

## Metodi statistici per lo studio dei fenomeni biologici

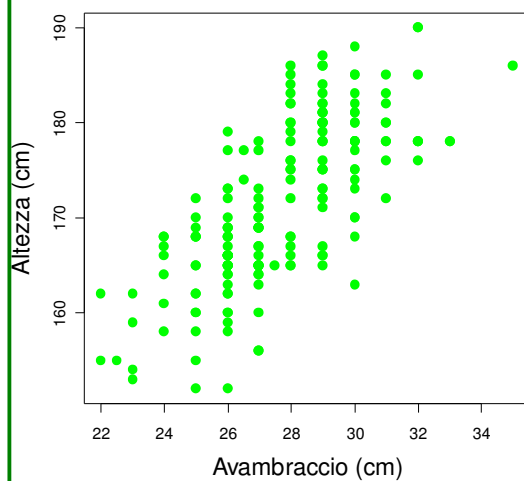
Alla fine di questa lezione dovrete essere in grado di:

- ❑ descrivere il significato del coefficiente di correlazione fra due variabili quantitative
- ❑ descrivere il significato della retta di regressione lineare fra due variabili quantitative



### Relazione tra due variabili quantitative

Popolazione



C'è una associazione tra le due variabili?

Come la misuro?

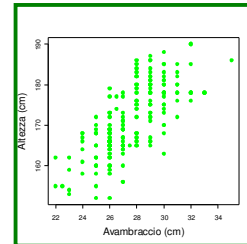


Coefficiente di Correlazione

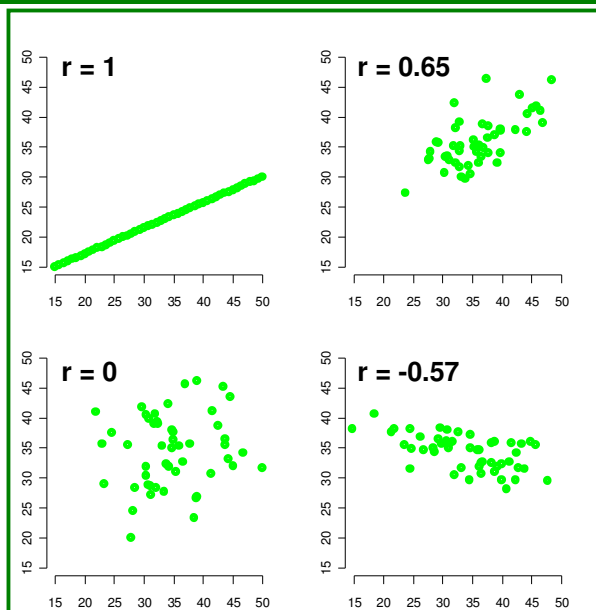
## Coefficiente di Correlazione ( $r$ )

### Caratteristiche

1. Permette di riassumere la forza della relazione lineare tra due variabili
2. Assume valori tra  $-1 \leq r \leq 1$
3. Il segno di  $r$  indica il senso dell'associazione
4. I valori  $r = 1$  e  $r = -1$  si verificano quando esiste una perfetta relazione lineare tra le due variabili
5. Il valore  $r = 0$  si verifica quando la relazione lineare tra le due variabili non esiste

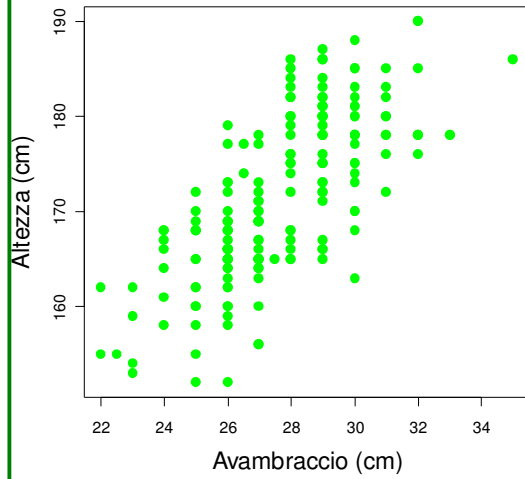


## Diverse correlazioni possibili



## Relazione tra due variabili quantitative

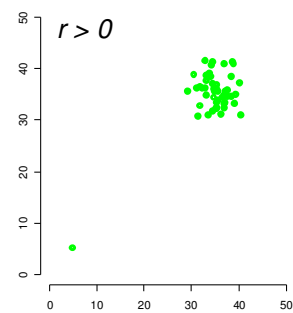
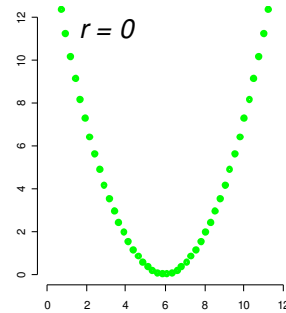
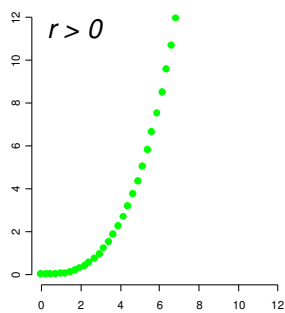
### Popolazione



E in questo caso  
quanto vale il  
coefficiente di  
correlazione?

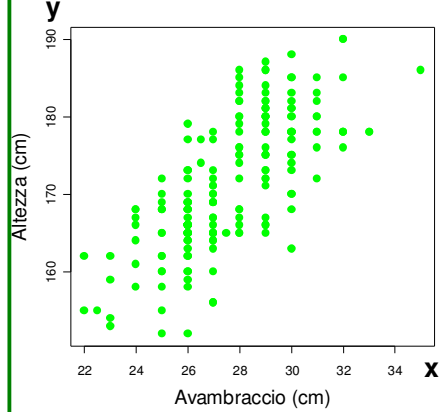
$$\rho = 0.72$$

## Quando il coefficiente di correlazione è inappropriato?



## Retta di regressione

### Popolazione



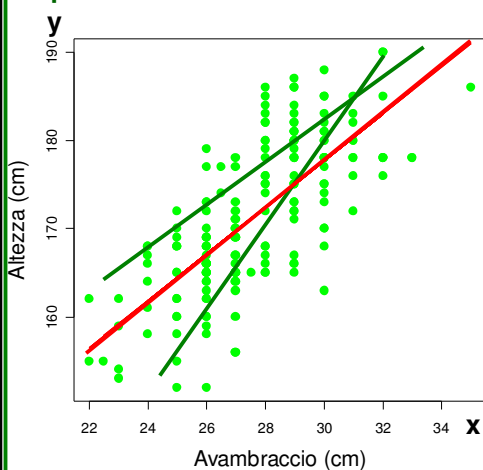
La retta di regressione consente di esaminare la variazione della *variabile dipendente* ( $y$ ), che corrisponde ad una determinata variazione della *variabile indipendente* ( $x$ )



$$y = \alpha + \beta \cdot x + \varepsilon$$

## Retta di regressione

### Popolazione



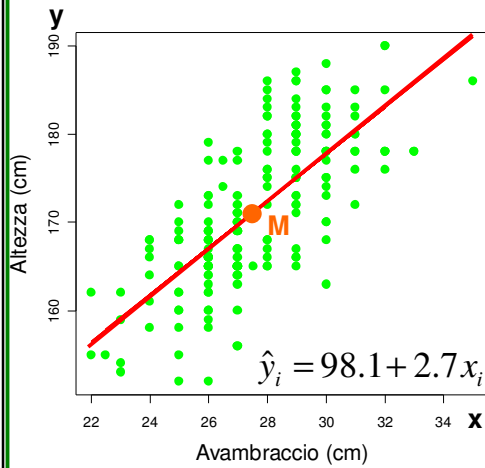
Qual è la *migliore* retta che può essere tracciata a partire dai dati ?



Scegliamo la retta che rende minima la variabilità totale tra i dati osservati e la retta stessa

## Calcolo dei parametri della retta

### Popolazione



Il metodo dei minimi quadrati determina i parametri ( $\alpha$   $\beta$ ) della retta di regressione che rendono minima la somma:

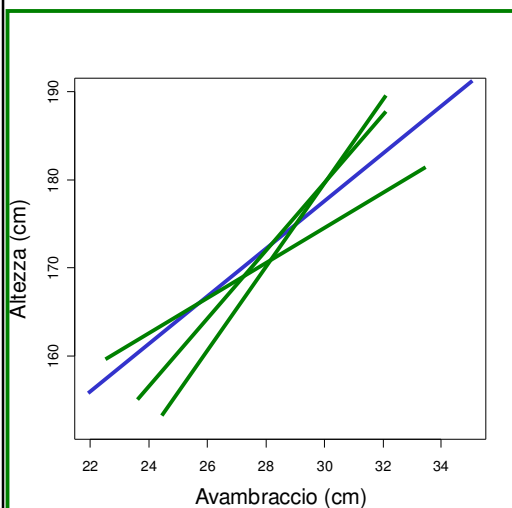
$$\sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$\alpha$ : Intercetta della retta

$\beta$ : Pendenza della retta

Questa retta passa per **M** il punto di coordinate  $(\bar{x}, \bar{y})$

## Popolazione e Campione



Retta di regressione della popolazione

Stime campionarie della Retta di regressione

$$\hat{y}_i = a + bx_i$$

$a$  è una stima puntuale di  $\alpha$

$b$  è una stima puntuale di  $\beta$

## Test d'ipotesi e Intervalli di confidenza

### Test d'ipotesi

Il più diffuso test d'ipotesi è quello sulla verifica della mancanza di relazione lineare tra le due variabili

Il test ha come ipotesi nulla  $\beta = 0$

### Intervallo di confidenza

Abbiamo visto come la variabilità della stima di  $a$  e  $b$  portano a andamenti differenti della retta di regressione stimata sul campione rispetto a quella della popolazione

È possibile calcolare gli intervalli di confidenza per  $a$ ,  $b$  e per l'intera retta di regressione

## Intervallo di confidenza della retta di regressione

