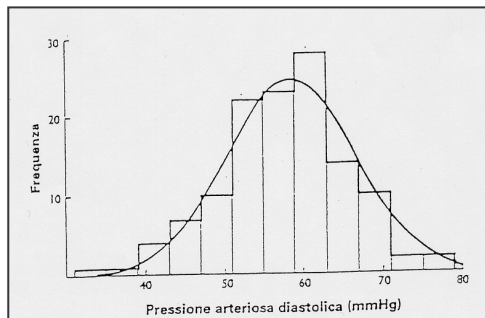
**Esempio:**

*altezza di giovani maschi adulti*



**Esempio:**

*pressione arteriosa diastolica del sangue di scolari*



**STATISTICHE CAMPIONARIE:**  
 $\bar{x}$  = media nel campione  
s = dev. std. nel campione



**CARATTERISTICHE DELLA POPOLAZIONE:**  
 $\mu$  = media nella popolazione  
 $\sigma$  = dev. std. nella popolazione

**Esempio (outcome quantitativo con distribuzione simmetrica):**

livello di emoglobina in g/100 ml (X) misurato in un campione di donne

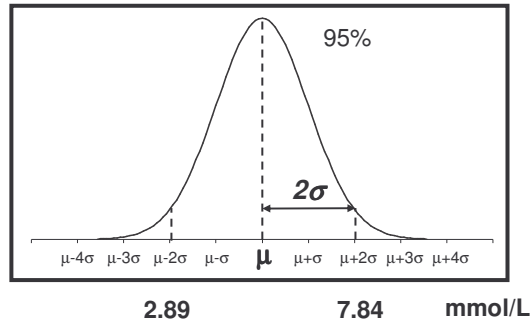
media ( $\pm$  deviazione standard) = 14.00 ( $\pm$  1.02) g/100 ml

MISURA DI POSIZIONE = PARAMETRO DI OCCORRENZA





## VALORI DI NORMALITA'



### Esempio:

I valori di normalità dell'UREA sono compresi tra **2.89** e **7.84** mmol/L.



## OUTCOME QUANTITATIVO (DISTRIBUZIONE ASIMMETRICA)

esempio:

durata della sopravvivenza di 9 individui affetti da tumore al pancreas (in mesi)

CAMPIONE 4 5 6 9 12 12 13 17 84

Media = 18 (senza il soggetto *outlier* sarebbe 9.7)

Mediana = 12

MISURA DI POSIZIONE = PARAMETRO DI OCCORRENZA

La media aritmetica è poco "robusta" in presenza di **valori anomali** (outliers)

