LEZIONI DI STATISTICA MEDICA

Prof. SIMONE ACCORDINI

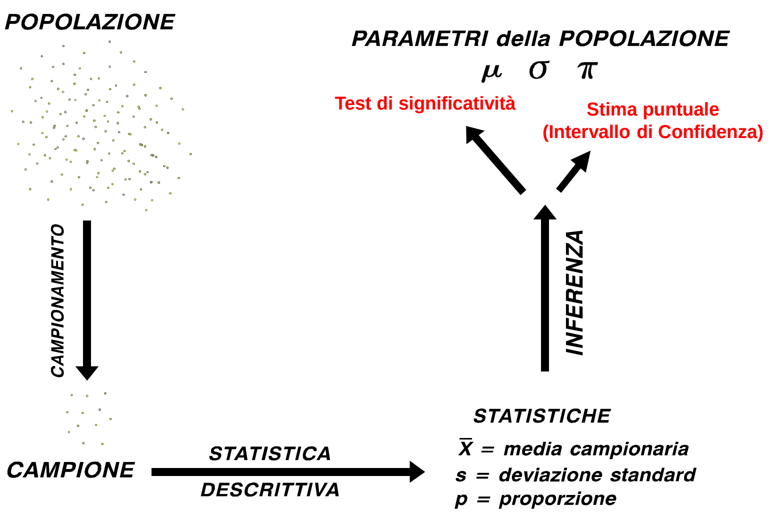
Lezione n.11

- Principi dell'inferenza statistica

- Intervallo di confidenza

Sezione di Epidemiologia & Statistica Medica
Università degli Studi di Verona

INFERENZA



METODI STATISTICI DELL'INFERENZA

1. Stimare il parametro di occorrenza (μ, π, I) in una o più popolazioni e/o stimare la misura di associazione (RR, OR, RD)

$\Rightarrow$ **STIMA PUNTUALE**

2. Associare alla stima puntuale una **misura di precisione**
$\rightarrow$ **misura dell'errore di stima**

$\Rightarrow$ **INTERVALLO DI CONFIDENZA**

METODI STATISTICI DELL'INFERENZA

3. Verificare se il parametro di occorrenza (μ, π, I) in una popolazione ha un valore diverso da quello ipotizzato

[ad esempio: la prevalenza di asma negli adulti italiani (π) è diversa da 0.02?] …

… verificare se il parametro di occorrenza varia tra due o più popolazioni:
($\mu_1 \neq \mu_0$, $\pi_1 \neq \pi_0$, $I_1 \neq I_0$ $\Leftrightarrow$ RR $\neq$ 1, OR $\neq$ 1, RD $\neq$ 0)

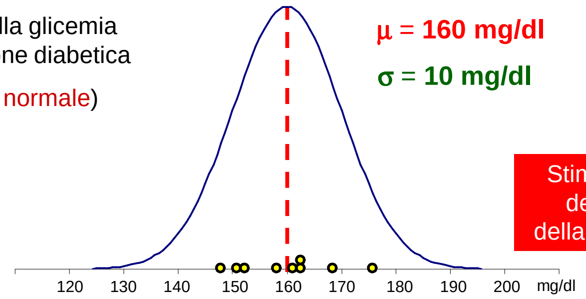
$\rightarrow$ **la differenza osservata è dovuta al caso oppure è dovuta ad altri fattori (trattamento, fattori di rischio, …)?**

$\Rightarrow$ **TEST STATISTICO**

INFERENZA SULLA MEDIA DI UNA POPOLAZIONE

Distribuzione della glicemia
in una popolazione diabetica

(modello teorico normale)



μ = 160 mg/dl

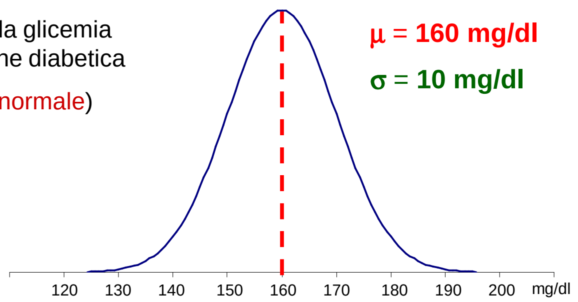
σ = 10 mg/dl

Stima puntuale
della media
della popolazione

campione: **162 152 158 162 168 161 148 176 150** → $\bar{x}$ = **159.7**

Distribuzione della glicemia
in una popolazione diabetica

(modello teorico normale)



μ = 160 mg/dl

σ = 10 mg/dl

Campione
studiato

I° campione: **162 152 158 162 168 161 148 176 150** → $\bar{x}$ = **159.7**

II° campione: **152 164 157 180 156 163 165 166 178** → $\bar{x}$ = **164.6**

III° campione: **157 142 163 162 152 149 152 180 151** → $\bar{x}$ = **156.4**

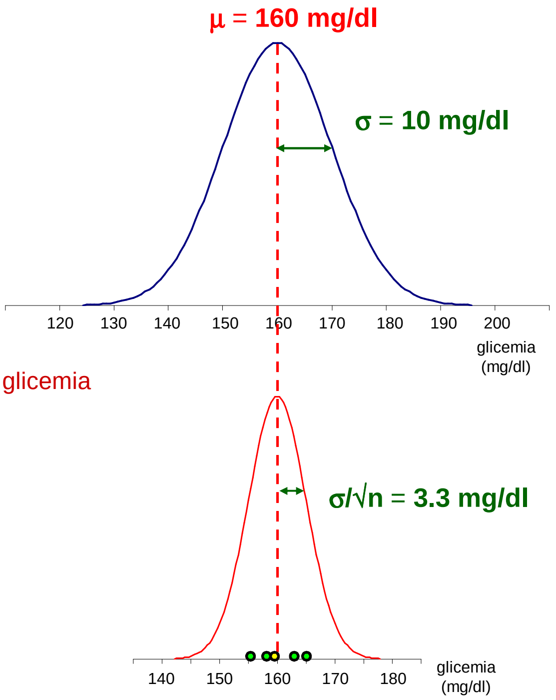
IV° campione: **162 154 168 160 155 172 162 152 140** → $\bar{x}$ = **158.3**

V° campione: **163 169 152 147 158 163 173 160 181** → $\bar{x}$ = **162.8**

PRINCIPIO DEL CAMPIONAMENTO RIPETUTO

Distribuzione della glicemia
in una popolazione diabetica

Distribuzione della media della glicemia
in campioni di numerosità = 9



DISTRIBUZIONE CAMPIONARIA DI UNA MEDIA

Sia $\bar{x}$ la media stimata in un campione casuale di dimensione *n* selezionato da una popolazione con media μ e deviazione standard σ:

1-2) la distribuzione campionaria di $\bar{X}$ ha:

$E[\bar{X}] = \mu$

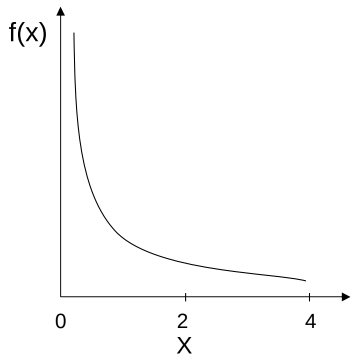
$DS[\bar{X}] = ES[\bar{X}] = \sigma/\sqrt{n}$

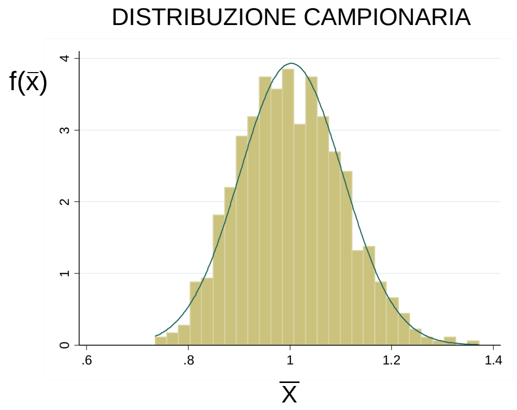
ERRORE STANDARD della media $\rightarrow$ misura della precisione della stima

3) TEOREMA DEL LIMITE CENTRALE: se la dimensione campionaria è sufficientemente grande ($n \geq 30$), allora la distribuzione campionaria di $\bar{X}$ è <u>*approssimativamente*</u> *normale*, indipendentemente dalla distribuzione della variabile nella popolazione

Esempio (teorema del limite centrale - media):

800 campioni di dimensione n = 100 generati casualmente da una distribuzione esponenziale

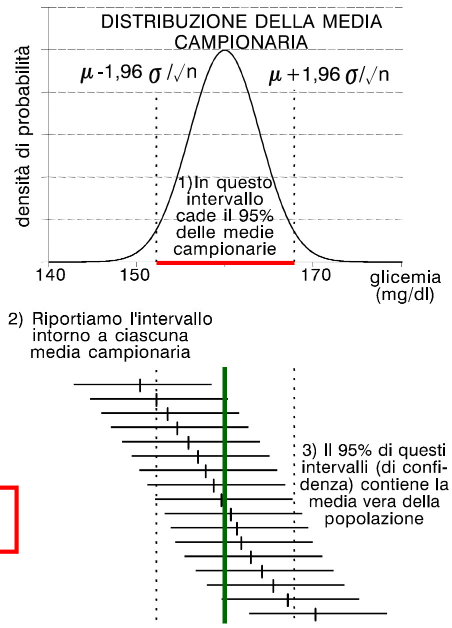




Intervallo di confidenza della media in una popolazione: IC95%(μ)

Per intervallo di confidenza della media μ, si intende un intervallo delimitato da due limiti L_{inf} (limite inferiore) ed L_{sup} (limite superiore) che abbia una definita probabilità (livello di confidenza) di contenere il vero valore (ignoto) del parametro nella popolazione:

$$\boxed{\text{prob}(L_{inf} < \mu < L_{sup}) = 0.95}$$



IC95%(μ)

$$\underbrace{\bar{x} - 1.96\ \sigma/\sqrt{n}}_{L_{inf}}\ ,\ \underbrace{\bar{x} + 1.96\ \sigma/\sqrt{n}}_{L_{sup}}$$

Esempio: Inferenza sulla media della glicemia in una popolazione diabetica

1. Stimare il parametro di occorrenza (μ)

⇒ **STIMA PUNTUALE** (n = 9) **$\bar{x}$ = 159.7** mg/dl

2. Associare alla stima puntuale una **misura di precisione**

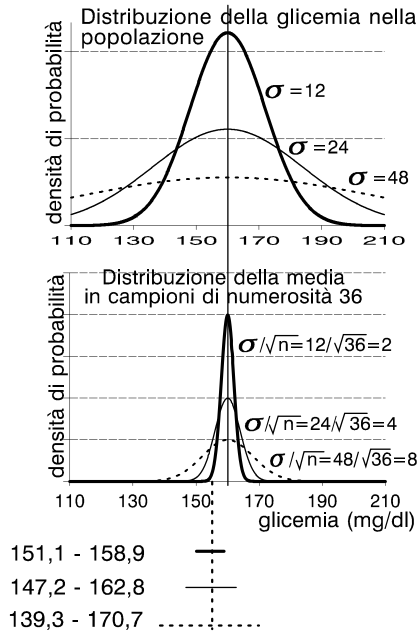
⇒ **INTERVALLO DI CONFIDENZA**

IC95%(μ) = **159.7 ± 1.96*10/√9** = [**153.2** mg/dl, **166.2** mg/dl]

L'IC diminuisce se diminuisce la variabilità nella popolazione (σ)

$IC95\%(\mu) = \bar{x} \pm 1.96\, \sigma / \sqrt{n}$

n = 36



L'IC diminuisce se aumenta la numerosità del campione (n)

$IC95\%(\mu) = \bar{x} \pm 1.96\, \sigma / \sqrt{n}$

σ = 24