CHE COS'E' LA PROBABILITA'

- La probabilità è la **MISURA** dell'incertezza di un evento, cioè come noi classifichiamo gli eventi rispetto alla loro incertezza.

- La SCALA di Probabilità varia tra 0.00 e 1.00.

0.00 = Evento Impossibile	.50 = Evento massimamente incerto	1.00 = Evento Certo

CHE COS'E' LA PROBABILITA'

- Gli statistici Misurano la Probabilità di un evento come:
  - Il **rapporto** tra i casi (teoricamente) favorevoli all'evento ed i casi (teoricamente) possibili.
  - La **frequenza relativa** dell'evento in un esperimento (cioè, quante volte si verifica l'evento relativamente ad un certo numero di repliche di un esperimento).
  - Il **grado di fiducia** che un individuo ha sul verificarsi di un evento (cioè, quanto egli è disposto a scommettere sull'evento).

DISTRIBUZIONE DI PROBABILITA'

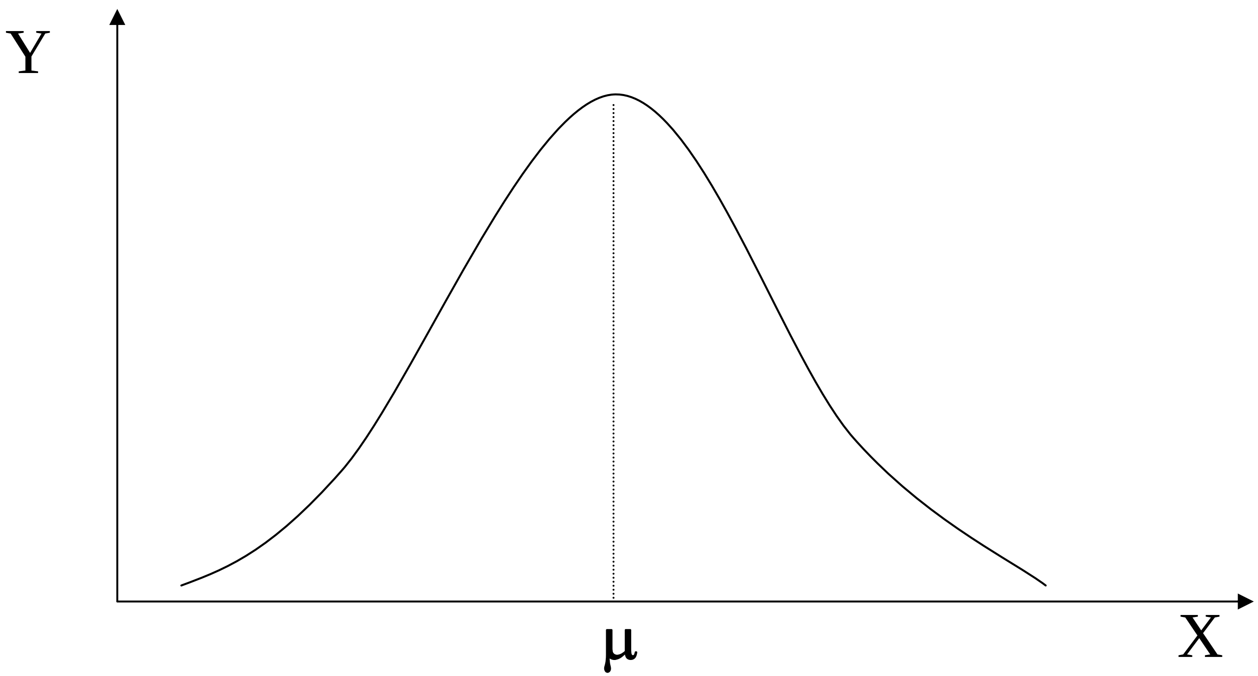
- I possibili risultati di un esperimento costituiscono uno **spazio campionario** di ***n*** ***eventi*** ⇨A ciascun evento possiamo associare la probabilità del suo verificarsi

⬇

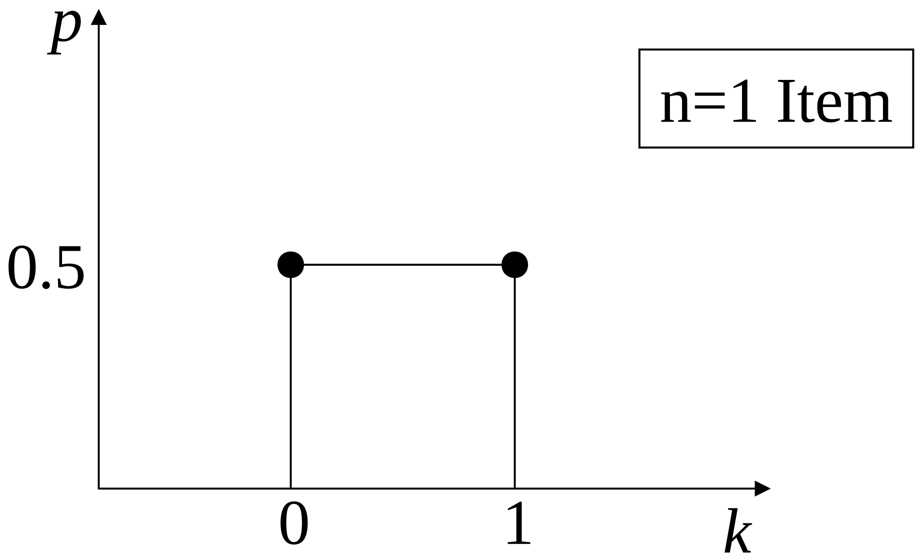
- **DISTRIBUZIONE DI PROBABILITA'** definita da *tutti i possibili risultati* e le *corrispondenti probabilità*

DISTRIBUZIONE NORMALE

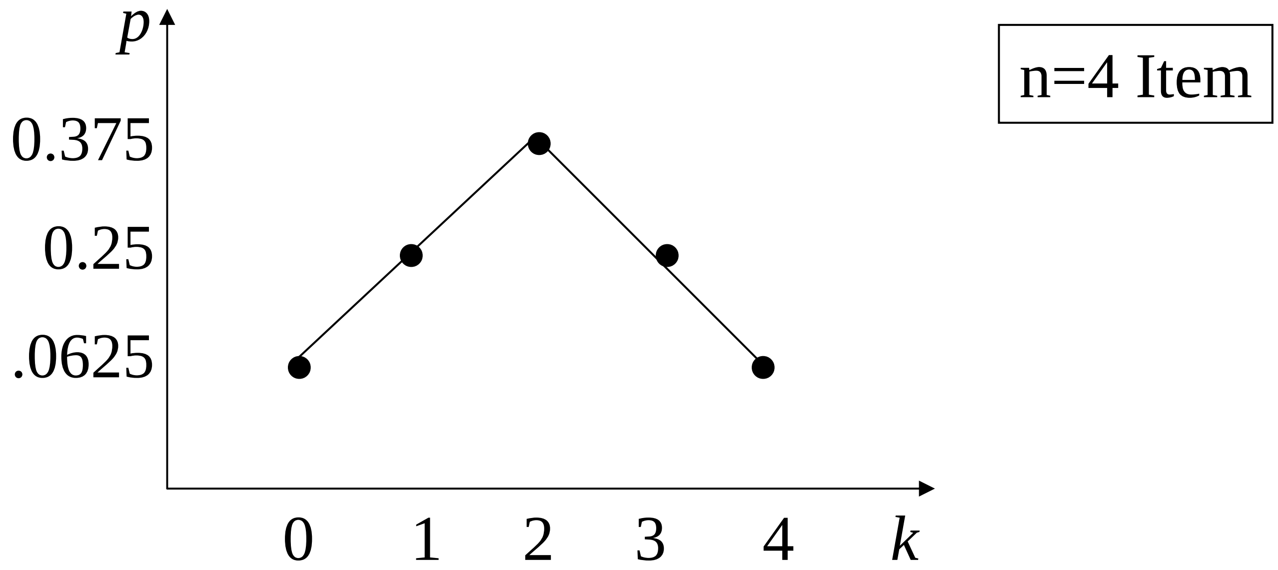
- La distribuzione NORMALE è rappresentata da una particolare curva continua a forma campanulare (gaussiana)



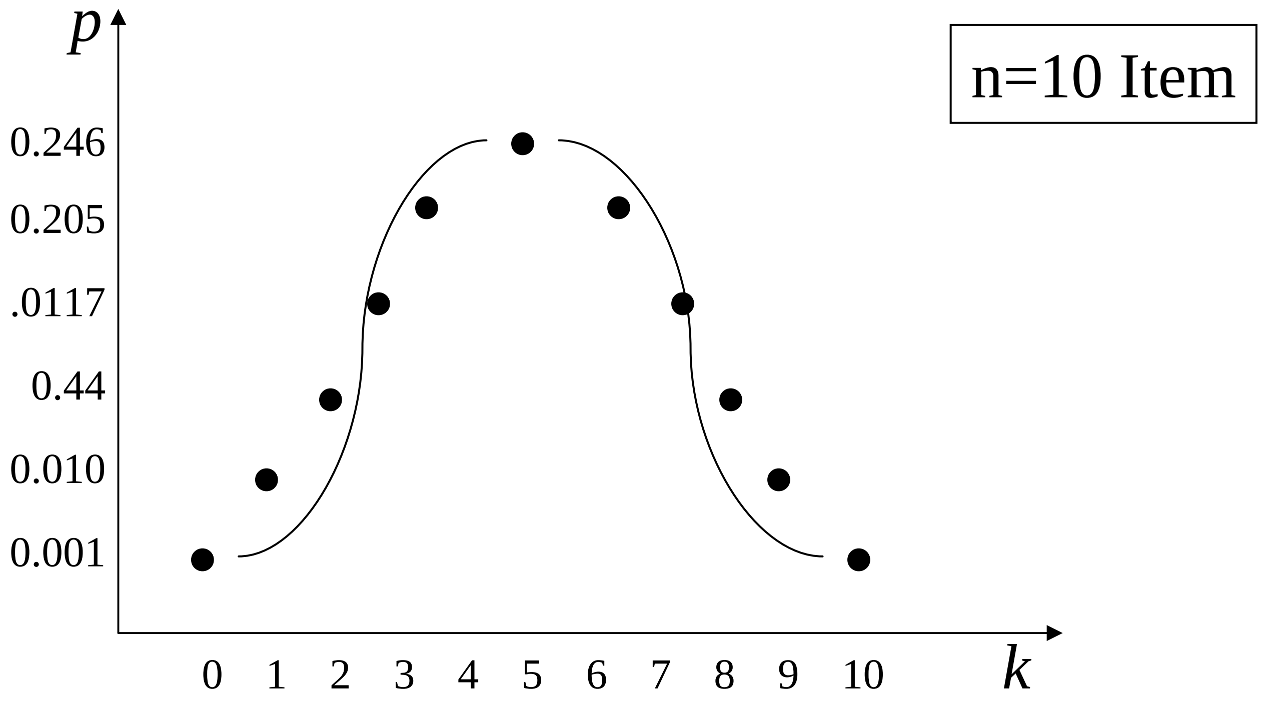
- Risposta ad un Item Vero-Falso: *k*= “Totale Risposte Corrette” con *p*=0.5



- Risposta a 4 Item Vero-Falso: *k*= “Totale Risposte Corrette” con *p*=0.5

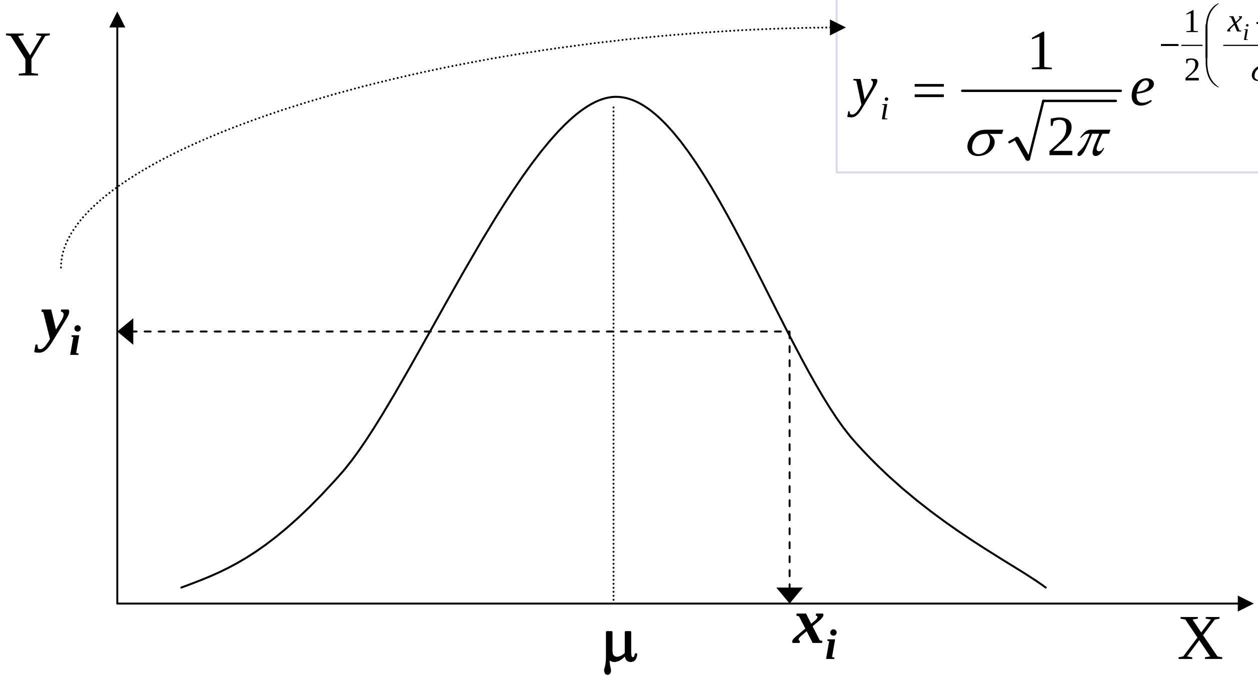


Risposta a 10 Item Vero-Falso: *k*= "Totale Risposte Corrette" con *p*=0.5



DISTRIBUZIONE NORMALE

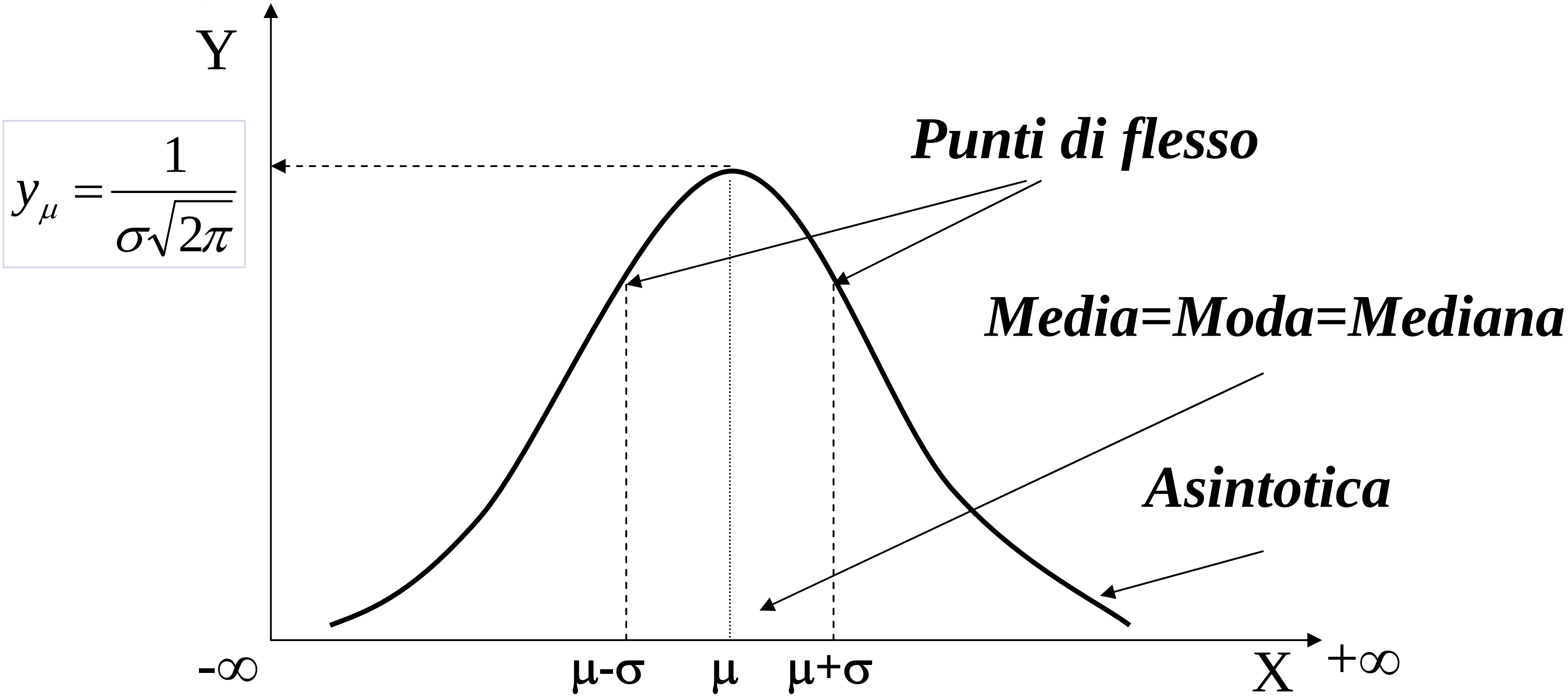
Per qualsiasi valore ***x*** che la variabile può assumere, attraverso la funzione si calcola la ***y*** corrispondente, cioè la probabilità



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

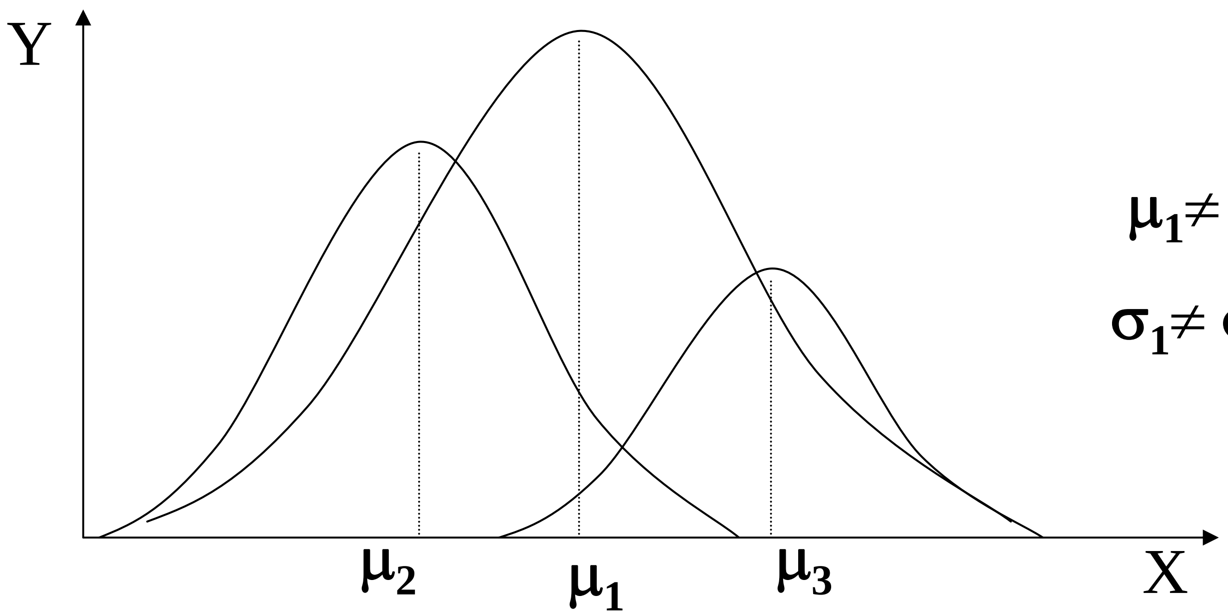
DISTRIBUZIONE NORMALE

CRESCENTE per $-\infty<x<\mu$ e DECRESCENTE per $\mu<x<+\infty$ ⇨ due punti di flesso a $\pm\,\sigma$ da μ



DISTRIBUZIONE NORMALE

- La curva NORMALE è definita dai parametri μ e σ ⇨ famiglia di distribuzioni normali con medie e deviazioni standard diverse



$\mu_1 \neq \mu_2 \neq \mu_3$

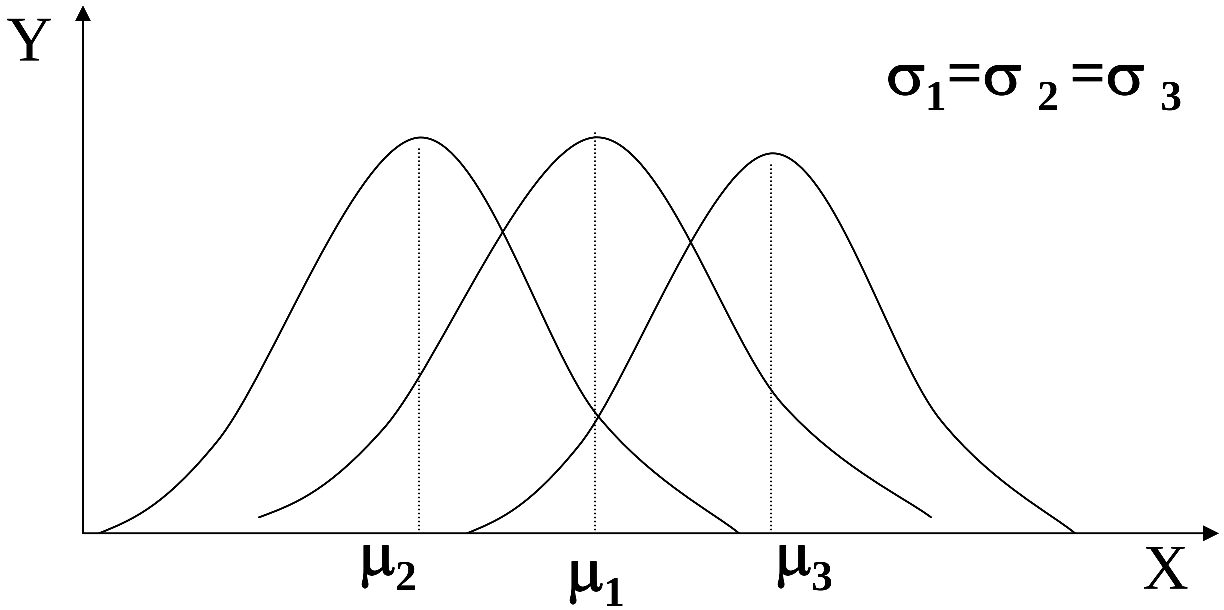
$\sigma_1 \neq \sigma_2 \neq \sigma_3$

DISTRIBUZIONE NORMALE

⇨ famiglia di distribuzioni normali con una diversa media e con la stessa deviazione standard

$\mu_1 \neq \mu_2 \neq \mu_3$

$\sigma_1 = \sigma_2 = \sigma_3$



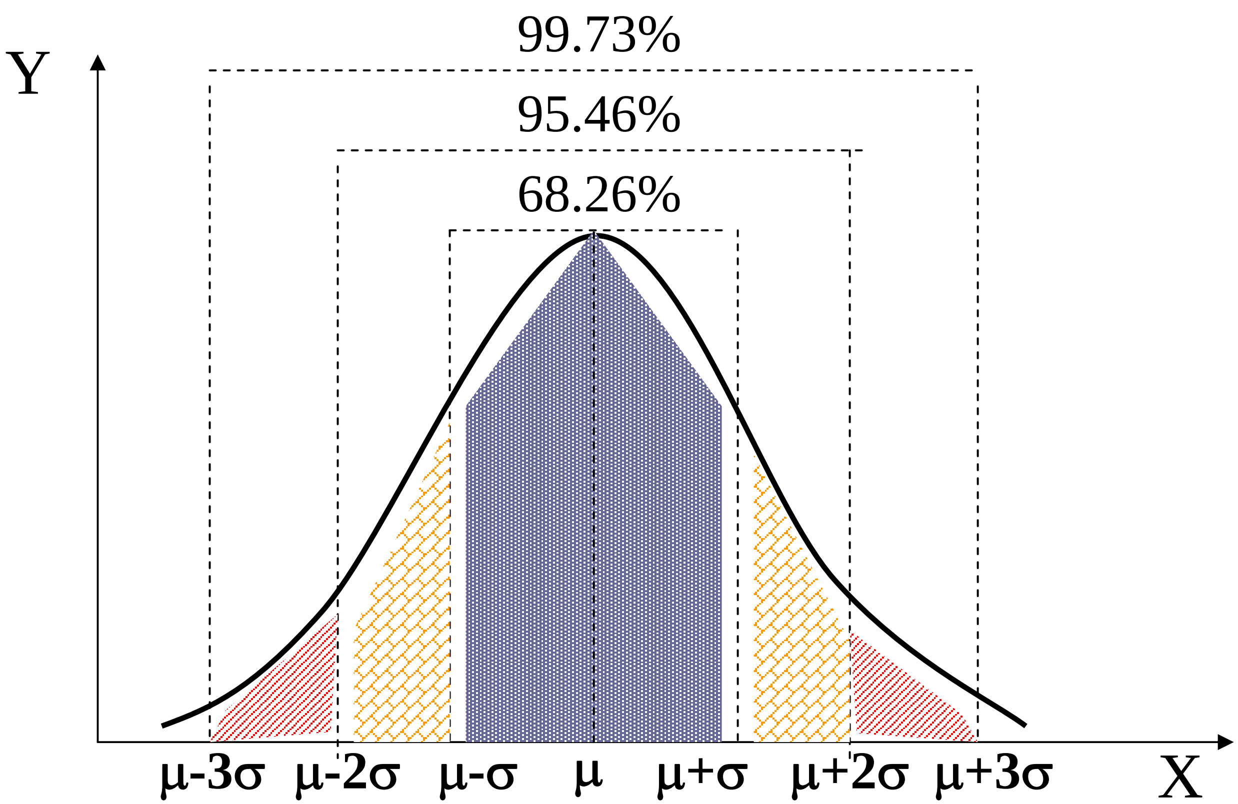
DISTRIBUZIONE NORMALE

- Qualsiasi siano i parametri μ e σ**,** l'area della porzione di curva delimitata dalla media e un ordinata espressa in termini di deviazioni standard è ***costante***
  - ⇨ $\boldsymbol{\mu+\sigma}$= 34.13% della distribuzione
  - ⇨ $\boldsymbol{\mu+2\sigma}$= 47.73% della distribuzione
  - ⇨ $\boldsymbol{\mu+3\sigma}$= 49.86% della distribuzione

$$\Downarrow$$

DISTRIBUZIONE NORMALE

Porzioni della distribuzione comprese tra ± 1,2,3 deviazioni standard da μ (in %)



STANDARDIZZAZIONE
PUNTI *Z*

- Consentono riferire una misura ad una scala standard con media uguale a zero e deviazione standard uguale a 1.
- La trasformazione dei valori x_i in valori z_i significa esprimere i valori come ***distanza dalla media in termini di deviazioni standard*** (cioè, usare la deviazione standard come unità di misura)
- Il segno del punteggio z indica immediatamente la posizione del soggetto sopra (+) o sotto la media (-)

STANDARDIZZAZIONE PUNTI *Z*

- Il punto z_i indica *la distanza dalla media* del valore $\boldsymbol{x_i}$ dell'*i-esimo* soggetto, distanza espressa in deviazioni standard.
- Per calcolare i punti z:

$$\boxed{z_i = \frac{x_i - M}{s}}$$

STANDARDIZZAZIONE
PUNTI *Z*

- **Esempio:**

In un test di percezione visiva la media è 21.25 con una deviazione standard di 6.74. Trasformare in punti *z* i seguenti punteggi ottenuti da 6 soggetti dislessici.

Sogg.	1	2	3	4	5	6
x_i	8	14	17	20	25	28

STANDARDIZZAZIONE PUNTI *Z*

- Si standardizzano i punteggi:

$$z_1 = \frac{8-21.25}{6.74} = -1.97 \qquad z_2 = \frac{14-21.25}{6.74} = -1.10$$

$$z_3 = \frac{17-21.25}{6.74} = -0.63 \qquad z_4 = \frac{20-21.25}{6.74} = -0.18$$

$$z_5 = \frac{25-21.25}{6.74} = 0.56 \qquad z_6 = \frac{28-21.25}{6.74} = 1.00$$

STANDARDIZZAZIONE
PUNTI *Z*

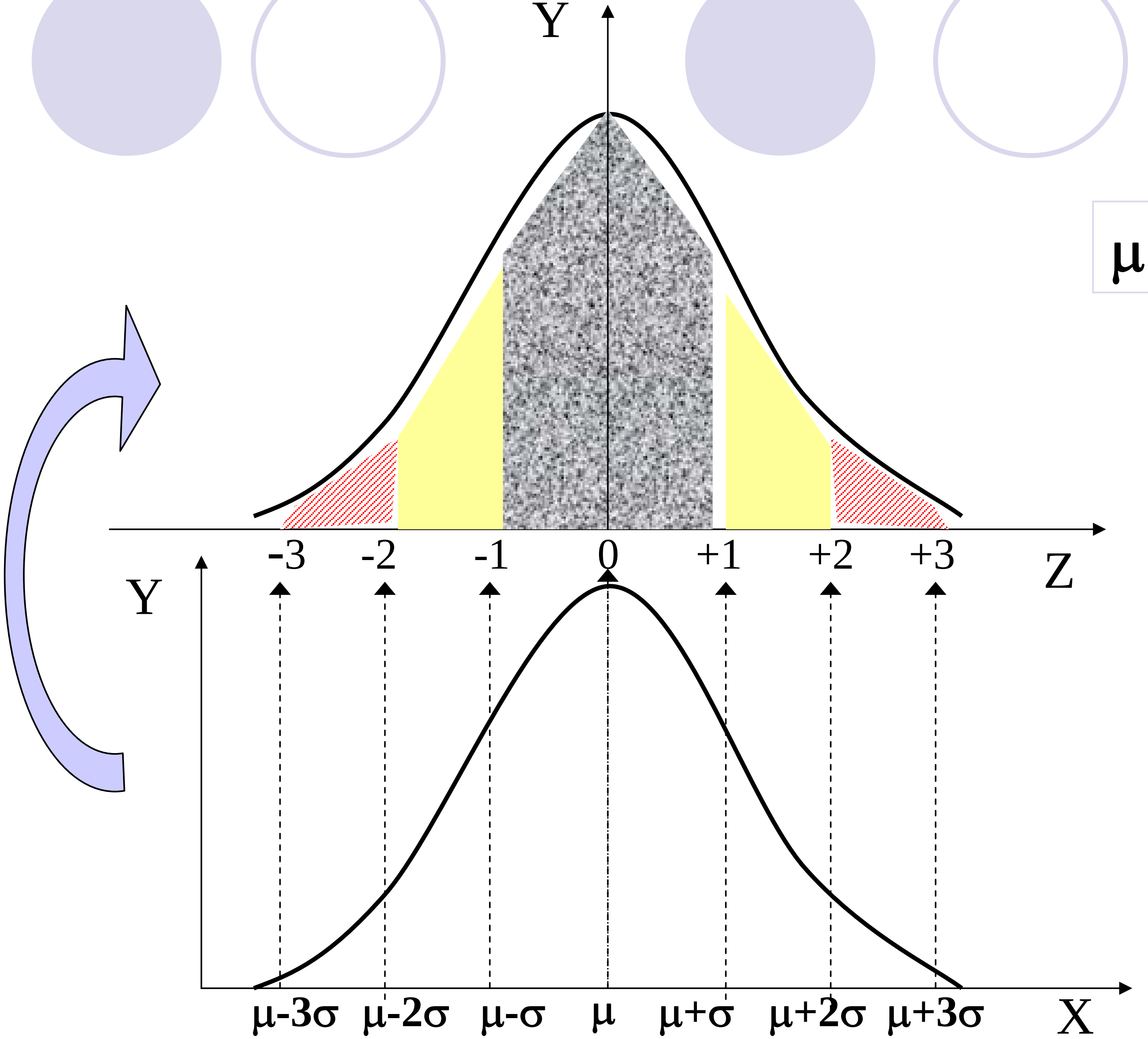
x_i	8	14	17	20	**21.25**	25	28
z_i	-1.97	-1.10	-0.63	-0.18	**0**	0.56	1.00
					(**MEDIA**)		

- Il soggetto n°6 con 25 è circa **mezza** deviazione standard ***sopra*** la media e dista da questa circa ***la metà*** rispetto al soggetto n°7.
- Il soggetto n°1 con 8 è circa **due** deviazioni standard ***sotto*** la media e dista da questa circa ***il doppio*** rispetto al soggetto n°2.

DISTRIBUZIONE NORMALE STANDARDIZZATA

- Trasformando i valori di ***x*** in punti ***z*** si ottiene una distribuzione **STANDARDIZZATA** con $\mu=0$ e $\sigma=1$

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



$\mu=0 \quad \sigma=1$

$\mu \quad \sigma$

DISTRIBUZIONE NORMALE

- Data una variabile *continua* possiamo associare ai possibili valori la **probabilità** del loro verificarsi ⇨ La probabilità associata ai valori di una variabile continua è sempre definita entro un ***intervallo*** i cui estremi delimitano un'area

⇩

DISTRIBUZIONE DI PROBABILITA' NORMALE

DISTRIBUZIONE NORMALE

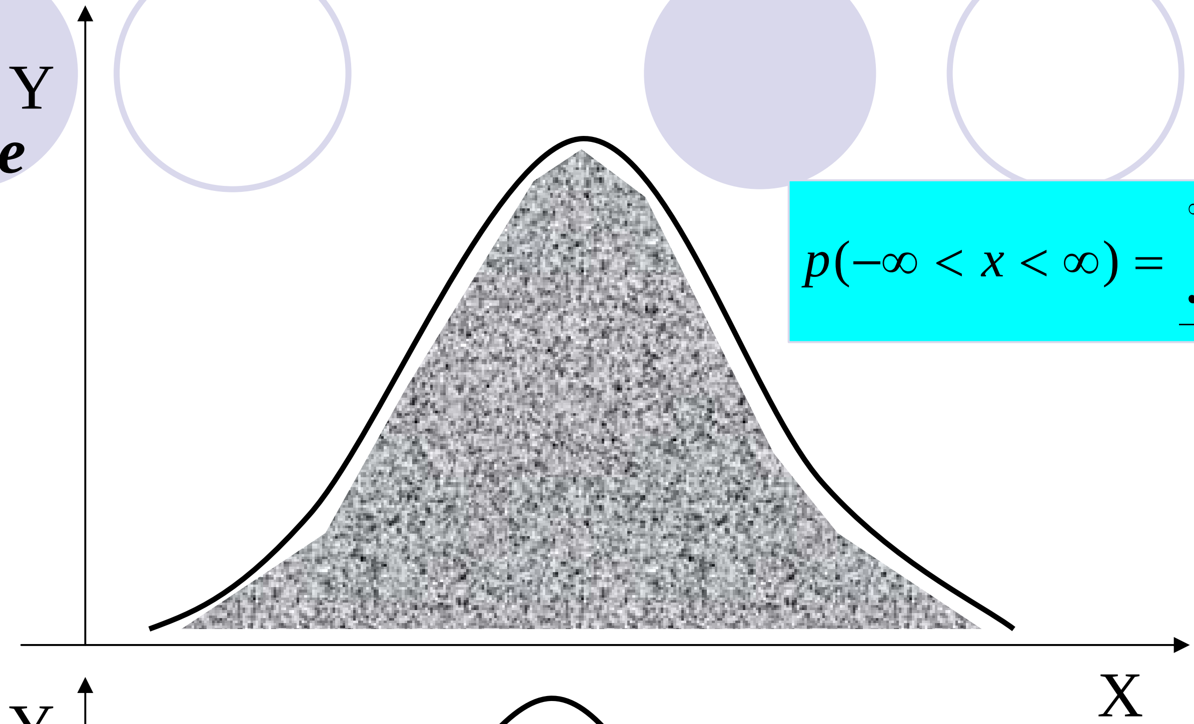
- L'area si ottiene risolvendo un'integrale

$$p(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x)dx$$

Il valore che si ottiene è sempre compreso tra 0 e 1(⇦ probabilità)

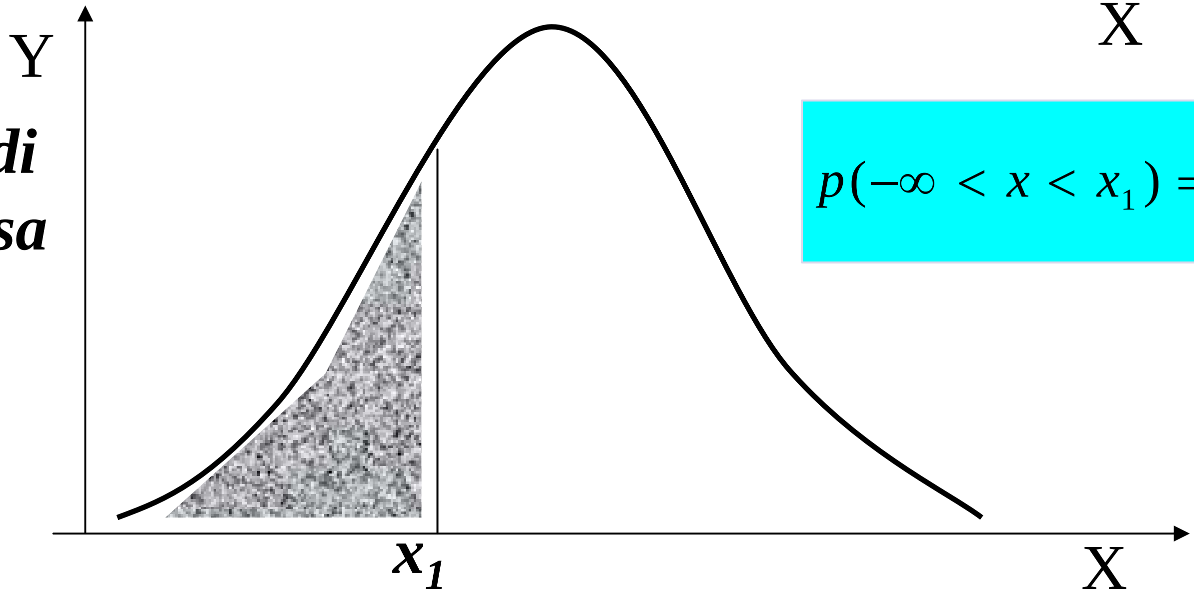
Moltiplicando tale valore per 100 si ottiene la ***percentuale*** della distribuzione compresa tra x_1 e x_2

Area totale sottesa alla curva



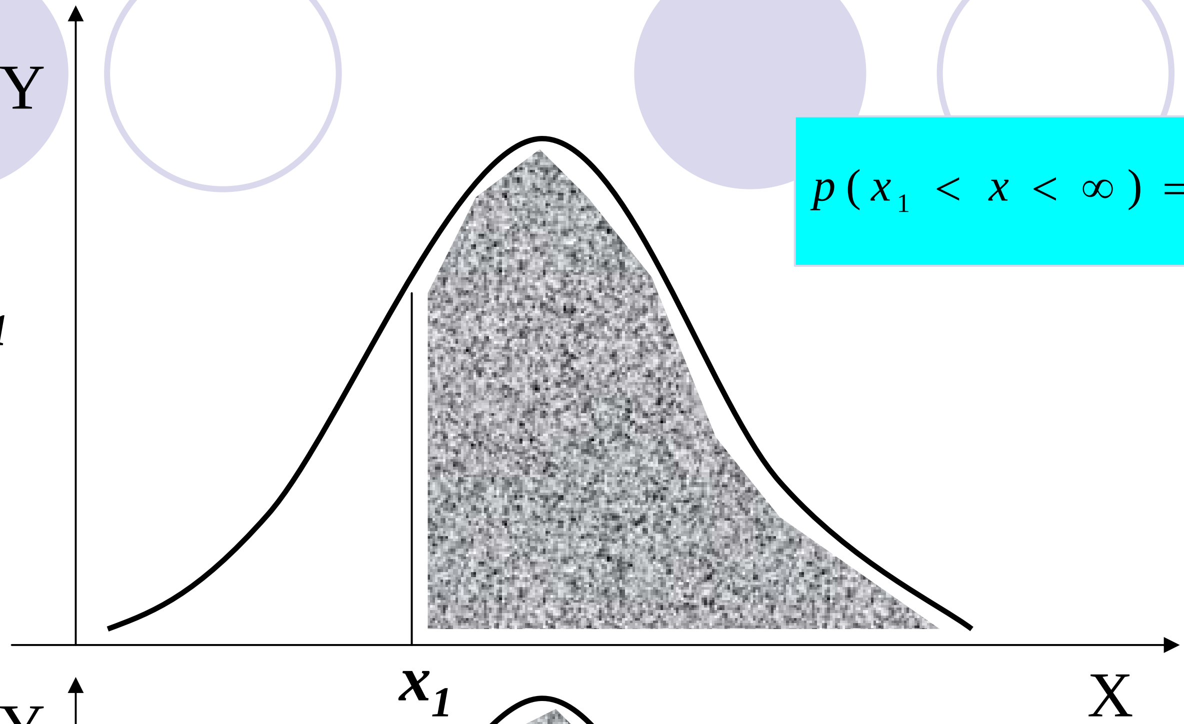
$$p(-\infty < x < \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$$

Porzione di area sottesa alla curva



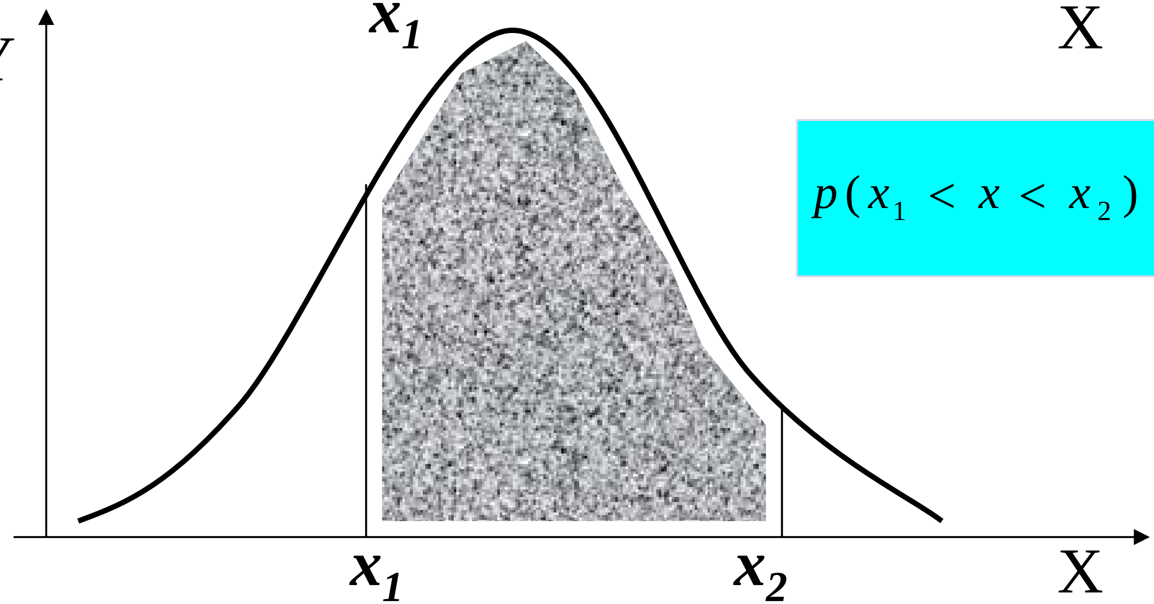
$$p(-\infty < x < x_1) = \int_{-\infty}^{x_1} f(x)dx$$

Porzione di area sottesa alla curva



$$p(x_1 < x < \infty) = \int_{x_1}^{\infty} f(x)\,dx$$

Porzione di area sottesa alla curva



$$p(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x)\,dx$$

DISTRIBUZIONE NORMALE

- Per la curva normale standardizzata (μ=0; σ=1) sono stati tabulati i valori degli integrali per tutti i valori di *z*, cioè di tutte le aree comprese tra 0 (media) e un qualsiasi *z*

$\Downarrow$

DISTRIBUZIONE DI PROBABILITA' NORMALE STANDARDIZZATA

TAVOLA DI Z

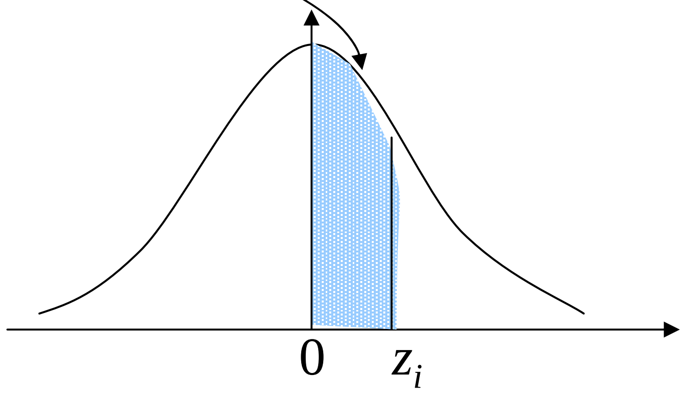
Riporta le aree comprese tra μ=0 e z

2° cifra decimale di z

***Valore** di z con la 1° cifra decimale*

Z	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
0.0										
0.1										
0.2										
0.3										
.					area tra 0 e z_i					
.										
1.0										
1.1										
.										
.										
.										
3.9										
∞										

$$p(0 < z < z_i) = \int_0^{z_i} f(z)dz$$



Esempio:

Se z = 1.03 ⇨ (0<z<1.03)=.3485

Z	0	1	2	3	4	5	6	7	8	9
0.0										
0.1										
0.2										
.										
.										
.										
1.0										
1.1										
.										
.										
.										
3.9										
∞										

.3485= area tra 0 e 1.03
(circa il 35% della distribuzione)

$$p(0 < z < 1.03) = \int_0^{1.03} f(z)dz = .3485$$

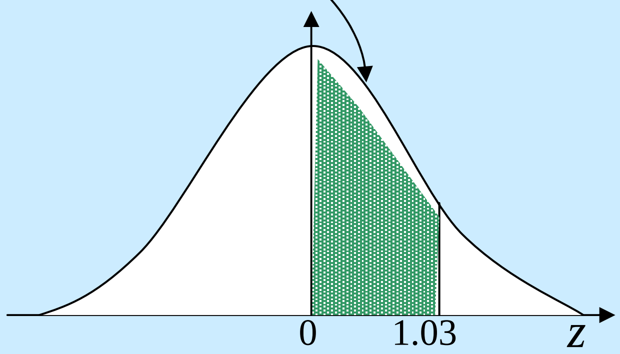


TAVOLA DI *Z*

Uso della Tavola di z:

Per avere la probabilità associata ad un intervallo i cui estremi sono valori *x*:

① trasformazione delle *x* in punti *z* ⇨

② attraverso la tavola si risale all'area delimitata da questi valori. Questa area rappresenta la probabilità ad essi associata

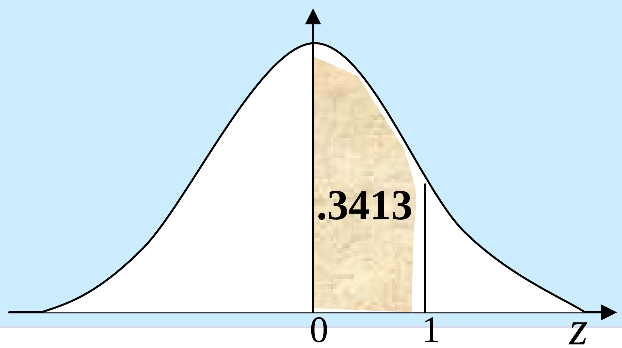
- **Esempio *a***: Data una distribuzione con μ=100 e σ=10 qual è la probabilità che un valore sia compreso tra la media e 110?

Si trasformano gli estremi dell'intervallo in punti *z*: μ=100 ⇨ *z* = 0

x=110 ⇨ $z = \frac{110-100}{10} = 1$

Si cerca sulla tavola l'area tra 0 e *z*=1.00

$$p(100 < x < 110) = p(0 < z < 1) = \int_0^1 f(z)dz = .3413$$



- **Esempio *b***: Qual è la probabilità che un valore sia compreso tra 90 e la media?

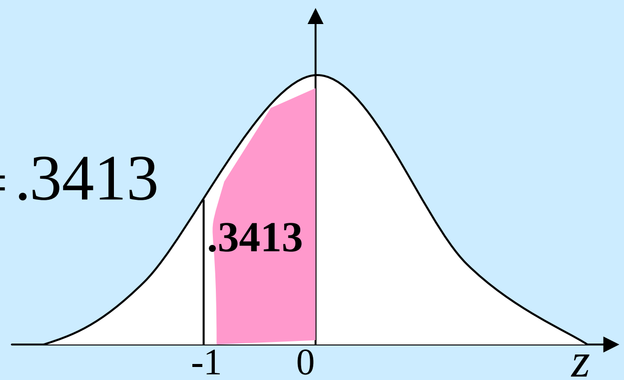
Si trasformano gli estremi dell'intervallo in punti *z*:

μ=100 ⇨ *z* = 0;

x=90 ⇨ $z = \frac{90 - 100}{10} = -1$

Si cerca sulla tavola l'area tra 0 e z=1.00 che data la simmetria della curva è identica a quella tra -1 e 0

$$p(90 < x < 100) = p(-1 < z < 0) = \int_{-1}^{0} f(z)dz = .3413$$



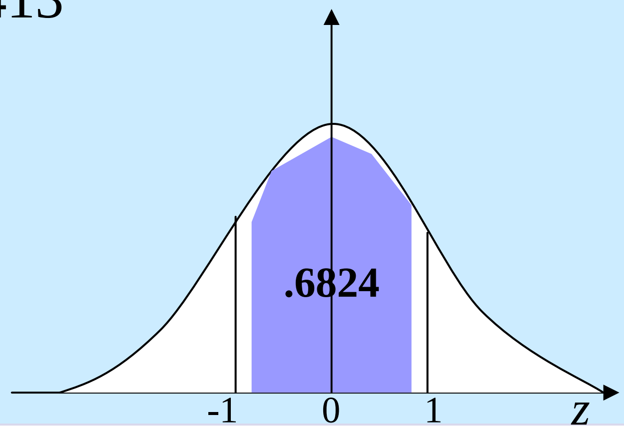
- **Esempio *c***: Qual è la probabilità che un valore sia compreso tra 90 e 110?
- $p(90<x<110) = p(90<x<100) + p(100<x<110)$ ⇨

$$z_1 = \frac{90-100}{10} = -1 \qquad z_2 = \frac{110-100}{10} = 1$$

$$p(90<x<100) = p(-1<z<0) = \int^{0} f(z)dz = .3413$$

$$p(100<x<110) = p(0<z<1) = \int_{0}^{1} f(z)dz = .3413$$

.3413+.3413=.6824



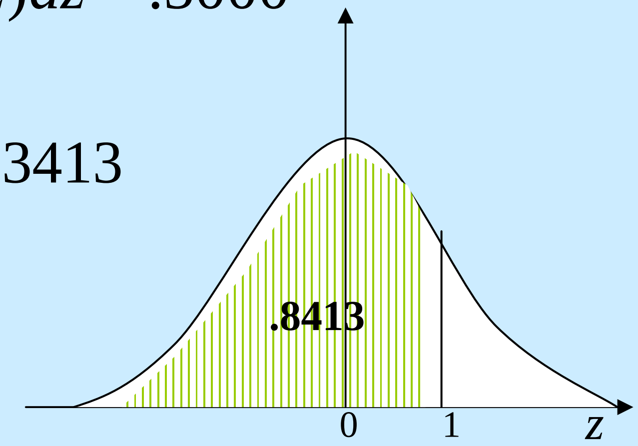
Esempio *d*: Qual è la probabilità che un valore sia minore di 110?

$p(x<110)= p(-\infty<x<100)+ p(100<x<110)$ ⇨

$$z = \frac{110-100}{10} = 1$$

$$p(-\infty < x < 100) = p(-\infty < z < 0) = \int_{-\infty}^{0} f(z)dz = .5000$$

$$p(100 < x < 110) = p(0 < z < 1) = \int_{0}^{1} f(z)dz = .3413$$



$p(x<110) = p(z<1) =$

$.5000 + .3413 = .8413$

Esempio *e*: Qual è la probabilità che un valore sia maggiore di 110?

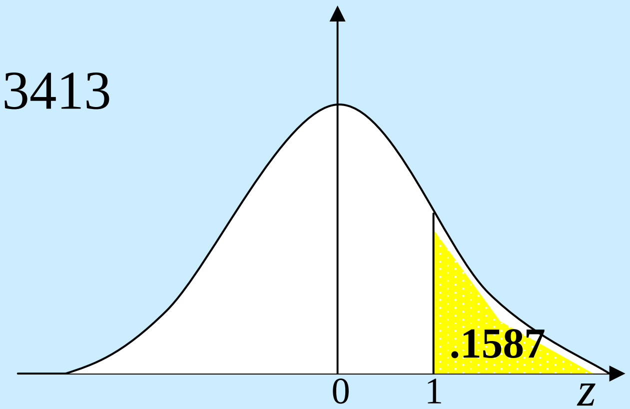
p(*x*>110)= *p*(100<*x*<∞)- *p*(100<*x*<110) ⇨

$$z = \frac{110-100}{10} = 1$$

$$p(100 < x < \infty) = p(0 < z < \infty) = \int_0^\infty f(z)dz = .5000$$

$$p(100 < x < 110) = p(0 < z < 1) = \int_0^1 f(z)dz = .3413$$

.5000 - .3413 = .1587



- **Esempio *f***: Qual è la probabilità che un valore sia compreso tra 108 e 110?

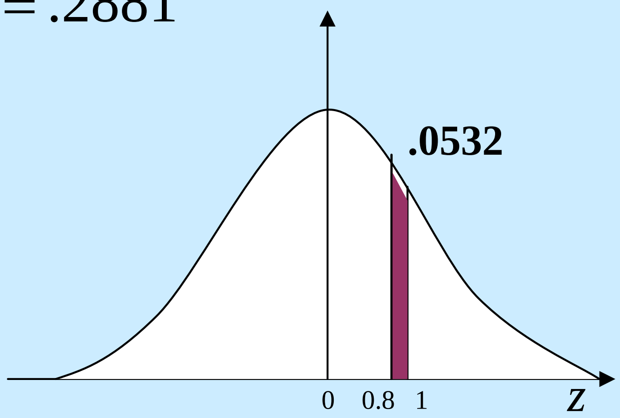
p(108<*x*<110)= *p*(100<*x*<110)- *p*(100<*x*<108)

⇨ $z_1 = \dfrac{108-100}{10} = 0.8$ $z_2 = \dfrac{110-100}{10} = 1$

$$p(100 < x < 108) = p(0 < z < 0.8) = \int_0^{0.8} f(z)dz = .2881$$

$$p(100 < x < 110) = p(0 < z < 1) = \int_0 f(z)dz = .3413$$

.3413 - .2881 = .0532



DISTRIBUZIONE NORMALE

- Se si conosce la porzione di area (*ovvero* la probabilità associata) delimitata da due valori (grezzi o standardizzati) di una distribuzione normale con N noto è possibile risalire alle ***frequenze teoriche*** comprese in quella porzione di area ⇨

$$f = Area \times N$$

- **Esempio *a***: Data una distribuzione normale con N=1000, l'area compresa tra 0 e z = .3413. Quali sono le *frequenze teoriche* comprese nell'intervallo?

$$f = .3413 \times 1000 = 341$$

$$\Downarrow$$

Le frequenze comprese nell'intervallo tra la media e z sono 341

- **Esempio *b***: Data una distribuzione normale con N=1000, μ=100 e σ=10 quali sono le frequenze teoriche comprese tra 108 e 110?

Dall'esempio ***f*** sappiamo che:

x_1=108 → z_1 =0.8

x_2=110 → z_2 =1

$$p(0.8 < z < 1) = .0532$$

$$\boxed{f = .0532 \times 1000 = 53}$$