

# *Università degli Studi di Napoli “Parthenope”*

*Corso di Laurea in Statistica e Informatica per l'Azienda, la Finanza e le  
Assicurazioni (SIAFA)*

## STATISTICA II MODULO

*Sergio LONGOBARDI*  
[longobardi@uniparthenope.it](mailto:longobardi@uniparthenope.it)

# PROGRAMMA

Il programma dell'insegnamento può essere suddiviso in più blocchi di lezioni:

**I blocco:** Campionamento e distribuzioni campionaria.

**II blocco:** Stima puntuale. Stima per intervallo. Teoria dei test statistici. Inferenza su medie, proporzioni, varianze.

**III blocco:** Il modello di regressione lineare semplice. Inferenza nel modello di regressione.

# Materiale didattico

## ***Libri di testo***

Borra S., Di Ciaccio A. (ediz. 2008 o ediz. 2014)

***Statistica***

***Metodologie per le scienze economiche e sociali***

McGraw-Hill

## ***Materiale online***

- **Slides delle lezioni**
- **Esercizi da svolgere**

# ESAME

- L'esame consiste in una prova scritta e una prova orale.
- La prova scritta, da completare entro un'ora, è costituita da un set di esercizi, ognuno articolato su più punti. Per la prova scritta è necessario l'utilizzo della calcolatrice.
- La prova orale riguarda gli argomenti del programma trattati nelle lezioni.

# RICEVIMENTO

*Martedì ore 12:00-13:00*

*Online su Teams canale:  
Z2d6134*

*In presenza:  
stanza 449/a*

*[sergio.longobardi@uniparthenope.it](mailto:sergio.longobardi@uniparthenope.it)*

## Variabili casuali

*Quantità numerica il cui valore non è noto a priori ma è determinato dal caso (cioè è legato al risultato di un esperimento casuale)*

Esempi:

- Il numero di ingressi in un museo in un dato giorno
- Il numero di clienti in un hotel
- Il risultato del lancio di due dadi

***Una variabile casuale  $X$  e' una funzione definita sullo spazio campionario  $\Omega$  che associa ad ogni evento  $E \subset \Omega$  un unico numero reale.***

## Variabili casuali

**Variabile casuale**

***discreta*** può assumere un insieme discreto (finito o numerabile) di numeri reali


































***continua*** può assumere tutti i valori compresi in un intervallo reale

$\Omega$  discreto  V.C. discreta

$\Omega$  continuo  V.C. continua

# Variabili casuali discrete

Lanciando due dadi si può osservare la variabile casuale discreta  $X$ : “somma dei due punteggi”. I risultati possibili (eventi elementari) sono le 36 diverse combinazioni delle facce dei due dadi. Ma i valori che può assumere la v.c. sono 11

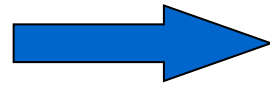
		.....	.....	.....	.....	.....	→ $X = 2$
			.....	.....	.....	.....	→ $X = 3$
				.....	.....	.....	→ $X = 4$
					.....	.....	→ $X = 5$
						.....	→ $X = 6$
$\Omega$							→ $X = 7$
						.....	→ $X = 8$
				.....	.....	.....	→ $X = 9$
			.....	.....	.....	.....	→ $X = 10$
		.....	.....	.....	.....	.....	→ $X = 11$
		.....	.....	.....	.....	.....	→ $X = 12$



# La funzione di probabilità

*La funzione di probabilità di una variabile casuale discreta  $X$  associa ad ognuno dei valori  $x_i$  la corrispondente probabilità  $P(X=x_i)$*

$$P(X=x_i)$$



*Probabilità che la v.c.  $X$  assuma il valore  $x_i$*

Proprietà

$$\sum_i P(x_i) = 1$$

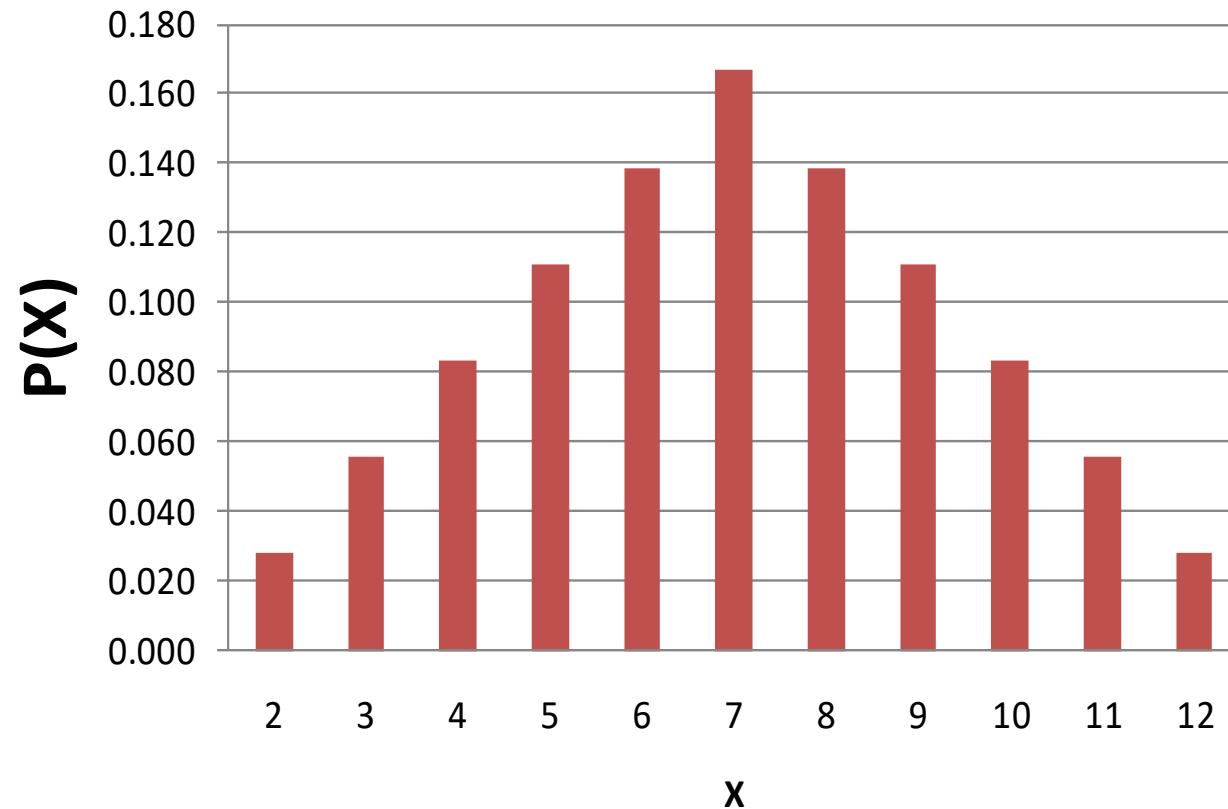
$$P(x_i) \geq 0$$

# Variabili casuali

## *Funzione di probabilità*

$x_i$	$P(X=x_i)$
2	$1/36 = 0.028$
3	$2/36 = 0.056$
4	$3/36 = 0.083$
5	$4/36 = 0.111$
6	$5/36 = 0.139$
7	$6/36 = 0.167$
8	$5/36 = 0.139$
9	$4/36 = 0.111$
10	$3/36 = 0.083$
11	$2/36 = 0.056$
12	$1/36 = 0.028$

## Rapp.grafica funzione di probabilità



# Funzione di Ripartizione

Probabilità che la v.c. assume uno specifico valore  $\longrightarrow$  Funz. di probabilità

Probabilità che la v.c. assume un valore inferiore o uguale ad un valore  $x_i$   $\longrightarrow$  Funz. di ripartizione

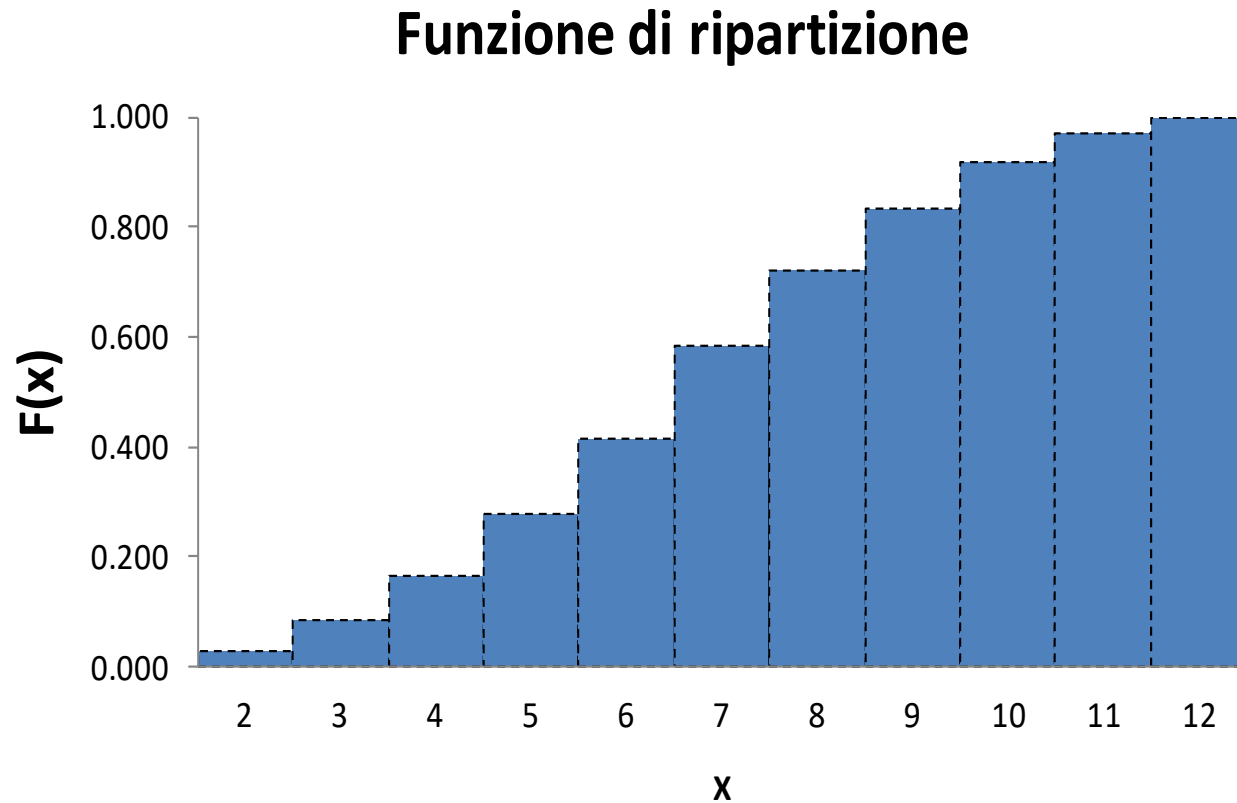
Data una v.c discreta  $\mathbf{X}$ , la funzione che fa corrispondere ai valori  $\mathbf{x}$  le probabilità cumulate viene detta **funzione di ripartizione** ed indicata con:

$$F(x) = P(X \leq x) = \sum_{w \leq x} P(X = w)$$

# Variabili casuali

## *Funzione di ripartizione*

$x_i$	$F(x)$
2	0.028
3	0.083
4	0.167
5	0.278
6	0.417
7	0.583
8	0.722
9	0.833
10	0.917
11	0.972
12	1.000



## Valori caratteristici di una v.c. discreta

Il valore atteso di una v.c. discreta  $X$  è definito come

$$E(X) = \sum_i x_i P(x_i)$$

La varianza di una v.c. discreta  $X$  è definito come

$$V(X) = \sum_i (x_i - E(X))^2 P(x_i)$$

# Esempio: calcolo del valore atteso

<b>x<sub>i</sub></b>	<b>P(X=x<sub>i</sub>)</b>
2	1/36= 0.028
3	2/36= 0.056
4	3/36= 0.083
5	4/36= 0.111
6	5/36= 0.139
7	6/36= 0.167
8	6/36= 0.139
9	4/36= 0.111
10	3/36= 0.083
11	2/36= 0.056
12	1/36= 0.028

$$E(X) = \sum_i x_i P(x_i)$$

$$E(X) = 2 \frac{1}{36} + 3 \frac{2}{36} + 4 \frac{3}{36} + 5 \frac{4}{36} + 6 \frac{5}{36} + 7 \frac{6}{36} + 8 \frac{5}{36} + 9 \frac{4}{36} \\ + 10 \frac{3}{36} + 11 \frac{2}{36} + 12 \frac{1}{36} = 7$$

## Variabili casuali continue

***Le v.c. continue*** possono assumere tutti i valori (infiniti) compresi in un intervallo reale

***Non è possibile come nel caso delle v.c. discrete associare una probabilità ad ognuno degli infiniti valori che la v.c. può assumere***

***Di conseguenza si assegna una misura di probabilità a tutti i possibili intervalli sull'asse reale mediante la funzione di densità***

## Variabili casuali continue

Chiameremo **Funzione di densità**, la funzione matematica  $f(x)$  per cui l'area sottesa alla funzione, corrispondente ad un certo intervallo, è uguale alla probabilità che  $X$  assuma un valore in quell'intervallo.

Mediante la funzione di densità è quindi possibile calcolare la probabilità che la v.c.  $X$  assuma valori compresi tra  $a$  e  $b$  è:

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$



## Variabili casuali continue

È possibile, per esempio, calcolare la probabilità che la v.c.  $X$  con funzione di densità:

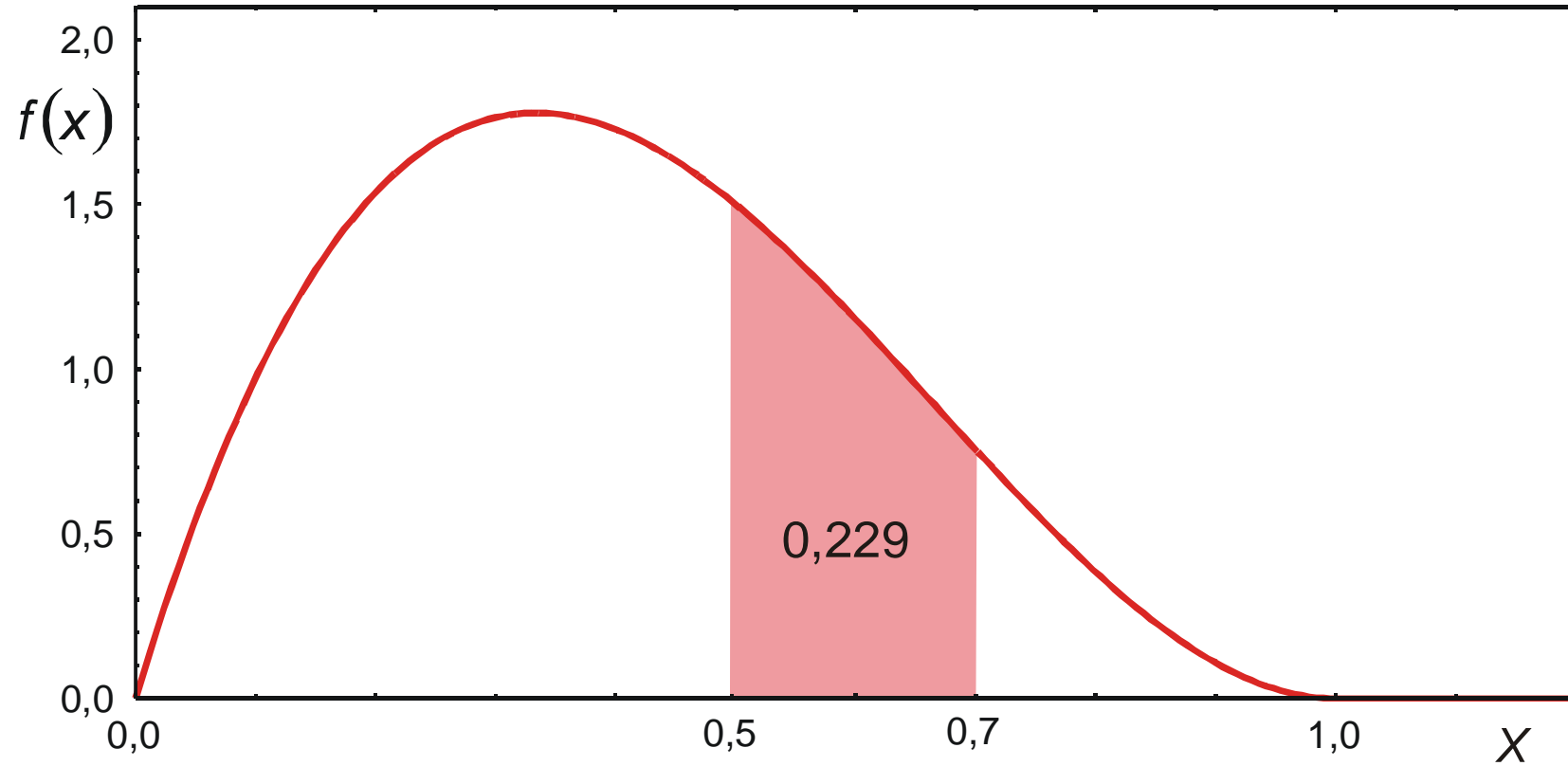
$$f(x) = 12x(1-x)^2$$

assuma un valore nell'intervallo  $[0,5 ; 0,7]$

$$P(0,5 \leq X \leq 0,7) = \int_{0,5}^{0,7} f(x) dx$$

$$\int_{0,5}^{0,7} 12x(1-x)^2 dx = 0,29$$

# Variabili casuali continue



# Funzioni di densità

## Proprietà

**$f(x) \geq 0$**  sempre

**L'area totale sottesa alla  
funzione = 1, ossia**

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

**La probabilità che la v.c.  
assuma un particolare valore  
dell'intervallo è zero.**

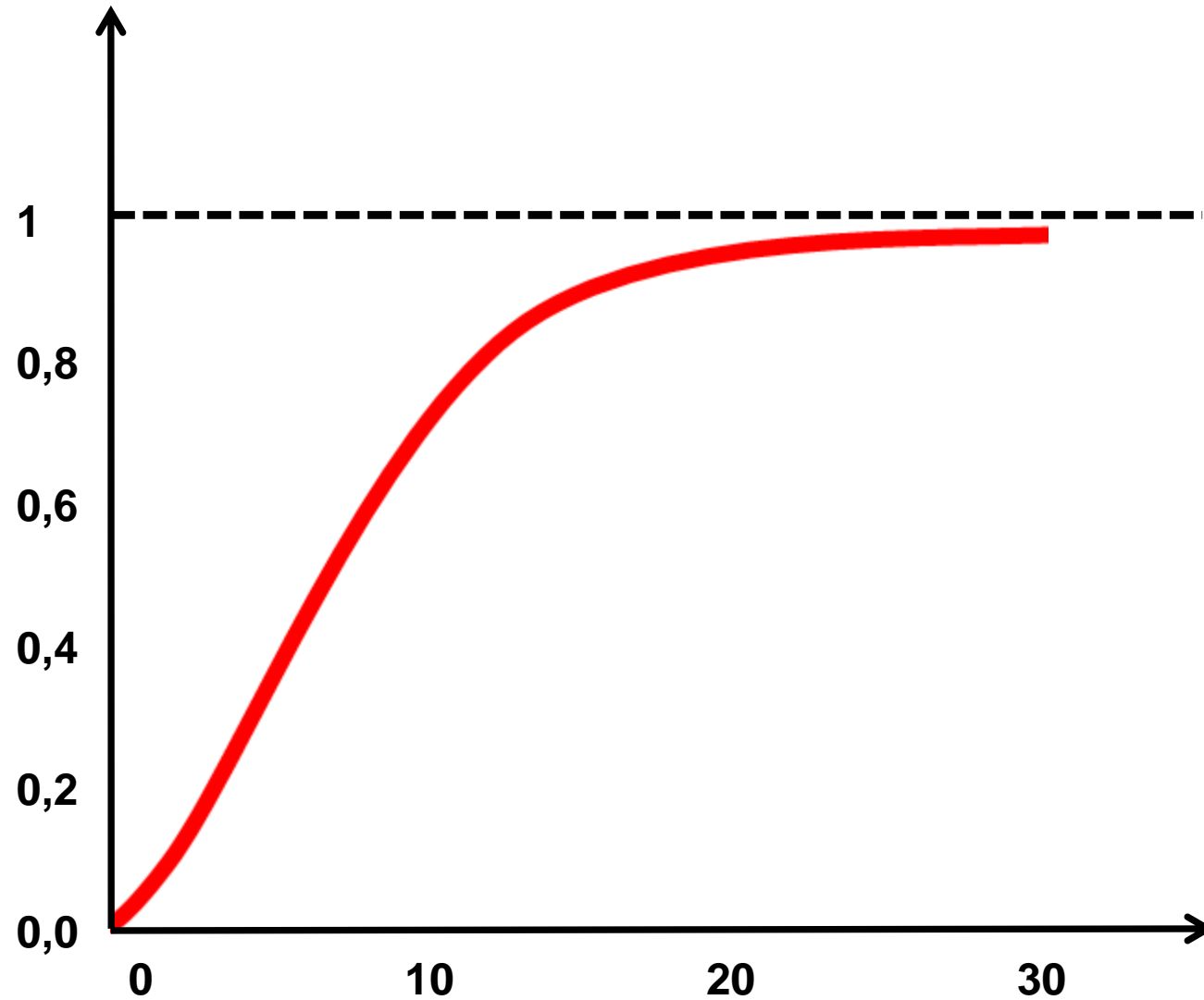
# Funzione di Ripartizione

Data una v.c. continua  $X$ , la funzione che fa corrispondere ai valori  $x$  le probabilità cumulate  $P(X \leq x)$  viene detta funzione di ripartizione.

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(w)dw$$

# Variabili casuali continue

Funzione di ripartizione della v.c X reddito annuo



## Valori caratteristici di una v.c. continua

Il valore atteso di una v.c. continua  $X$  è definito come

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

La varianza di una v.c. continua  $X$  è definito come

$$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx$$

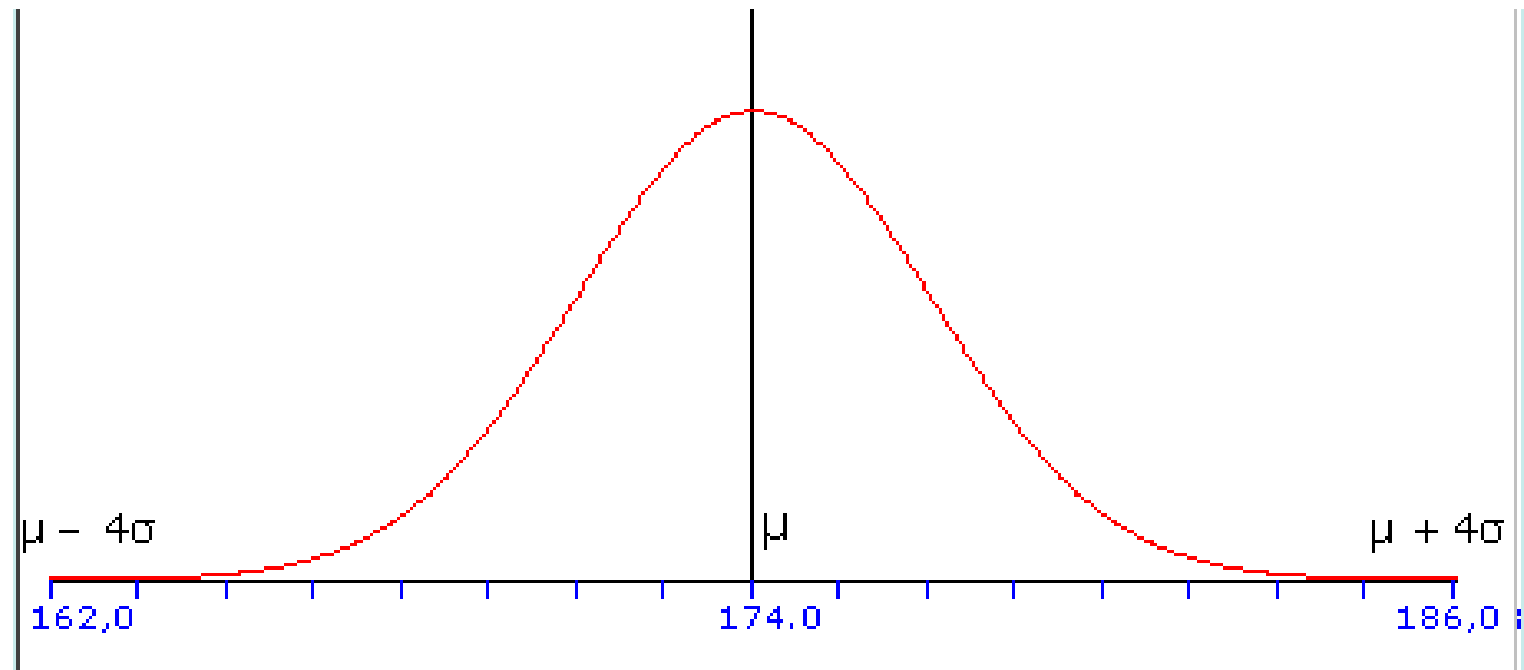
# ***La variabile casuale normale***

# Variabile casuale normale

La maggior parte dei fenomeni reali si approssima in modo adeguato con una particolare funzione di densità denominata *funzione di densità Normale*.

La v.c. Normale venne sviluppata nel 1733 da De Moivre e successivamente è stata resa “famosa” da Gauss per l'analisi di dati astronomici tanto da essere anche denominata come v.c. Gaussiana.

Esempio di v.c.  
normale:  
altezza di un  
collettivo di  
studenti di sesso  
maschile





## Variabile casuale normale

La v.c. Normale  $X$ , è una v.c. continua che può assumere valori su tutto l'asse reale.

La funzione di densità è data da:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$-\infty < X < +\infty$$

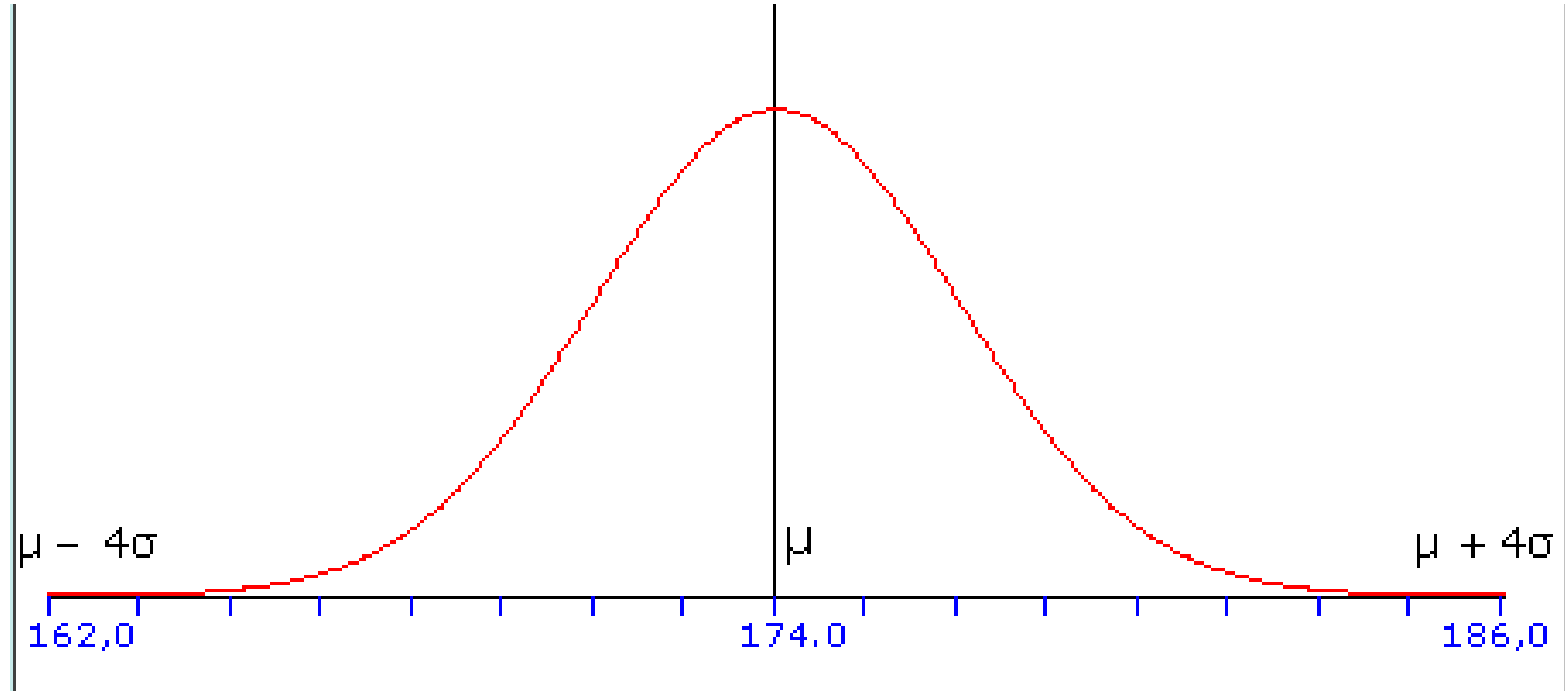
$$-\infty < \mu < +\infty$$

$$\sigma^2 > 0$$

La v.c. Normale  $X$  si indica con  $X \sim N(\mu, \sigma^2)$  in quanto i parametri che la caratterizzano sono 2:

- 1) il valore atteso  $E(X) = \mu$
- 2) la varianza  $V(X) = \sigma^2$

# Variabile casuale normale



## Caratteristiche della Normale

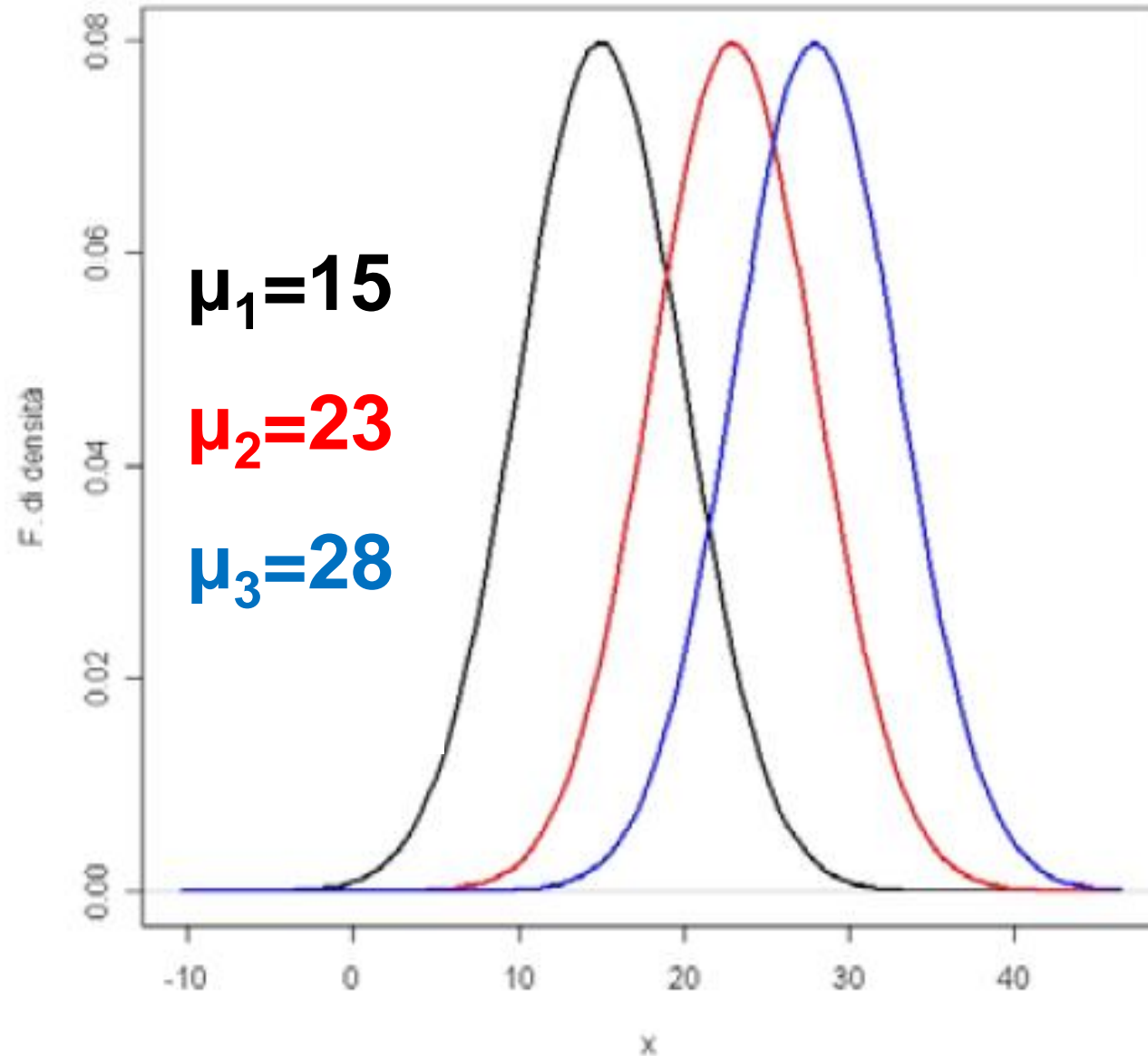
- Forma campanulare, unimodale e simmetrica
- Media=Mediana=Moda= $\mu$
- Punti di flesso a distanza  $\sigma$  dalla media ( $\mu - \sigma$  e  $\mu + \sigma$ )

$$\lim_{x \rightarrow +\infty} f(x) = 0$$

$$\lim_{x \rightarrow -\infty} f(x) = 0$$

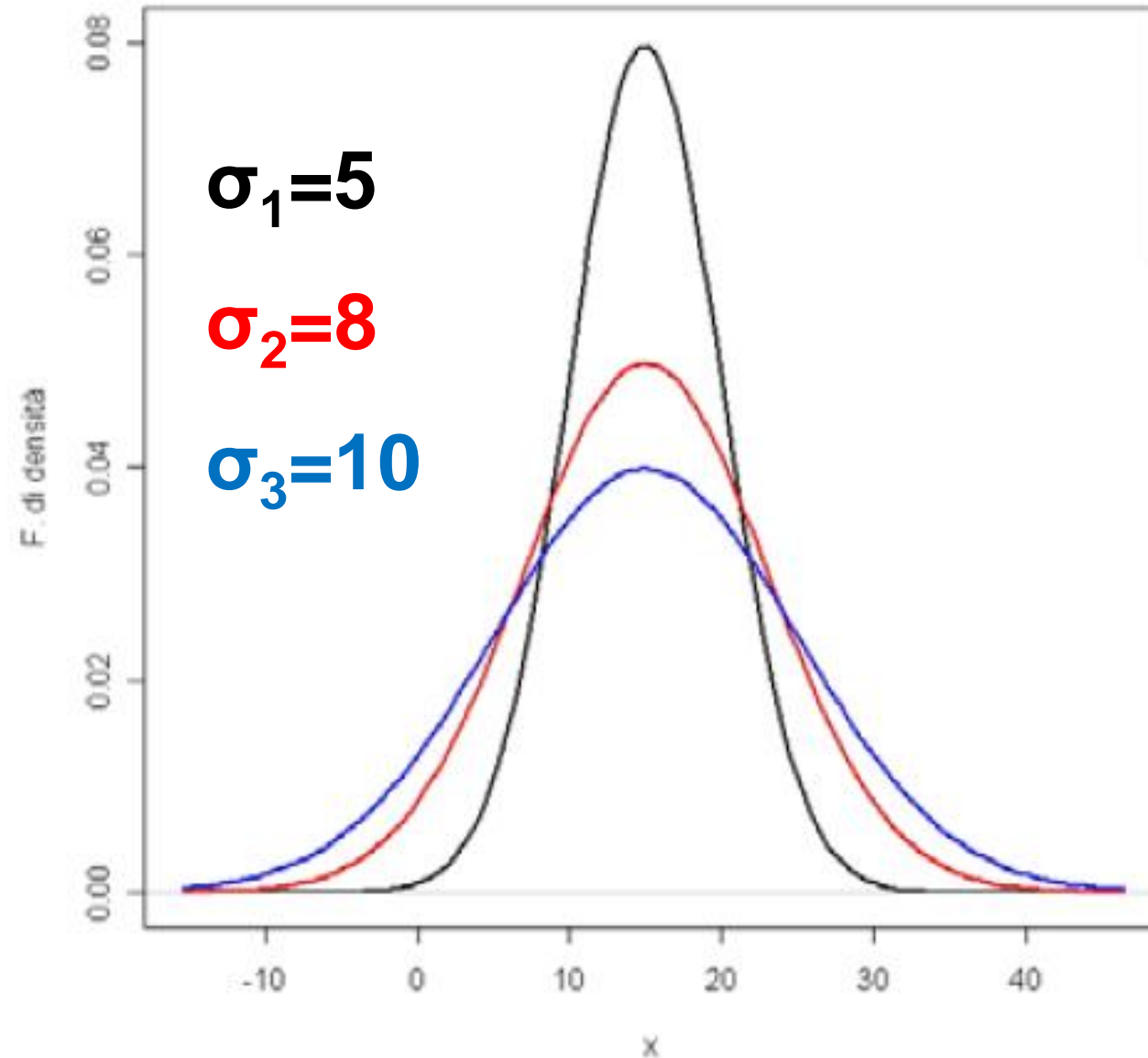
## Variabile casuale normale

La media  $\mu$   
determina la  
posizione della  
curva sull'asse  
delle ascisse



## Variabile casuale normale

La deviazione standard  $\sigma$  determina la dispersione della curva intorno al valore medio



# Variabile casuale normale

## Proprietà

- *Ogni trasformazione lineare di una v.c. Normale è ancora una v.c. Normale*
- *La somma di due v.c. Normali indipendenti è ancora una v.c. Normale con media e varianza pari, rispettivamente, alla somma delle medie e delle varianze delle due v.c. Normali.*

## Variabile casuale normale standardizzata

Applicando la seguente trasformazione lineare alla v.c. Normale  $X \sim N(\mu, \sigma^2)$  :

$$Z = \frac{X - \mu}{\sigma}$$

Si ottiene la *v.c. normale standardizzata*  $Z$  caratterizzata da media nulla e varianza unitaria:

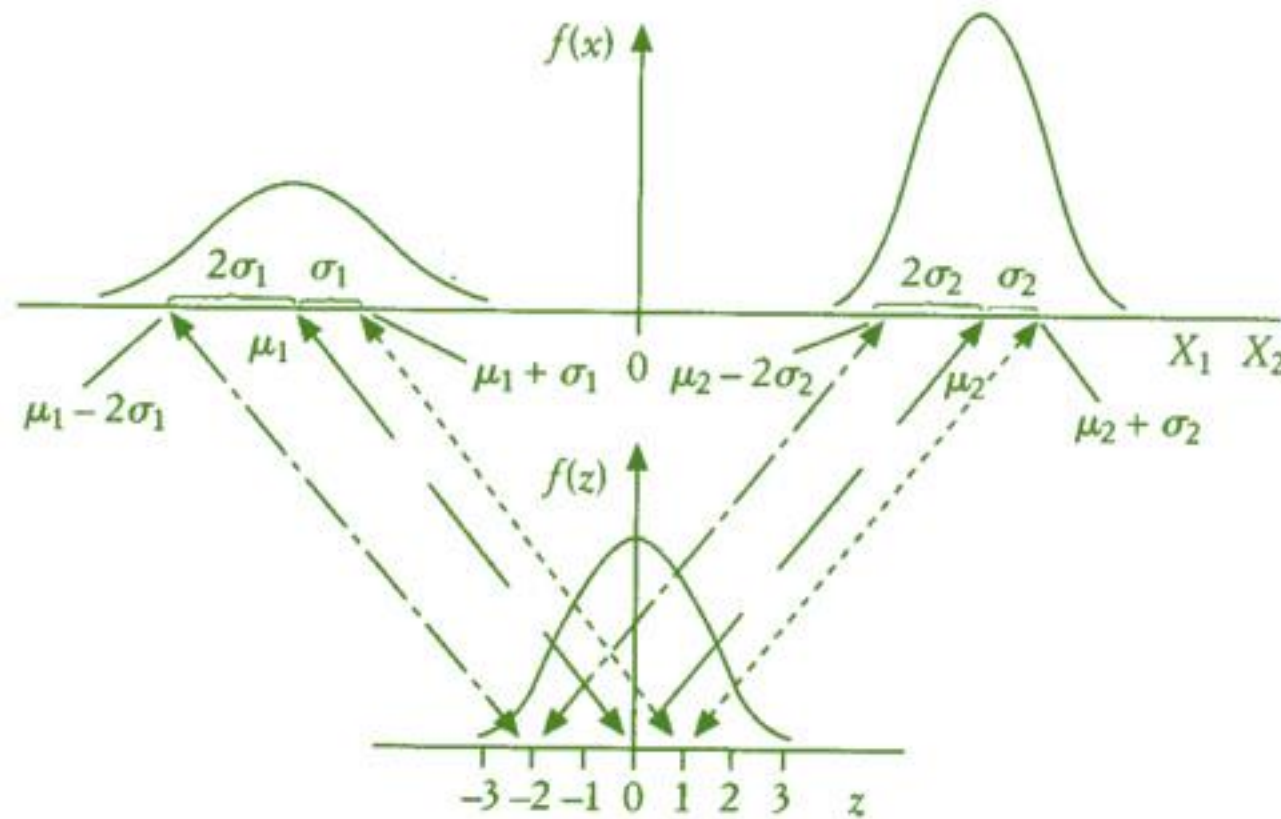
$$Z \sim N(0, 1)$$

Con la seguente funzione di densità:

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

# Variabile casuale normale standardizzata

Ogni variabile distribuita normalmente può essere trasformata in una v.c. standardizzata indipendentemente dal fenomeno, dalla sua unità di misura, dalla sua media e deviazione standard



# Variabile casuale normale standardizzata

L'utilità di ricorrere alla standardizzata risiede nel fatto che è laborioso calcolare i valori della funzione di ripartizione di ogni variabile casuale normale (che sono diversi per ogni  $\mu$  e  $\sigma$ ) mentre i valori della funzione di ripartizione della normale standardizzata sono riportati su delle apposite tavole

**Esempio di  
procedura di  
standardizz.**

$$\Pr(x_1 \leq X \leq x_2) = \Pr\left(\frac{x_1 - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{x_2 - \mu}{\sigma}\right) =$$

$$= \Pr(z_1 \leq Z \leq z_2) = F(z_2) - F(z_1)$$



## Variabile casuale normale standardizzata

In particolar modo si utilizza la standardizzazione allo scopo di calcolare le probabilità che la v.c. assuma un determinato intervallo di valori, “sfruttando” alcune proprietà della normale standardizzata.

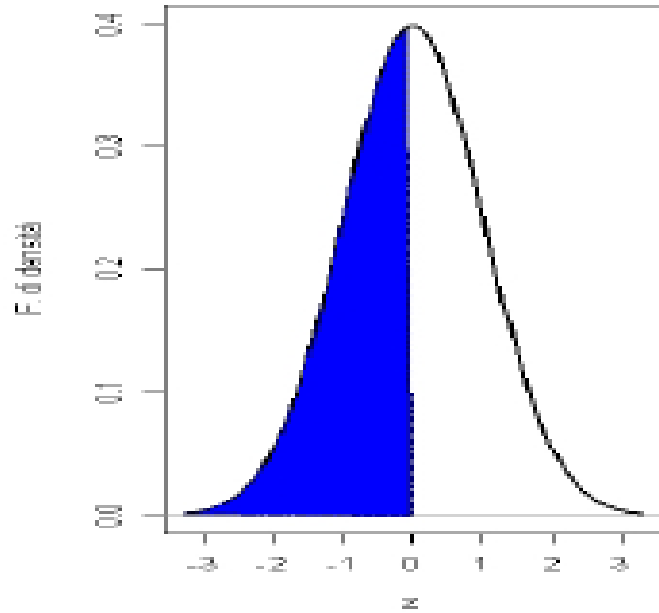
Per la proprietà della simmetria rispetto all'asse  $z=0$  si ha che :

$$\Phi(0) = P(Z \leq 0) = 0,5$$

$$\Phi(-z) = 1 - \Phi(z)$$

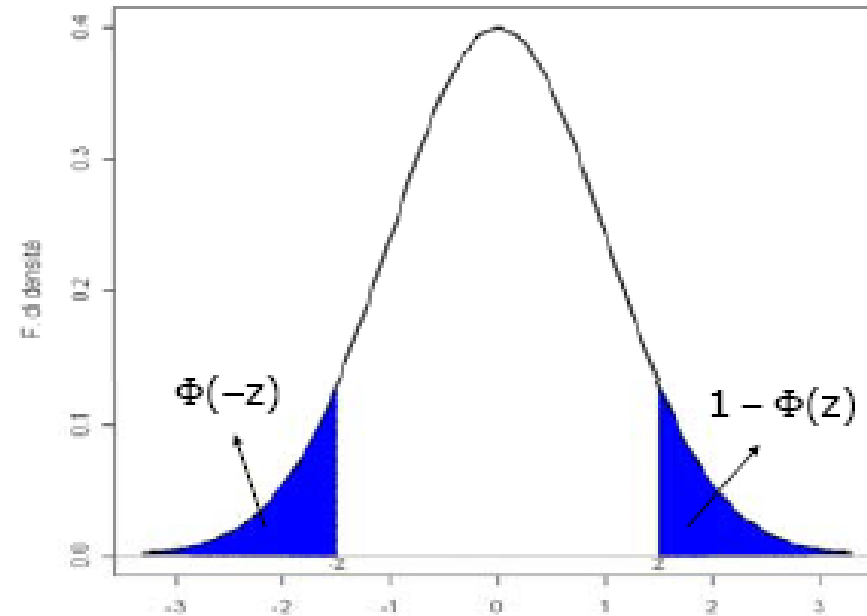
$$P(Z > z) = 1 - \Phi(z)$$

# Variabile casuale normale

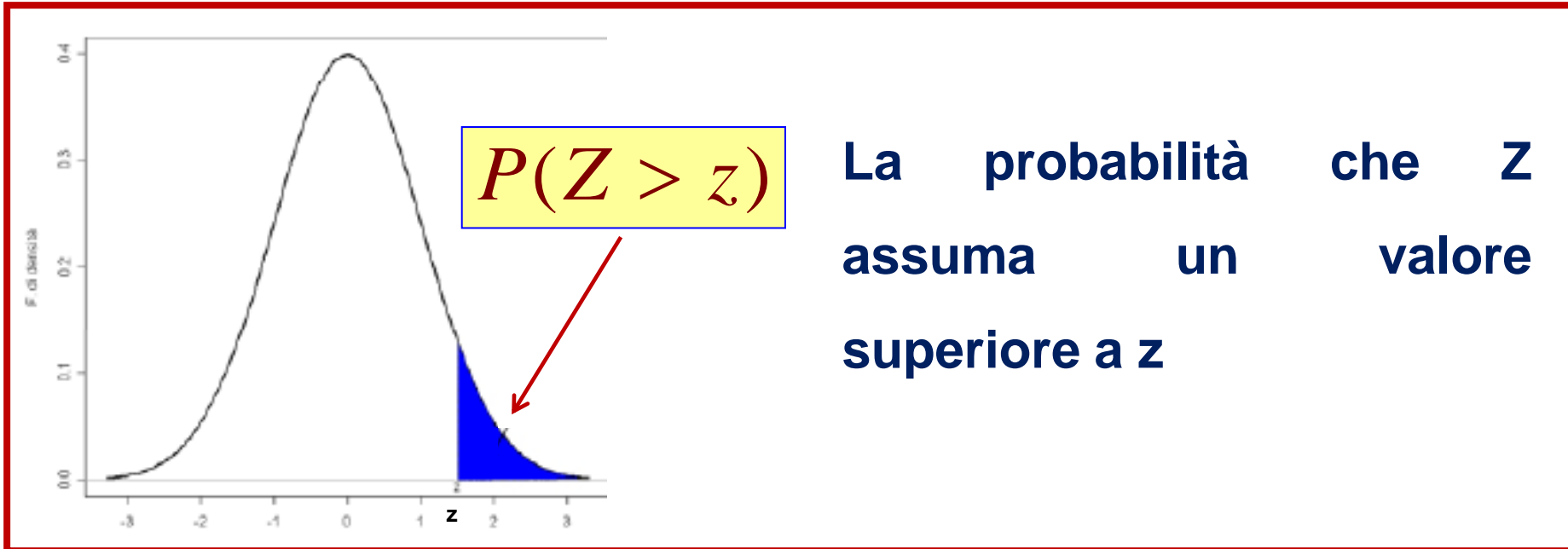


$$\Phi(0) = P(Z \leq 0) = 0,5$$

$$\Phi(-z) = 1 - \Phi(z)$$

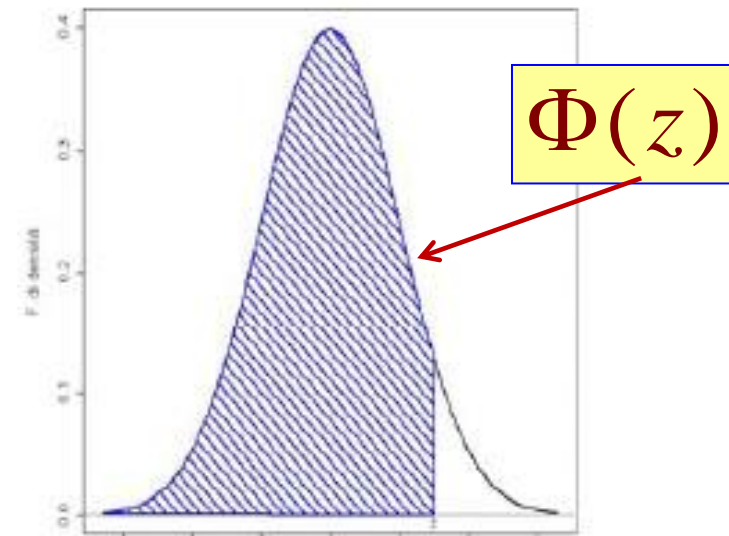


## Variabile casuale normale stand.



Può essere calcolata come la differenza tra 1 e la prob. che Z assumi un valore inferiore a z

$$P(Z > z) = 1 - \Phi(z)$$



# Utilizzo delle tavole

## Esempio

Data una v.c  $X \sim N(8,16)$  si calcoli la probabilità che  $X$  assuma un valore inferiore a 15:

$$\Pr(X \leq 15) = \Pr\left(\frac{X - \mu}{\sigma} \leq \frac{15 - \mu}{\sigma}\right) =$$

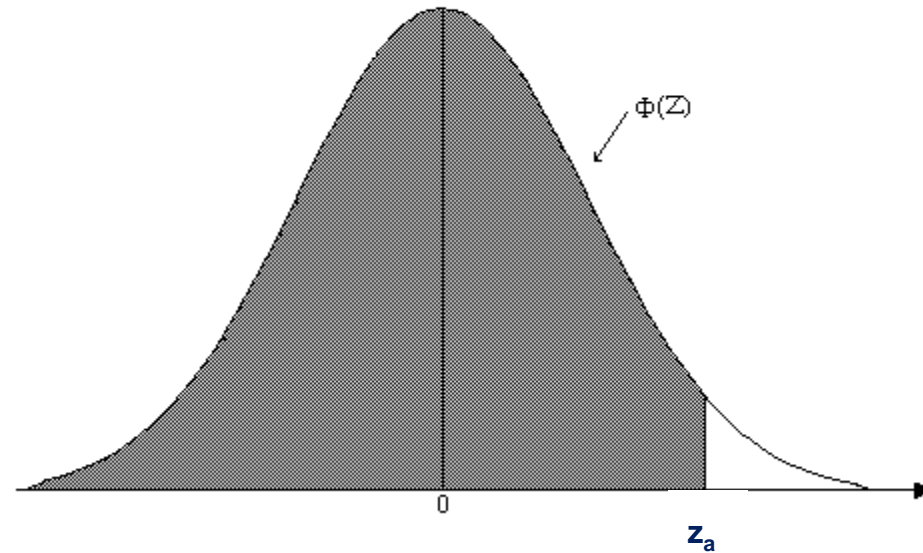
$$= \Pr\left(Z \leq \frac{15 - 8}{4}\right) = \Pr(Z \leq 1,75)$$

Quindi calcolare la prob. che  $X$  sia inferiore a 15 equivale a calcolare la probabilità che  $Z$  sia inferiore a 1,75 (cioè l'area della funzione di ripartizione per  $Z \leq 1,75$  (valore riportato sulle tavole))

# Variabile casuale normale

## Funzione di ripartizione di Z

$$\Pr(z \leq z_a) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_a} e^{-z^2/2} dz$$



# Utilizzo delle tavole

I valori della funzione di ripartizione di Z sono riportati su delle apposite tavole per ogni valore di  $z_a$  (vedi Borra di Ciaccio pag.491)

Tavole della funzione di ripartizione della variabile Normale Standardizzata:

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} t^2} dt$$

Mediante le  
tavole si ricava  
che:  
  
 $\Pr(Z \leq 1,75) =$   
  
 $\Phi(1,75) = 0,959$

z	Seconda cifra decimale di z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169

# Utilizzo delle tavole

UNA SOCIETA' ALBERGHIERA INTERNAZIONALE HA STABILITO CHE IN MEDIA IN OGNUNO DEI PROPRI ALBERGHI SI REGISTRANO 100 CLIENTI AL GIORNO CON UNA VARIANZA DI 64. NELL'IPOTESI CHE IL NUMERO DI CLIENTI GIORNALIERI SI DISTRIBUISCA SECONDO UNA NORMALE E SCELTO A CASO UN ALBERGO QUAL È LA PROBABILITÀ CHE IN UN DETERMINATO GIORNO:

**1. IL NUMERO DI CLIENTI È INFERIORE A 90**

**2. IL NUMERO DI CLIENTI È COMPRESO TRA 90 E 120**

# Variabile casuale normale

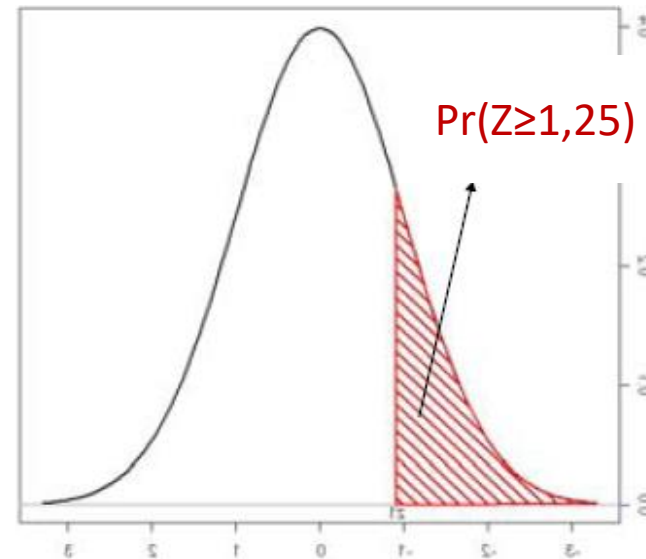
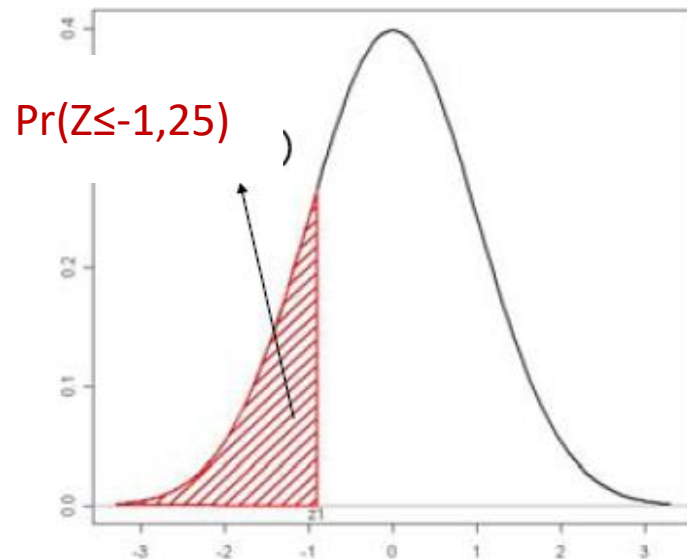
X=clienti giornalieri

Si ipotizza che:  $X \sim N(\mu, \sigma^2)$   ~~$\rightarrow N(100, 64)$~~

A) Probabilità che in un albergo ci siano meno di 90 clienti

$$\Pr(X \leq 90) = \Pr\left(Z \leq \frac{90 - 100}{8}\right) = \Pr(Z \leq -1,25)$$

$$\begin{aligned}\Pr(Z \leq -1,25) &= \Pr(Z \geq 1,25) = 1 - \Pr(Z \leq 1,25) = \\ &= 1 - 0,8944 = 0,1056\end{aligned}$$





## Variabile casuale normale

B) Probabilità che in un albergo ci sia un numero di clienti compreso tra 90 e 120

$$\begin{aligned}\Pr(90 \leq X \leq 120) &= \Pr\left(\frac{90 - 100}{8} \leq Z \leq \frac{120 - 100}{8}\right) \\ &= \Pr(-1,25 \leq Z \leq 2,5)\end{aligned}$$

$$\begin{aligned}\Pr(-1,25 \leq Z \leq 2,5) &= \Pr(z \leq 2,5) - \Pr(z \leq -1,25) = \\ &= \Pr(z \leq 2,5) - [1 - \Pr(z \leq 1,25)] = 0,9938 - 0,1056 = 0,8882\end{aligned}$$