

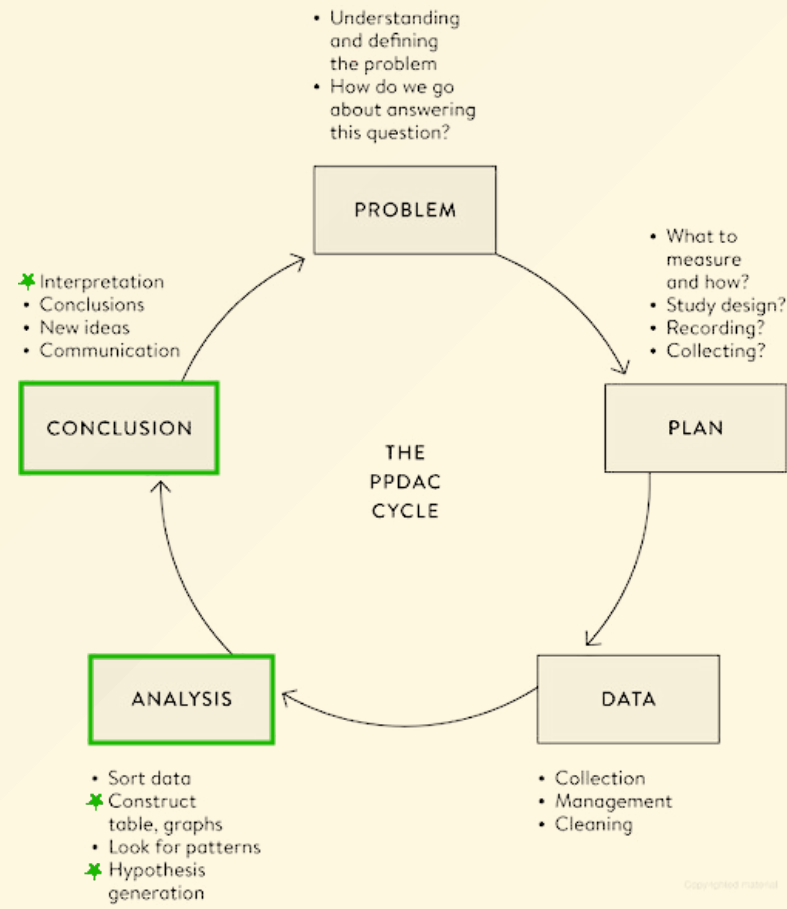
## **Lezione 6**

# **La distribuzione Normale**

# Obiettivi di apprendimento

- Conoscere le caratteristiche della distribuzione Normale
- Conoscere le caratteristiche della distribuzione Normale Standardizzata

# Le fasi della ricerca



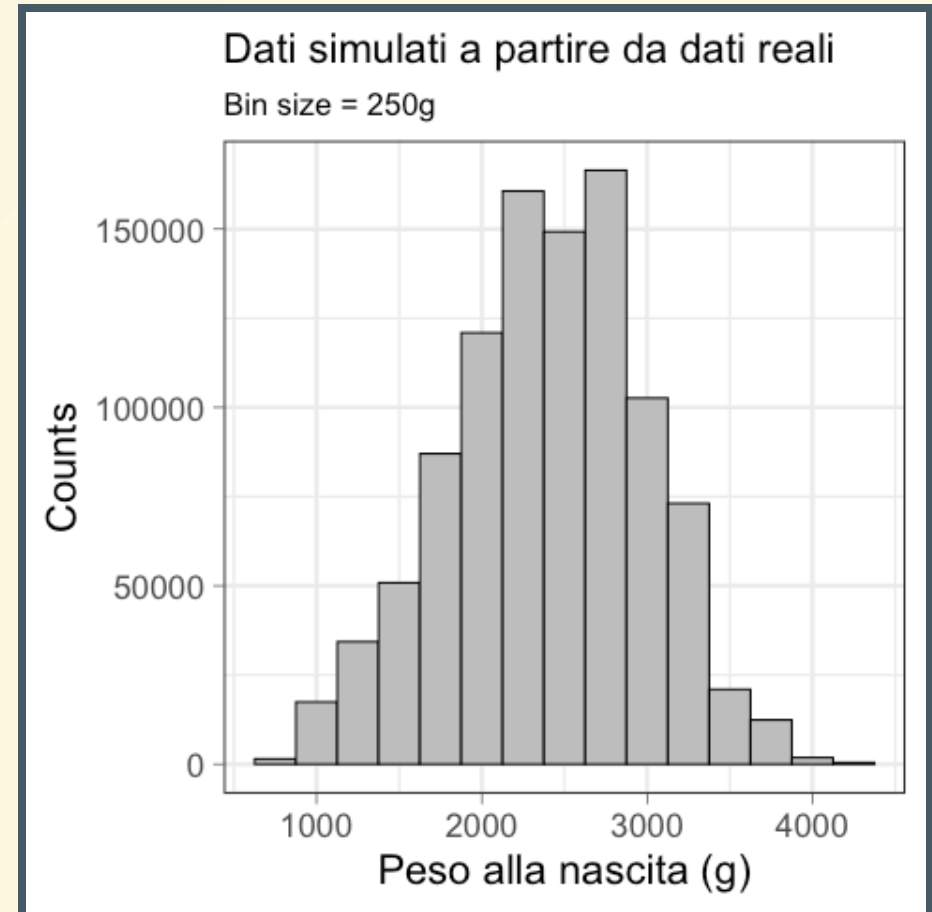
# La distribuzione della popolazione

“ Qual è la distribuzione del peso alla nascita per i gemelli inglesi? ”

# La distribuzione della popolazione

“ Qual è la distribuzione del peso alla nascita per i gemelli inglesi? ”

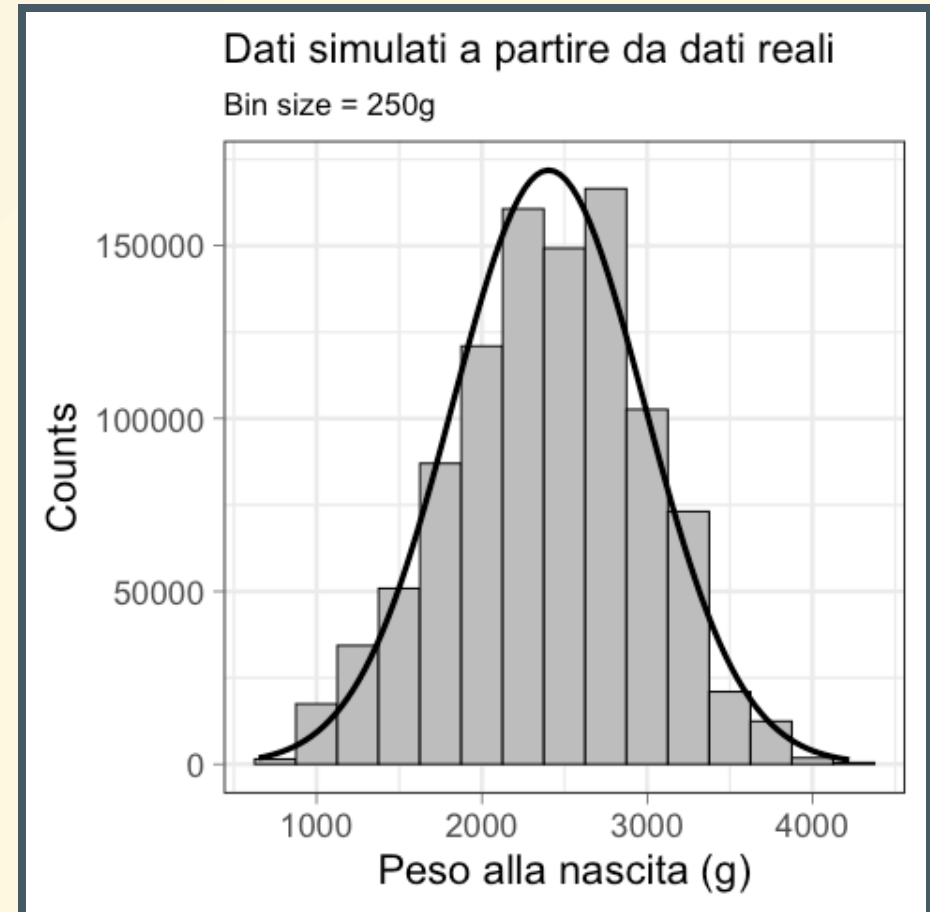
$N = 1,000,000$   
 $\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$   
*mediana* = 2408 g



# La distribuzione della popolazione

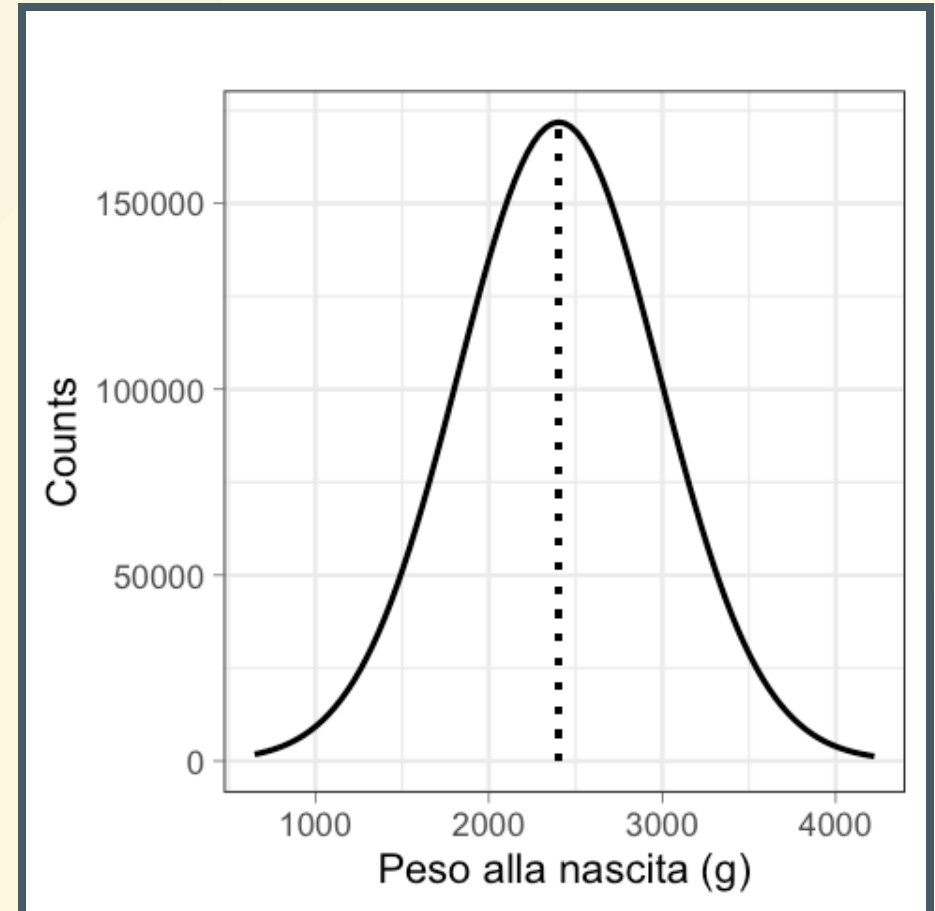
“ Qual è la distribuzione del peso alla nascita per i gemelli inglesi? ”

$N = 1,000,000$   
 $\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$   
*mediana* = 2408 g



# La distribuzione Normale

- $\mathcal{N} = (\mu, \sigma^2)$
- moda  $\equiv$  media  $\equiv$  mediana
- Simmetrica



# Parametri vs statistiche

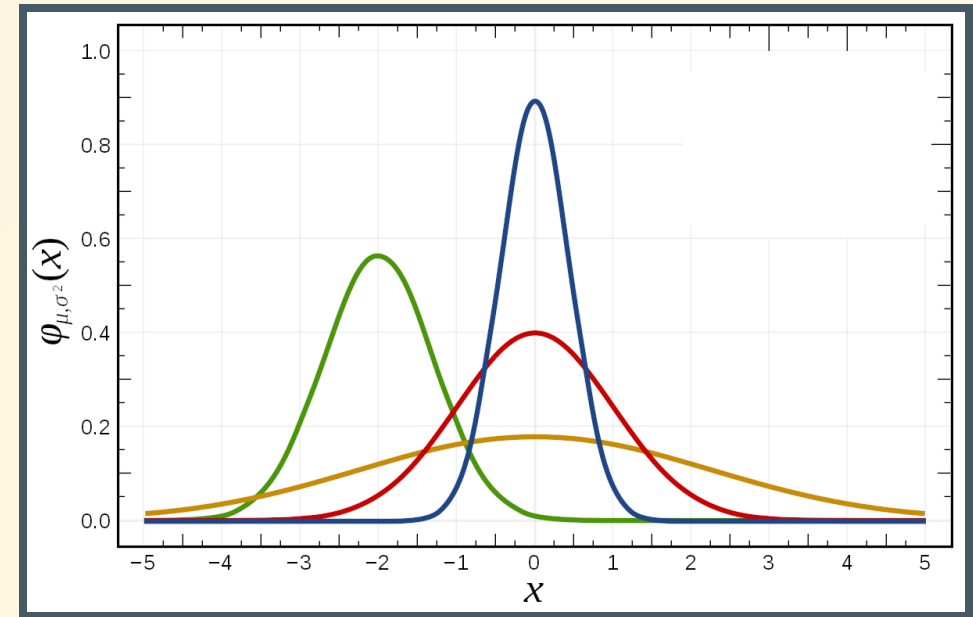
Statistica	Popolazione	Campione
Numerosità	$N$	$n$
Media	$\mu$	$\bar{x}$
Deviazione Standard	$\sigma$	$s$
Proporzione	$\pi$	$p$



# Esercizio #1

? Qual è la curva con la media maggiore?

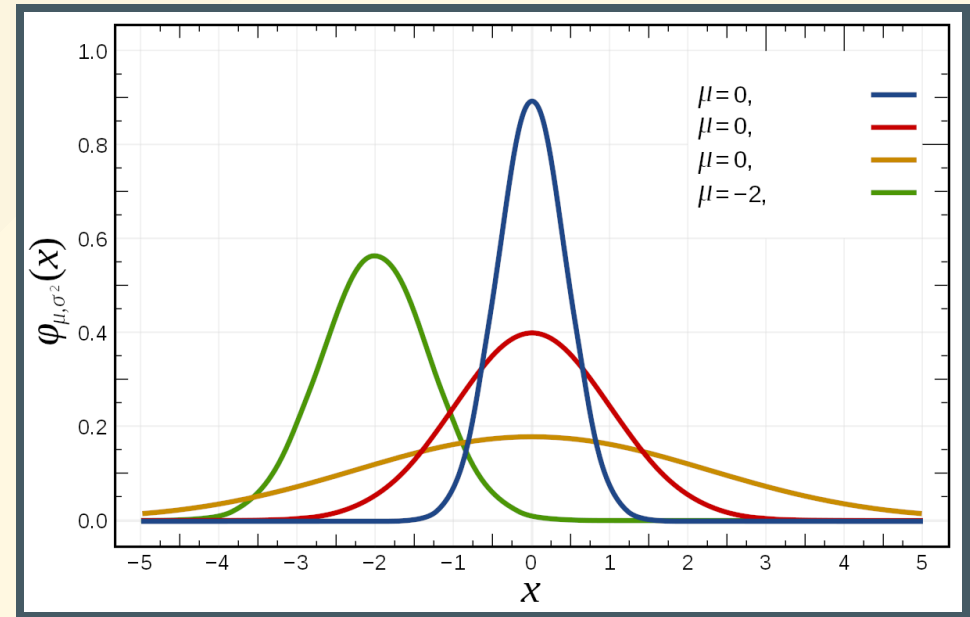
- a) Verde
- b) Blu
- c) Gialla
- d) Non lo posso sapere
- e) Nessuna delle precedenti



# Esercizio #1 -- Soluzione

? Qual è la curva con la media maggiore?

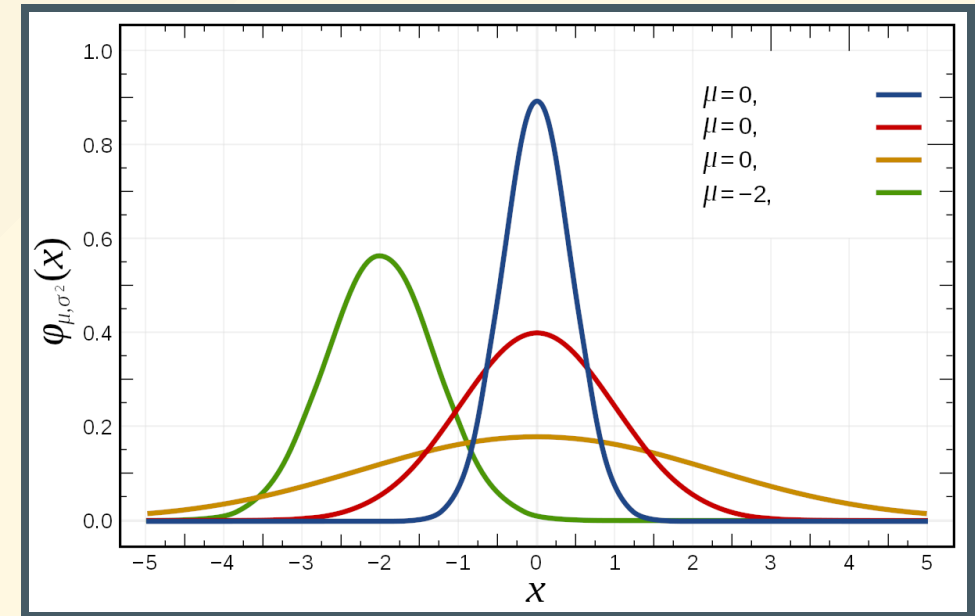
- a) Verde
- b) Blu
- c) Gialla
- d) Non lo posso sapere
- e) Nessuna delle precedenti



## Esercizio #2

? Qual è la curva con la deviazione standard maggiore?

- a) Verde
- b) Blu
- c) Gialla
- d) Non lo posso sapere
- e) Nessuna delle precedenti



## Esercizio #2 -- Soluzione

? Qual è la curva con la deviazione standard maggiore?

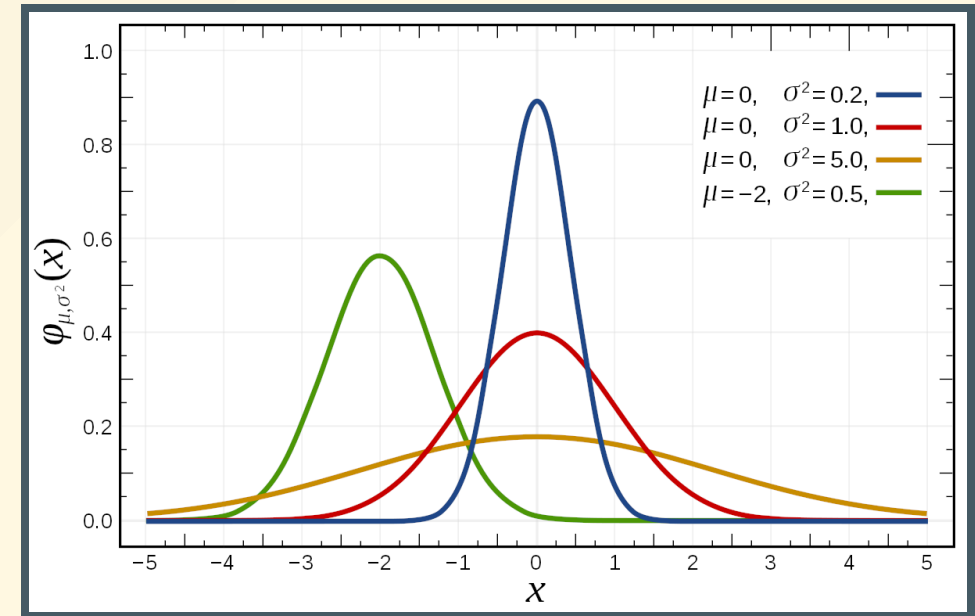
a) Verde

b) Blu

c) Gialla ☒

d) Non lo posso sapere

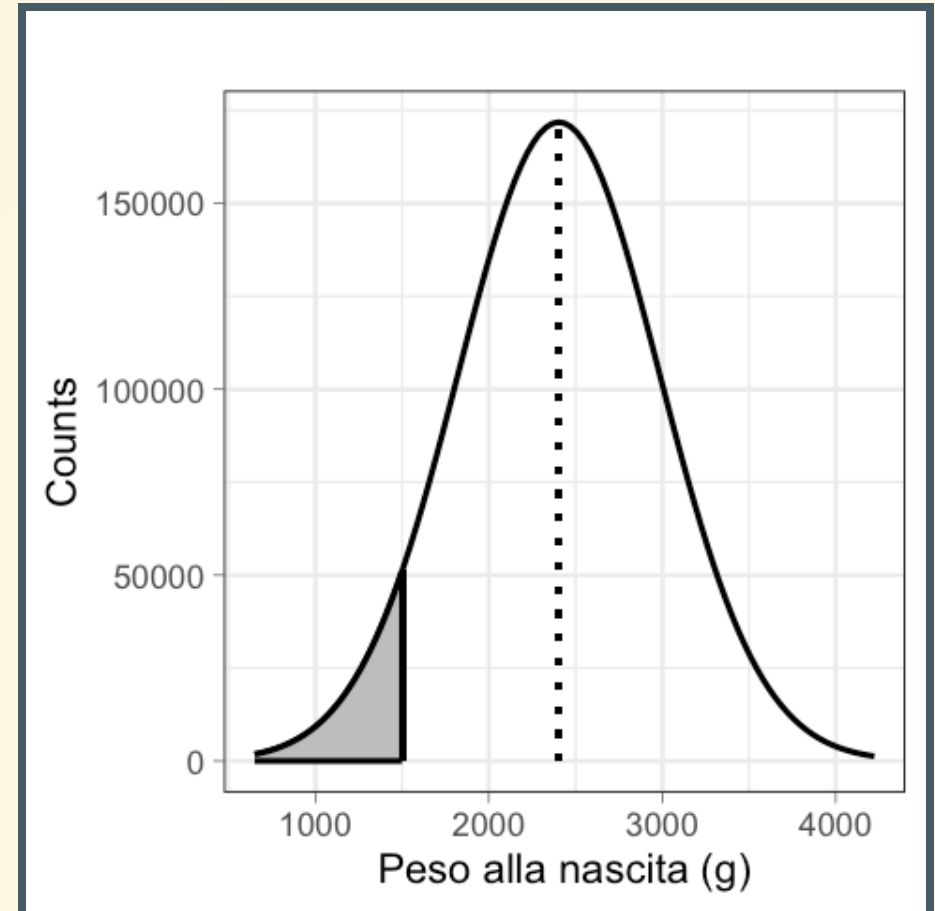
e) Nessuna delle precedenti



# La distribuzione Normale

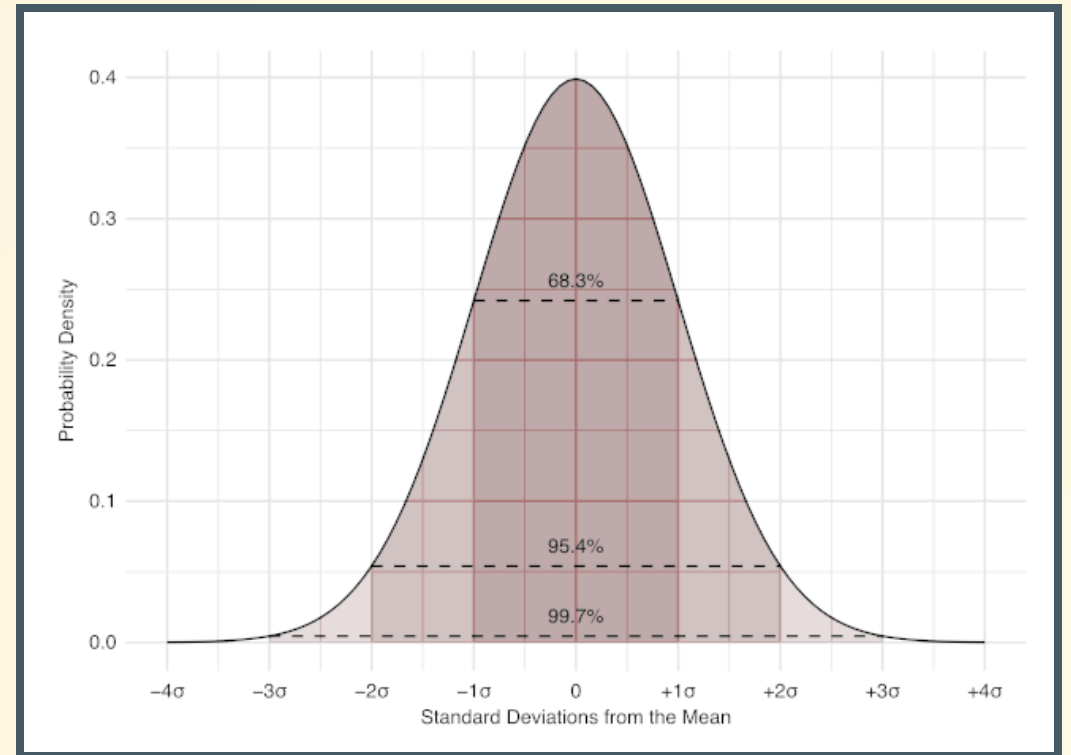
- Area sottesa alla curva = 1
- proporzione  $\equiv$  probabilità

“very low birth weight”  $< 1500$  g  
Gemelli “very low birth weight” = 6%  
 $\mathcal{P}(\text{“}\beta \text{ very low birth weight”}) = 0.06$

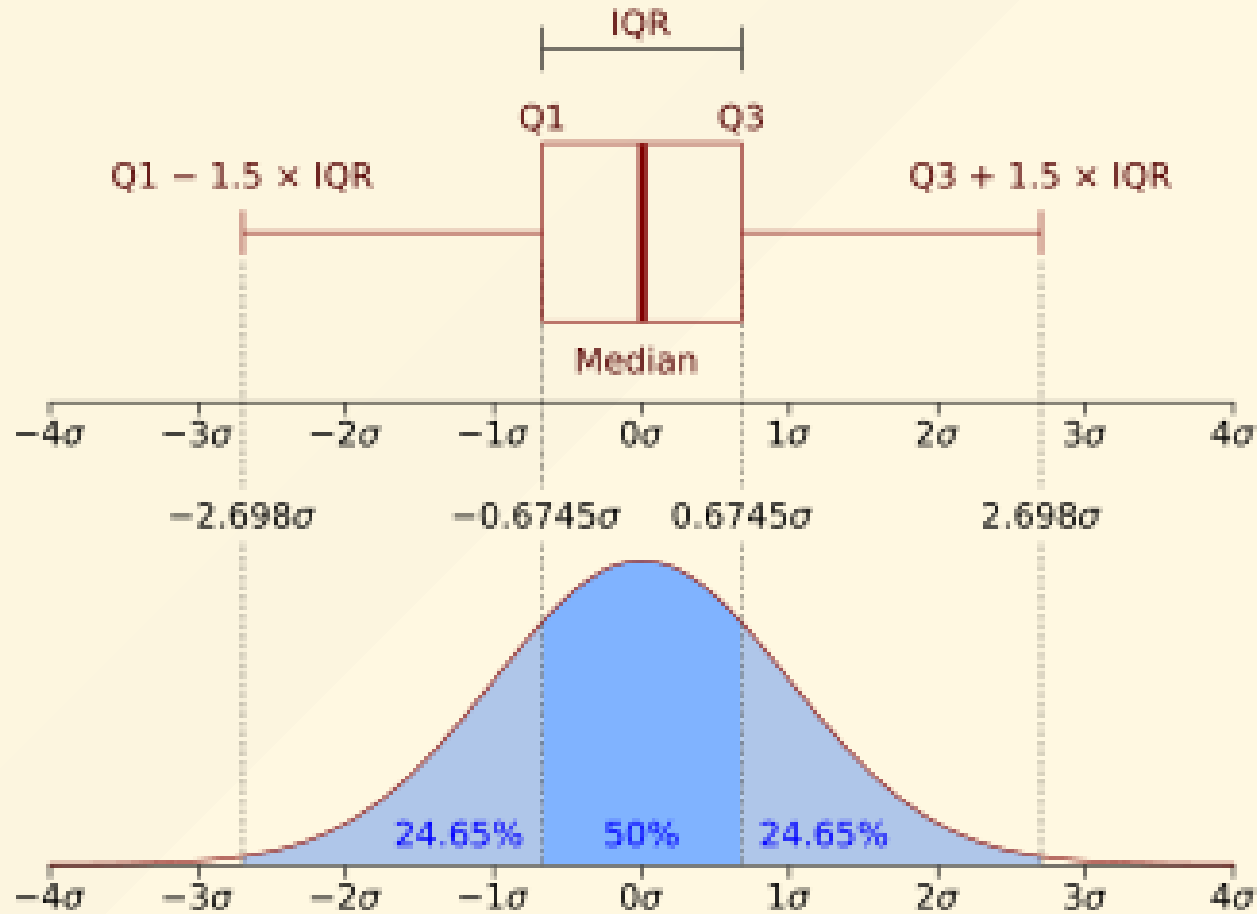


# La distribuzione Normale

- Regola del 3  $\sigma$ :
  - 68% dei valori osservati sono a 1  $\sigma$  dalla media
  - 95% sono a 2  $\sigma$
  - 99.7% sono a 3  $\sigma$
- Regola empirica:
  - valori  $< 2\sigma$  sono "*comuni*"
  - valori  $> 2\sigma$  sono "*inusuali*"
  - valori  $> 3\sigma$  sono "*estremi*"



# I valori estremi



## Esercizio #3

? L'altezza della popolazione maschile italiana si distribuisce secondo una normale con media 170 cm e deviazione standard 9.5 cm

E' possibile calcolare i seguenti valori? Se sì, quali sono?

- a) La mediana
- b) La proporzione di italiani con altezza  $> 170$  cm
- c) Le altezze inusuali
- d) L'altezza più comune
- e) L'italiano più alto di sempre



## Esercizio #3 -- Soluzione

? L'altezza della popolazione maschile italiana si distribuisce secondo una normale con media 170 cm e deviazione standard 9.5 cm

E' possibile calcolare i seguenti valori? Se sì, quali sono?

a) La mediana  $\rightarrow$  coincide con la media = 170 cm

## Esercizio #3 -- Soluzione

? L'altezza della popolazione maschile italiana si distribuisce secondo una normale con media 170 cm e deviazione standard 9.5 cm

E' possibile calcolare i seguenti valori? Se sì, quali sono?

a) La mediana  $\rightarrow$  170cm

b) La proporzione di italiani con altezza  $> 170$  cm  $\rightarrow$  sono quelli a destra della mediana, la metà dell'area sottesa dalla curva = 50%

## Esercizio #3 -- Soluzione

? L'altezza della popolazione maschile italiana si distribuisce secondo una normale con media 170 cm e deviazione standard 9.5 cm

E' possibile calcolare i seguenti valori? Se sì, quali sono?

a) La mediana  $\rightarrow$  170cm

b) La proporzione di italiani con altezza  $> 170$  cm  $\rightarrow$  50%

c) Le altezze inusuali  $\rightarrow$  sono quelle  $> 2$  deviazioni standard dalla media  
 $= 170 - 9.5 \times 2 = 151$  cm  $\wedge$   $170 + 9.5 \times 2 = 189$  cm

## Esercizio #3 -- Soluzione

? L'altezza della popolazione maschile italiana si distribuisce secondo una normale con media 170 cm e deviazione standard 9.5 cm

E' possibile calcolare i seguenti valori? Se sì, quali sono?

a) La mediana  $\rightarrow$  170cm

b) La proporzione di italiani con altezza  $> 170$  cm  $\rightarrow$  50%

c) Le altezze inusuali  $\rightarrow < 151$  cm  $\wedge > 189$  cm

d) L'altezza più comune  $\rightarrow$  è la moda, che coincide con la media e la mediana = 170 cm

## Esercizio #3 -- Soluzione

? L'altezza della popolazione maschile italiana si distribuisce secondo una normale con media 170 cm e deviazione standard 9.5 cm

E' possibile calcolare i seguenti valori? Se sì, quali sono?

a) La mediana  $\rightarrow$  170cm

b) La proporzione di italiani con altezza  $> 170$  cm  $\rightarrow$  50%

c) Le altezze inusuali  $\rightarrow < 151$  cm  $\wedge > 189$  cm

d) L'altezza più comune  $\rightarrow$  170 cm

e) L'italiano più alto di sempre  $\rightarrow$  non si può calcolare

# Esercizio #4

**Table 1. Demographic Characteristics of the Participants**

Characteristic	All Participants (N = 277)	
	Oxytocin (N = 139)	Placebo (N = 138)
Age		
Mean — yr	10.4±4.1	10.4±4.0
Distribution — no. (%)		
3–6 yr	34 (24)	35 (25)
7–11 yr	54 (39)	53 (38)
12–17 yr	51 (37)	50 (36)
Sex — no. (%)		
Male	122 (88)	120 (87)
Female	17 (12)	18 (13)



Indicativamente, in quale range di età è compreso il 70% dei pazienti nel gruppo di intervento?

- a) 3 – 17 anni
- b) 6.3 – 14.5 anni
- c) 4.1 – 16.7 anni
- d) Non è possibile desumerlo dalla tabella

# Esercizio #4 -- Soluzione

Table 1. Demographic Characteristics of the Participants		
Characteristic	All Participants (N = 277)	
	Oxytocin (N = 139)	Placebo (N = 138)
Age		
Mean — yr	10.4±4.1	10.4±4.0
Distribution — no. (%)		
3–6 yr	34 (24)	35 (25)
7–11 yr	54 (39)	53 (38)
12–17 yr	51 (37)	50 (36)
Sex — no. (%)		
Male	122 (88)	120 (87)
Female	17 (12)	18 (13)

- ? Indicativamente, in quale range di età è compreso il 70% dei pazienti nel gruppo di intervento?
- a) 3 – 17 anni
  - b) 6.3 – 14.5 anni ☒
  - c) 4.1 – 16.7 anni
  - d) Non è possibile desumerlo dalla tabella

Sikich, L. *et al.*, *Intranasal Oxytocin in Children and Adolescents with Autism Spectrum Disorder*, NEJM, 2021

## Esercizio #5

? Con quale probabilità si potrà trovare nella popolazione soggetti con valori superiori al terzo quartile?

a) 25%

b) 50%

c) 75%

d) Servono più informazioni per poter rispondere



## Esercizio #5 -- Soluzione

? Con quale probabilità si potrà trovare nella popolazione soggetti con valori superiori al terzo quartile?

a) 25% 

b) 50%

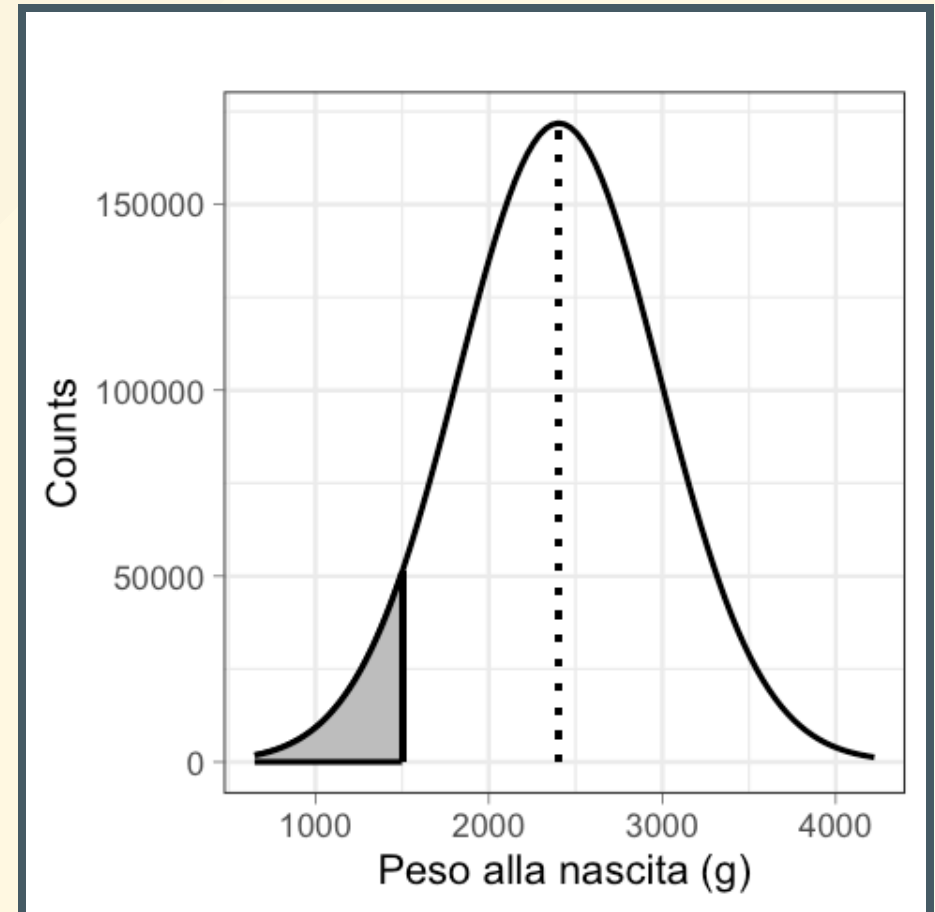
c) 75%

d) Servono più informazioni per poter rispondere

# Proporzione $\equiv$ probabilità

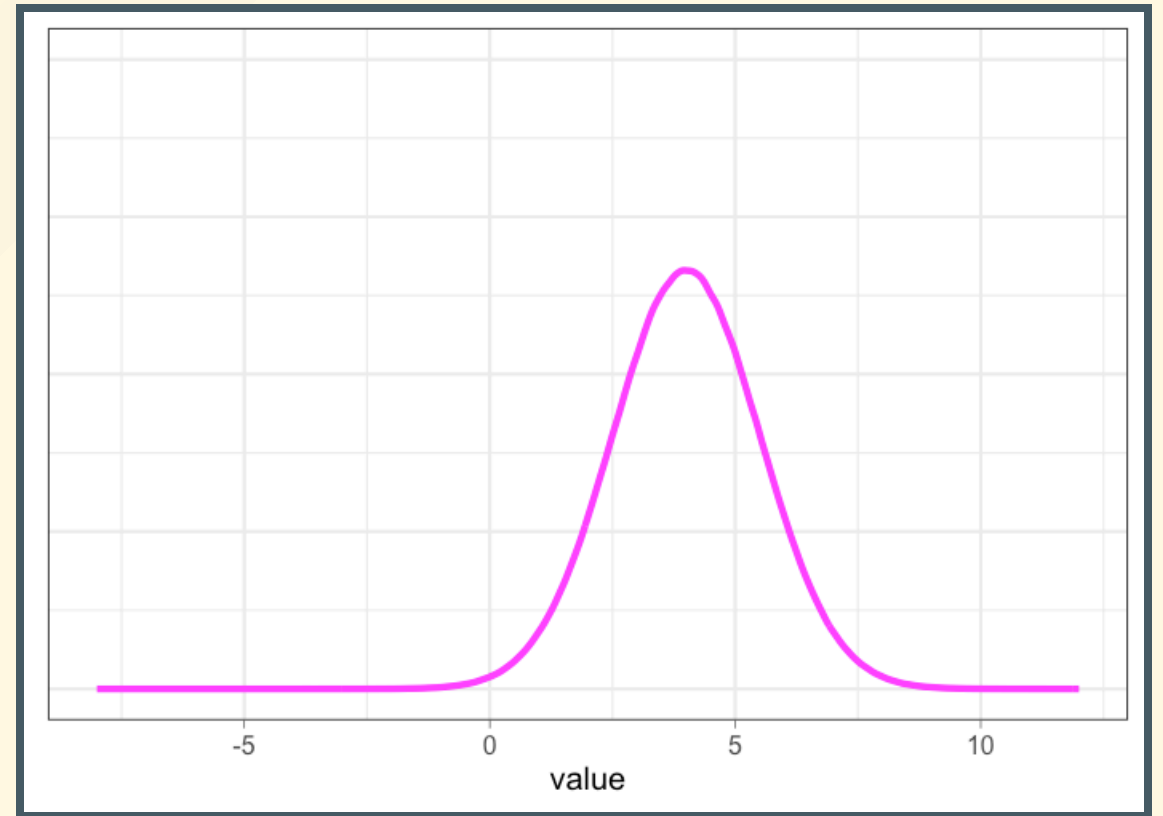
- 6% dei gemelli sono "very low birth weight"
- La probabilità essere "very low birth weight" è 0.06

Ma come è stato calcolato?



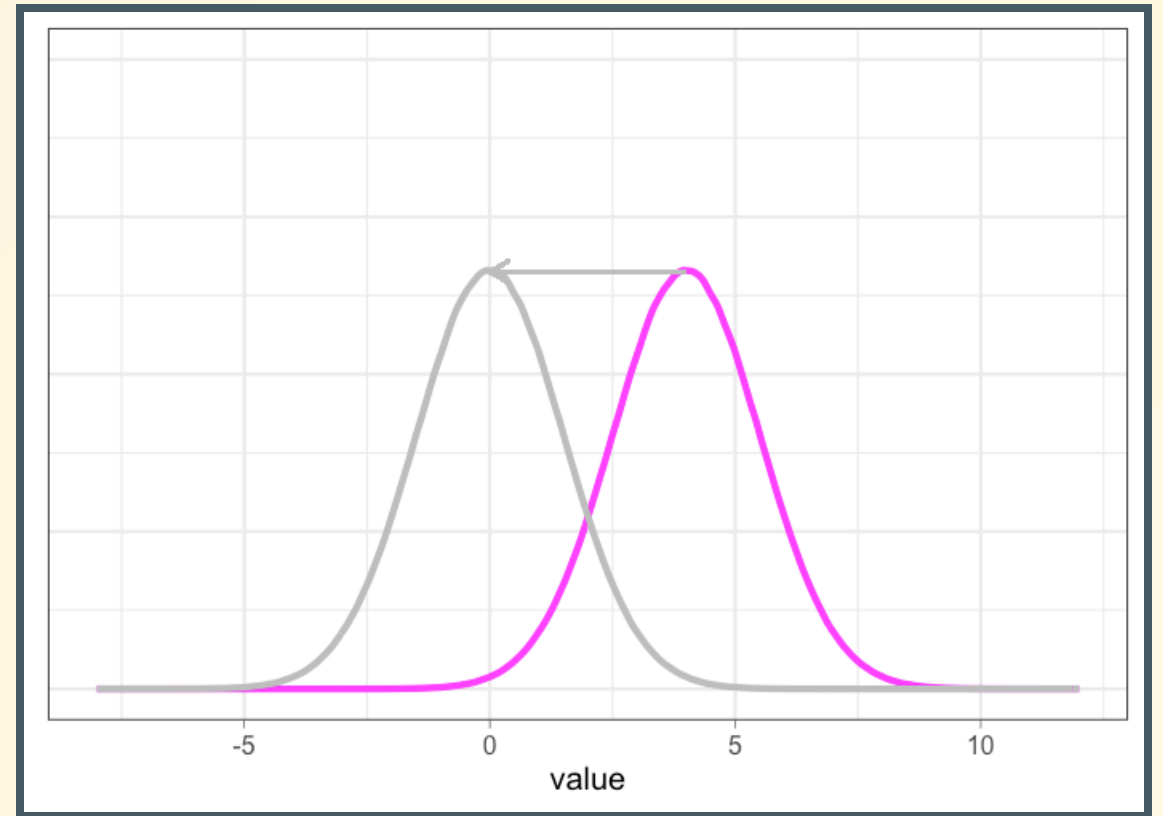
# La standardizzazione

- $\mathcal{N} = (\mu, \sigma^2) \rightarrow Z = (0, 1)$



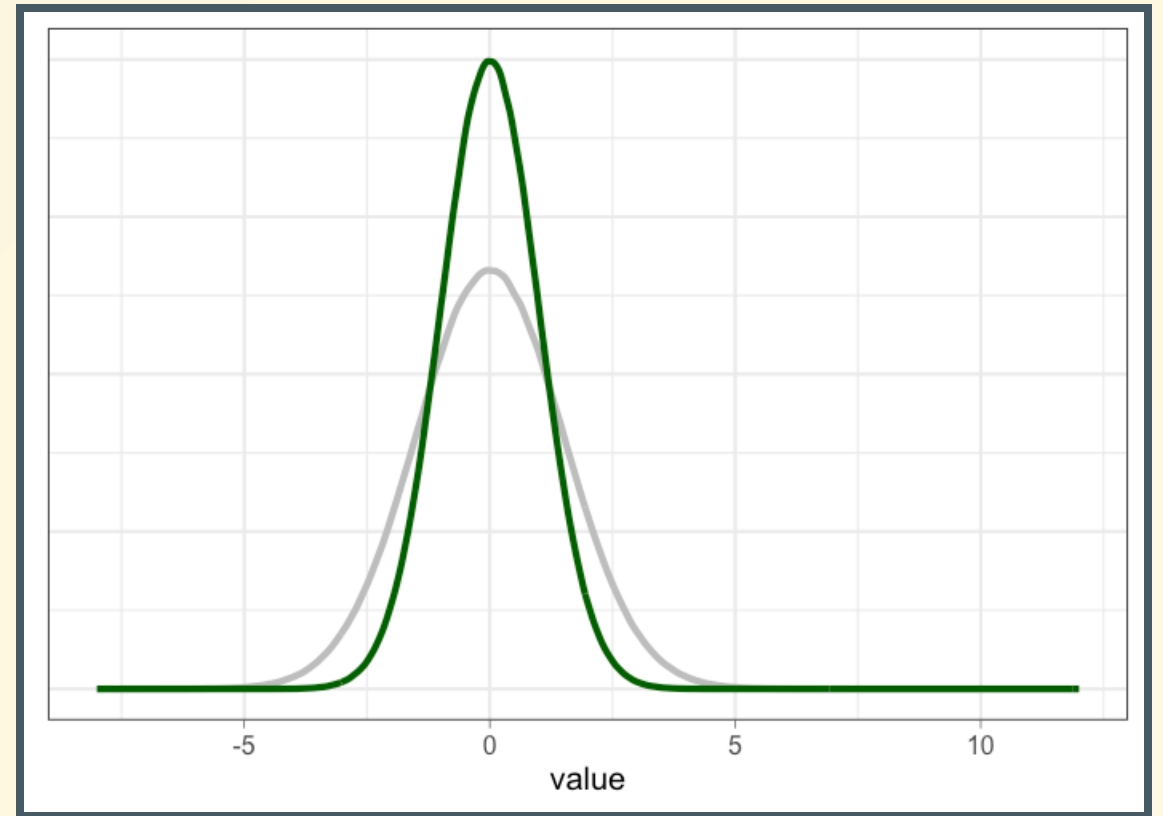
# La standardizzazione

- $\mathcal{N} = (\mu, \sigma^2) \rightarrow Z = (0, 1)$
- $z = \frac{x - \mu}{\sigma}$



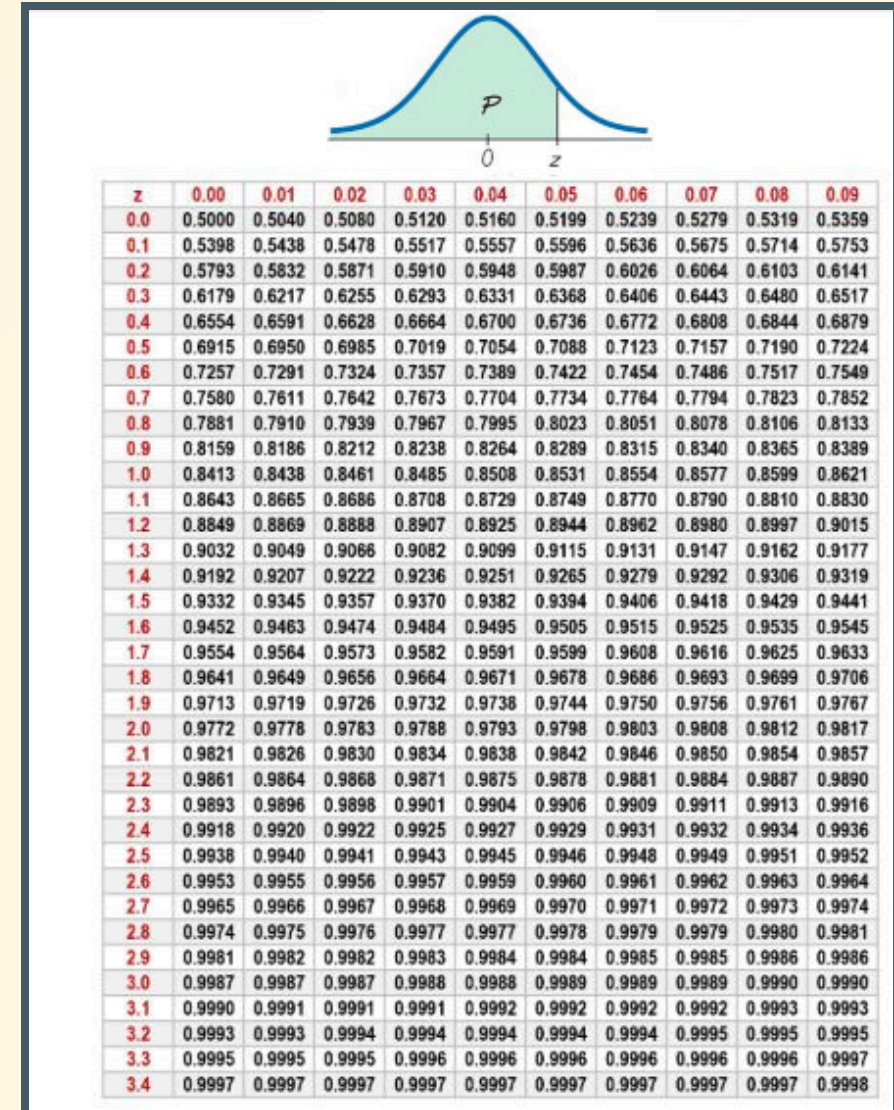
# La standardizzazione

- $\mathcal{N} = (\mu, \sigma^2) \rightarrow Z = (0, 1)$
- $z = \frac{x - \mu}{\sigma}$



# La distribuzione Normale standardizzata

- $\mathcal{N} = (\mu, \sigma^2) \rightarrow Z = (0, 1)$
- $z = \frac{x - \mu}{\sigma}$



# Calcoliamo la probabilità/proporzione

  $\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$

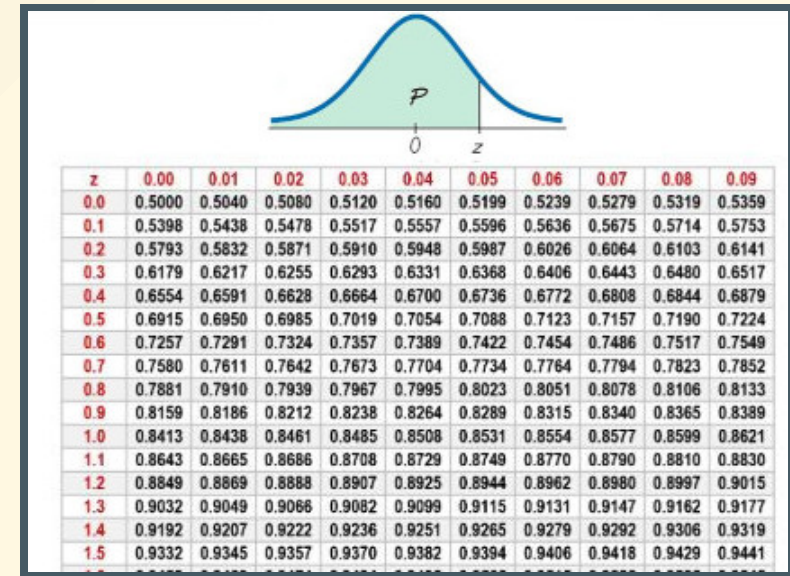
$$\mathcal{P}(x < 1500 \text{ g}) = ?$$

# Calcoliamo la probabilità/proporzione



$$\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$$

$$z = \frac{x - \mu}{\sigma} = \frac{1500 \text{ g} - 2404 \text{ g}}{580 \text{ g}} = -1.56$$



$$\mathcal{P}(x < 1500 \text{ g}) = ?$$

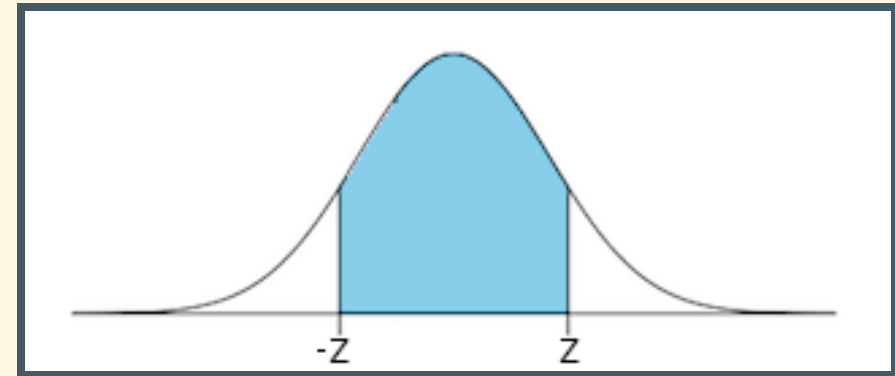


# Calcoliamo la probabilità/proporzione



$$\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$$

$$z = \frac{x - \mu}{\sigma} = \frac{1500 \text{ g} - 2404 \text{ g}}{580 \text{ g}} \\ = -1.56$$



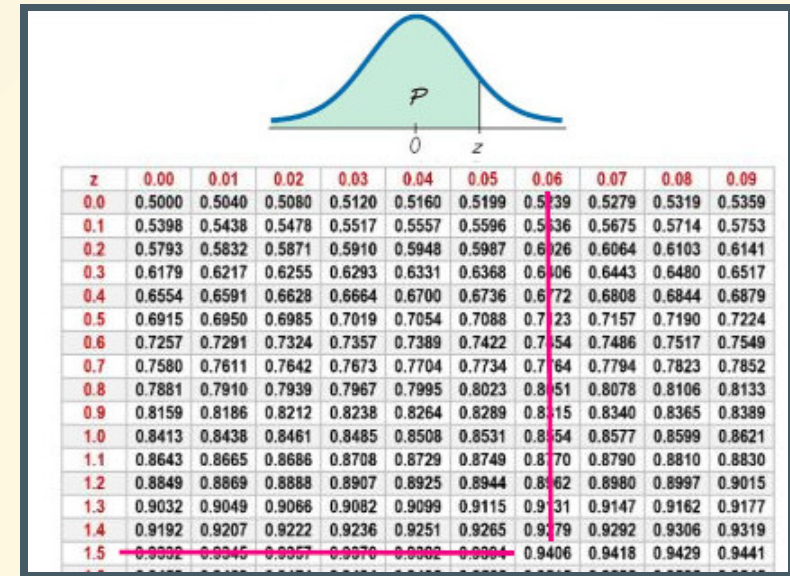
$$\mathcal{P}(x < 1500 \text{ g}) = ?$$

# Calcoliamo la probabilità/proporzione



$$\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$$

$$z = \frac{x - \mu}{\sigma} = \frac{1500 \text{ g} - 2404 \text{ g}}{580 \text{ g}} = -1.56$$

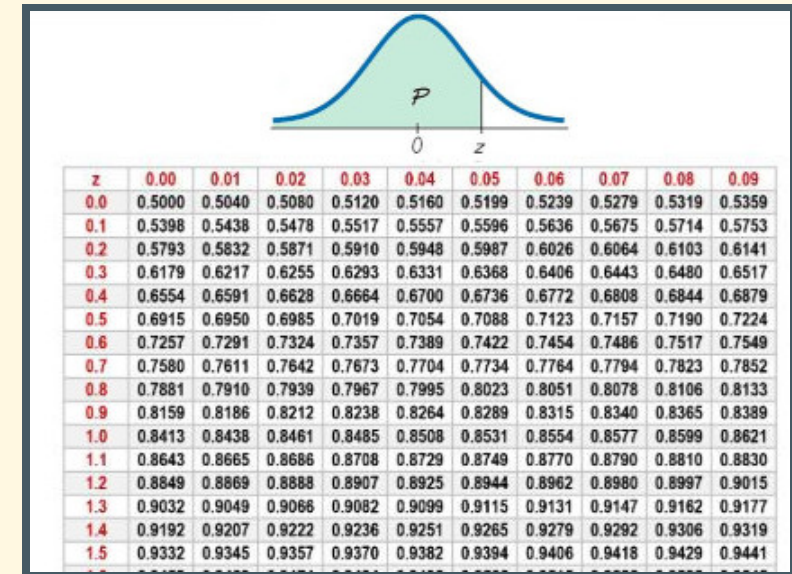


$$\mathcal{P}(x < 1500 \text{ g}) = 1 - 0.9406 = 0.0594 \rightarrow 5.94\%$$

## Esercizio #6

? Non sapendo che il bambino ha un gemello, il pediatra dice alla madre che un peso alla nascita inferiore ai 2500g è inusuale. La madre deve preoccuparsi?

$$\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$$



05:00

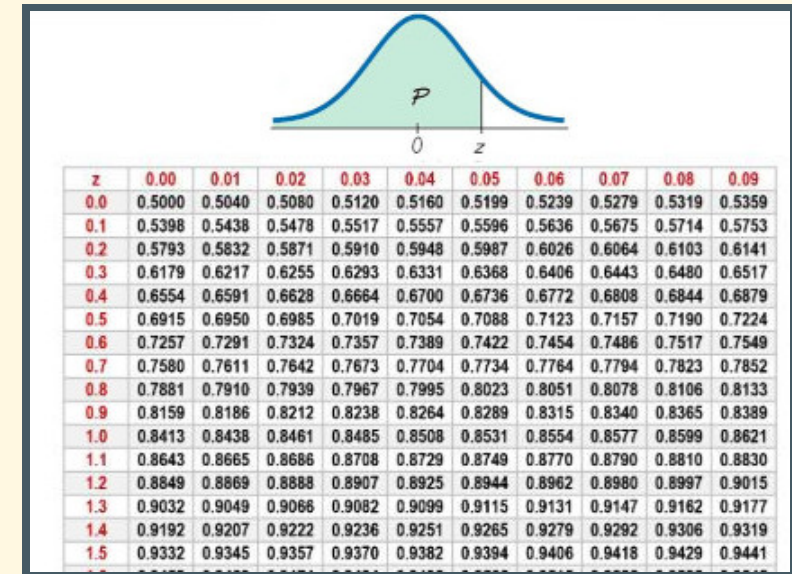
# Esercizio #6 -- Soluzione

? Non sapendo che il bambino ha un gemello, il pediatra dice alla madre che un peso alla nascita inferiore ai 2500g è inusuale. La madre deve preoccuparsi?

$$\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$$

$$z = \frac{x - \mu}{\sigma} = \frac{2500 - 2404}{580} = 0.17$$

$$\mathcal{P}(x < 2500) = 0.5675 \rightarrow 56.75\%$$



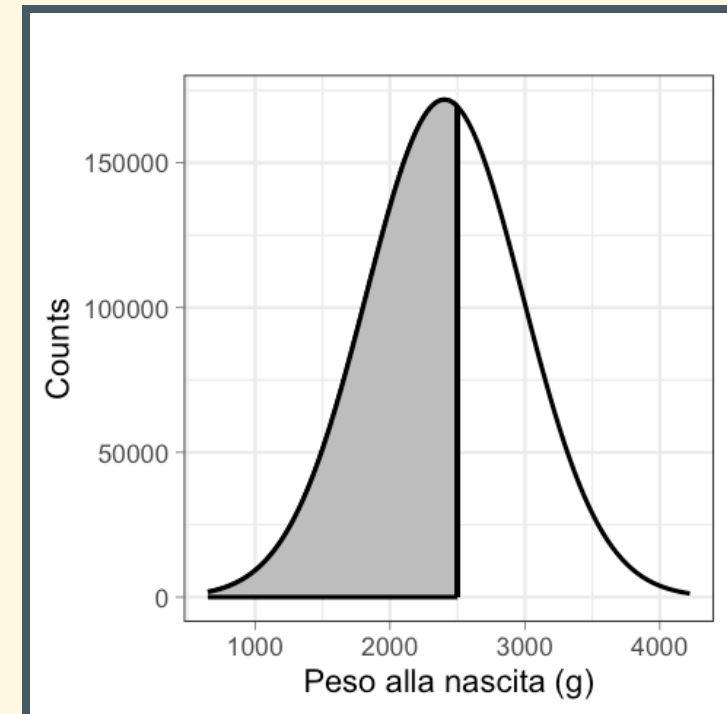
# Esercizio #6 -- Soluzione

? Non sapendo che il bambino ha un gemello, il pediatra dice alla madre che un peso alla nascita inferiore ai 2500g è inusuale. La madre deve preoccuparsi?

$$\mu = 2404 \text{ g}; \sigma = 580 \text{ g}$$

$$z = \frac{x - \mu}{\sigma} = \frac{2500 - 2404}{580} = 0.17$$

$$\mathcal{P}(x < 2500) = 0.5675 \rightarrow 56.75\%$$



## Esercizio #7

? Abbiamo una distribuzione Normale  $\mathcal{N} = (0, 1)$ . Qual è il valore della sua mediana?

- a) 0
- b) 1
- c) 2
- d) Servono più informazioni per poter rispondere

## Esercizio #7 -- Soluzione

? Abbiamo una distribuzione Normale  $\mathcal{N} = (0, 1)$ . Qual è il valore della sua mediana?

a) 0 ☒

b) 1

c) 2

d) Servono più informazioni per poter rispondere

# Cosa abbiamo imparato in questa lezione?

- La popolazione viene rappresentata con dei parametri equivalenti alle statistiche usate per i campioni
- Diversi fenomeni naturali sono normalmente distribuiti
- La normale è definita dalla sua media e deviazione standard e corrisponde a una distribuzione di probabilità
- La distribuzione (normale) di una popolazione ci fornisce la probabilità di estrarre un individuo da quella popolazione ma anche la sua frequenza nella popolazione
- Se i dati sono normalmente distribuiti, il 68% della popolazione si trova a 1 SD dalla media, il 95% a 2 SD e il 99.7% a 3 SD