

## **CURVA NORMALE O GAUSSIANA**

**(si presenta come una curva simmetrica e/o campanulare)**

*Gauss* individuò una funzione matematica parametrizzata in funzione delle grandezze che consentono di individuare UNIVOCAMENTE una distribuzione (numero delle osservazioni, valore medio e scarto quadratico medio)

**Equazione “Curva Normale” determinata da 3 parametri:**

**N** = area sotto la curva

**μ** = ascissa di massimo

**σ** = distanza in valore assoluto di ciascun punto di flesso dal punto di massimo

$$y = \frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

## CURVA NORMALE

passando dalla curva normale alla curva normale standardizzata si utilizzano le frequenze relative per cui  $N=1$  e si sostituisce all'unità di misura della distribuzione (determinata dai parametri  $\mu$  e  $\sigma$ ) una misura "standard"

ottenuta con l'impiego degli scarti standardizzati:  $z_i = \frac{x_i - \mu}{\sigma}$

$$z_i = \frac{\text{osservazione} - \text{valore medio}}{sqm}$$

Es.:

**Rossi** voto 28 esame di Politica economica  
Media classe= 25  
Sqm= 4

$$z_1 = \frac{28 - 25}{4} = 0,75$$

**Bianchi** voto 23 esame di Sociologia  
Media classe= 18  
Sqm= 5

$$z_2 = \frac{23 - 18}{5} = 1$$

Voto migliore **Rossi**.

Performance migliore **Bianchi**.

Equazione Curva Normale standardizzata:

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Sapendo che le frequenze corrispondono all'area sotto la curva e disponendo dell'equazione della curva, sia normale che standardizzata, è possibile stimare ~~(con alcune operazioni matematiche)~~ l'area sottesa ad essa in determinati intervalli (cioè la frequenza dei valori compresi tra due valori scelti).

Per una DISTRIBUZIONE NORMALE STANDARD

Alcune aree importanti sono espresse dalle frequenze dei valori compresi in intervalli simmetrici intorno alla media:

$$Fr\{\mu - \sigma \langle \bar{x} \langle \mu + \sigma\} = Fr\{-1 \langle z \langle +1\} = 0,34134 + 0,34134 = 0,68268 = 68,3 \%$$

Le frequenze comprese tra uno scarto quadratico medio al di sotto ed uno scarto quadratico medio al di sopra della curva rappresentano circa il 68% del totale delle osservazioni

$$Fr\{\mu - 2\sigma \langle \bar{x} \langle \mu + 2\sigma\} = Fr\{-2 \langle z \langle +2\} = 0,47725 + 0,47725 = 0,95450 = 95,4 \%$$

Le frequenze comprese tra due scarti quadratici medi al di sotto e due scarti quadratici medi al di sopra della curva rappresentano circa il 95% del totale delle osservazioni

$$Fr\{\mu - 3\sigma \langle \bar{x} \langle \mu + 3\sigma\} = Fr\{-3 \langle z \langle +3\} = 0,49865 + 0,49865 = 0,99730 = 99,7 \%$$

I valori standardizzati non sempre prendono valori interi (es. +1, +2,+3, etc..). Dunque per trovare l'area compresa tra il valore medio ed un qualunque valore standardizzato, dobbiamo usare le TAVOLE DELLA CURVA NORMALE (valori dell'integrale della curva normale standardizzata: TAVOLA pag.349).

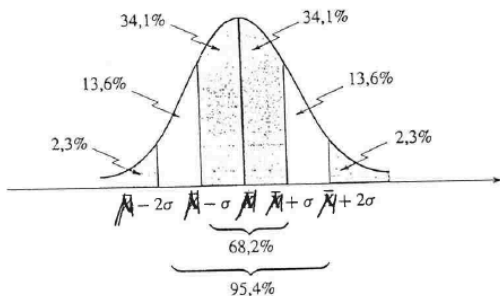


Figura 7.6 Alcune aree importanti sottese dalla curva normale.

## ESERCIZI

### 1) Area compresa tra due valori (uno inferiore e l'altro superiore alla media)

$$N=100$$

$$\mu = 176 \text{ cm}$$

$$\sigma = 4 \text{ cm}$$

Determinare area compresa tra  $x_1 \div x_2 = 174 \div 177$

PROCEDIMENTO:

$$z_1 = \frac{174 - 176}{4} = -0,5$$

$$z_2 = \frac{177 - 176}{4} = 0,25$$

*cercando i corrispettivi valori sulla Tavola B che riporta i valori dell'integrale della curva normale standardizzata, avremo:*

$$P_{(z_1)} = 0,1915$$

*(non si troveranno valori negativi data la simmetria della distribuzione)*

$$P_{(z_2)} = 0,0987$$

Concludendo, l'area compresa tra 174 e 177 cm sarà data dalla somma di queste due porzioni (perché si trovano una al di sotto e l'altra al di sopra del valore medio), è cioè:

$$P(z_1 \div z_2) = 0,1915 + 0,0987 = 0,2902$$

Il valore ottenuto si può moltiplicare per 100 per conoscere la percentuale di osservazioni comprese nell'intervallo tra 174 e 177 che rappresentano, quindi, il 29%.

**2) Area compresa tra due valori (a sinistra del valore medio)**

$$N=100$$

$$\mu = 26$$

$$\sigma = 3$$

Determinare area compresa tra  $\chi_1 \div \chi_2 = 20 \div 24$

PROCEDIMENTO:

$$z_1 = \frac{20 - 26}{3} = -2$$

$$z_2 = \frac{24 - 26}{3} = -0,66$$

*cercando i corrispettivi valori sulla Tavola B (pag.349) che riporta i valori dell'integrale della curva normale standardizzata, avremo:*

$$P_{(z_1)} = 0,4772$$

$$P_{(z_2)} = 0,2454$$

*(non si troveranno valori negativi data la simmetria della distribuzione)*

Concludendo, l'area compresa tra i voti 20 e 24 sarà data dalla differenza di queste due porzioni (perché si trovano ambedue al di sotto del valore medio), è cioè:

$$P(z_1 \div z_2) = 0,4772 - 0,2454 = 0,2318$$

Il valore ottenuto si può moltiplicare per 100 per conoscere la percentuale di osservazioni comprese nell'intervallo tra i voti 20 e 24 che racchiudono, quindi, il 23% dei voti presi dagli studenti osservati.

**3) Area residuale al di sotto o al di sopra di determinati valori (a destra o a sinistra del valore medio)**

$$N=100$$

$$\mu = 176$$

$$\sigma = 4$$

Determinare area occupata dalle osservazioni:  
al di sotto di 173 cm ovvero  $x_i < 173$

PROCEDIMENTO:

$$z_2 = \frac{173 - 176}{4} = -0,75$$

*cercando i corrispettivi valori sulla Tavola B che riporta i valori dell'integrale della curva normale standardizzata, avremo:*

$$P_{(z_2)} = 0,27337$$

*(non si troveranno valori negativi data la simmetria della distribuzione)*

e l'area al di sotto dei 173 cm sarà data dalla differenza tra 0,5 (che corrisponde alla metà dell'area della curva) ed il valore di  $P_{(z_2)}$  e cioè:

$$0,5 - P_{(z_2)} = 0,5 - 0,2734 = 0,226 * 100 = 22 \%$$

**4) Area compresa tra un certo valore al di sotto del valore medio in su**

$N=100$

$\mu = 142$

$\sigma = 5$

Determinare area occupata dalle osservazioni al di sopra di 130 Km, ovvero  $x_i > 130$

PROCEDIMENTO:

$$z = \frac{130 - 142}{5} = -2,4$$

*cercando il corrispettivo valore sulla Tavola B che riporta i valori dell'integrale della curva normale standardizzata, avremo:*

$$P_{(z)} = 0,4918$$

Concludendo, l'area tra i 130 km in su sarà data da 0,5 (che corrisponde alla metà dell'area della curva) cui si somma il valore di  $P_{(z)}$ , ovvero tra 130 ed il valore medio e cioè:

$$0,5 + P_{(z)} = 0,5 + 0,4918 = 0,991$$

Il valore ottenuto si può moltiplicare per 100 per conoscere la percentuale di osservazioni al di sopra del valore 130, che rappresentano, quindi, il 99% degli automobilisti osservati.