Distribuzione Normale

ADCOM 2025-2026

Filippo Gambarota PhD

filippo.gambarota@unipd.it

Università di Padova

Ultimo aggiornamento: 04-01-2026

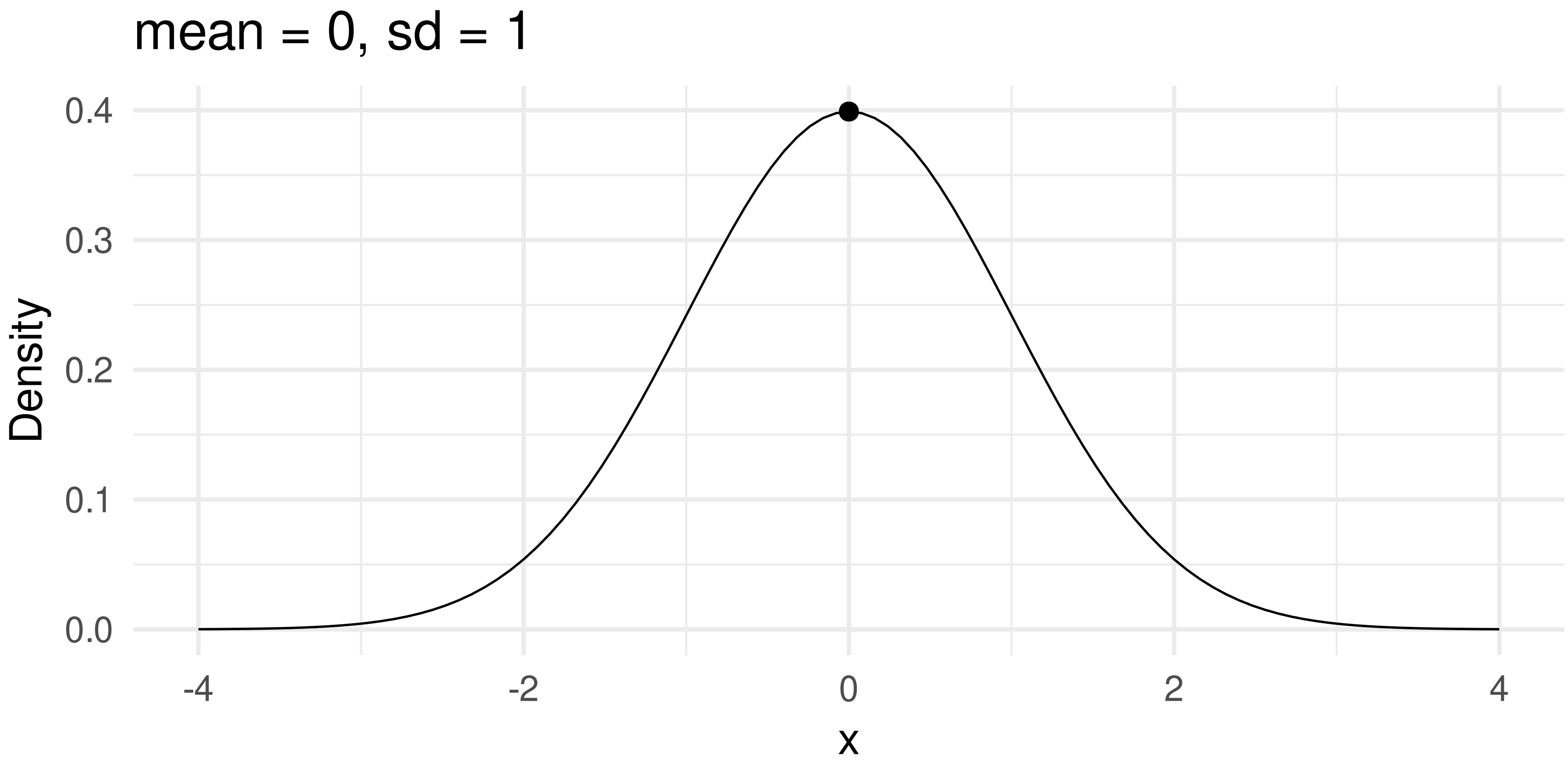
Distribuzione Normale

La distribuzione Normale o Gaussiana è una delle principali distribuzioni di probabilità. E’ una distribuzione di probabilità continua che dipende da due parametri e . La PDF (probability density function) è:

I valori ammessi sono compresi tra e e sono simmetrici rispetto alla media.

Distribuzione Normale

Possiamo calcolare la densità dati i valori di e . Per la densità è circa 4.



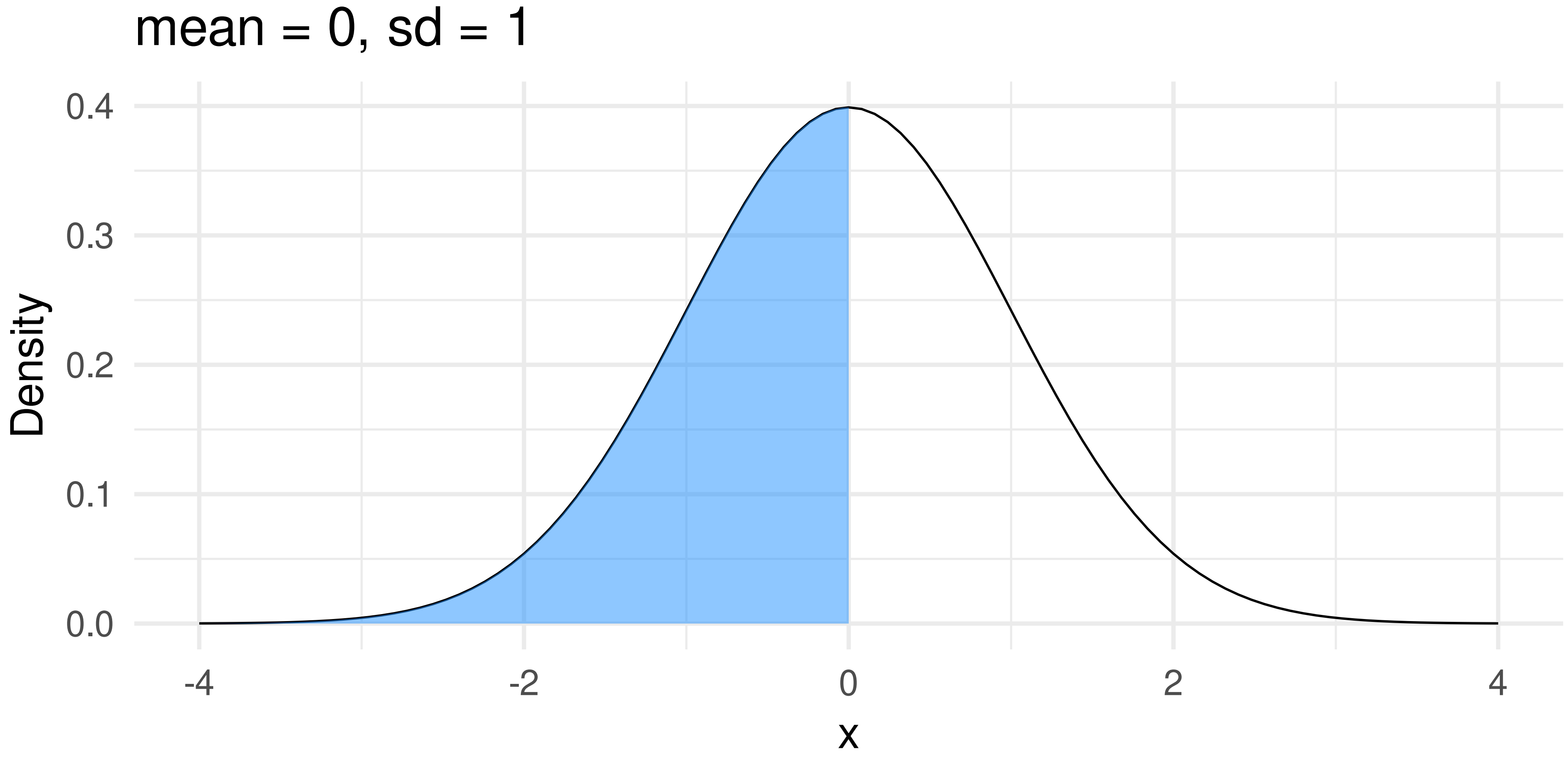
Distribuzione Normale, CDF

Oltre alla densità presentata prima, possiamo calcolare la probabilità cumulata fino al valore :

Questo è lo stesso concetto dei percentili e dei ranghi percentili. Questa funzione risponde alla domanda, quanta probabilità c'è di avere un valore (o) avendo una Normale con media e deviazione standard ?

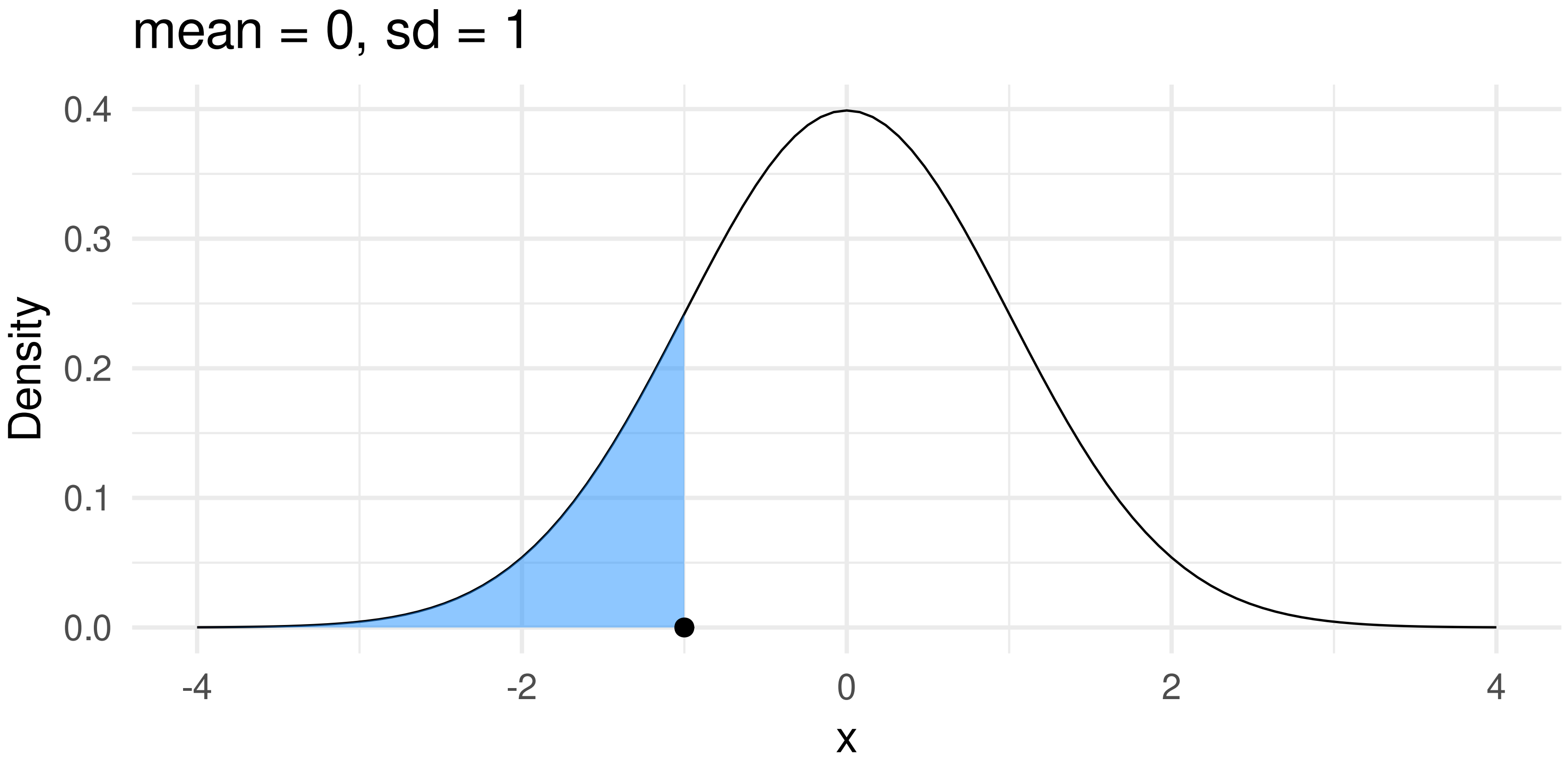
Distribuzione Normale, CDF

Probabilità cumulata fino a è del 50%.



Distribuzione Normale, CDF inversa

Ovviamente, esiste anche una funzione che data la probabilità cumulata restituisce il valore di associato. è il valore che si lascia a sinistra il ~16% di probabilità cumulata.



Distribuzione Normale

Tutto questo è possibile, ovviamente, assumendo di conoscere e . Sono formule che permettono di calcolare queste quantità (densità, probabilità cumulata e inverso) per normali *teoriche*.

Con dati *empirici* i quantili e ranghi percentili svolgono una funzione simile. Più la distribuzione è effettivamente normale, più queste approssimazioni teoriche sono vicine al dato empirico.

Distribuzioni in R

In R sono implementate le principali distribuzioni di probabilità, tra cui chiaramente quella Normale.

Inoltre ci sono 4 funzioni per lavorare con le distribuzioni normali:

- `d` + nome della distribuzione ad esempio `dnorm`: calcola la densità di un certo valore
- `p` + nome della distribuzione ad esempio `pnorm`: calcola la probabilità cumulata fino al valore
- `q` + nome della distribuzione ad esempio `qnorm`: calcola il valore associato ad una certa probabilità cumulata
- `r` + nome della distribuzione ad esempio `rnorm`: genera numeri casuali da una distribuzione normale.

Distribuzioni in R

Ad esempio, calcoliamo la densità di probabilità del valore per una normale con e :

```r
dnorm(x = 100, mean = 80, sd = 10)
```

```
[1] 0.005399097
```

```r
dnorm(x = 80, mean = 80, sd = 10)
```

```
[1] 0.03989423
```

```r
dnorm(x = 1, mean = 80, sd = 10)
```

```
[1] 1.118796e-15
```

Distribuzioni in R

Ora calcoliamo la probabilità cumulata fino a sempre per la stessa distribuzione:

```r
pnorm(q = 100, mean = 80, sd = 10)
```

```
[1] 0.9772499
```

```r
pnorm(q = 80, mean = 80, sd = 10)
```

```
[1] 0.5
```

```r
pnorm(q = 5, mean = 80, sd = 10)
```

```
[1] 3.190892e-14
```

Distribuzioni in R

Ora calcoliamo l'inverso:

```r
qnorm(p = 0.5, mean = 80, sd = 10)
```

```
[1] 80
```

```r
qnorm(p = 0.2, mean = 80, sd = 10)
```

```
[1] 71.58379
```

```r
qnorm(p = 0.9, mean = 80, sd = 10)
```

```
[1] 92.81552
```

Ovviamente c'è una corrispondenza diretta tra `qnorm` e `pnorm`:

```r
(q <- qnorm(0.2, 80, 10))
```

```
[1] 71.58379
```

```r
pnorm(q, 80, 10)
```

```
[1] 0.2
```

Normale standard

La distribuzione Normale standard è un tipo particolare di distribuzione normale con μ e σ . La cosa interessante è che tutte le distribuzioni normali possono trasformarsi in normali standard e viceversa.

I valori x di una normale standard (o punti z):

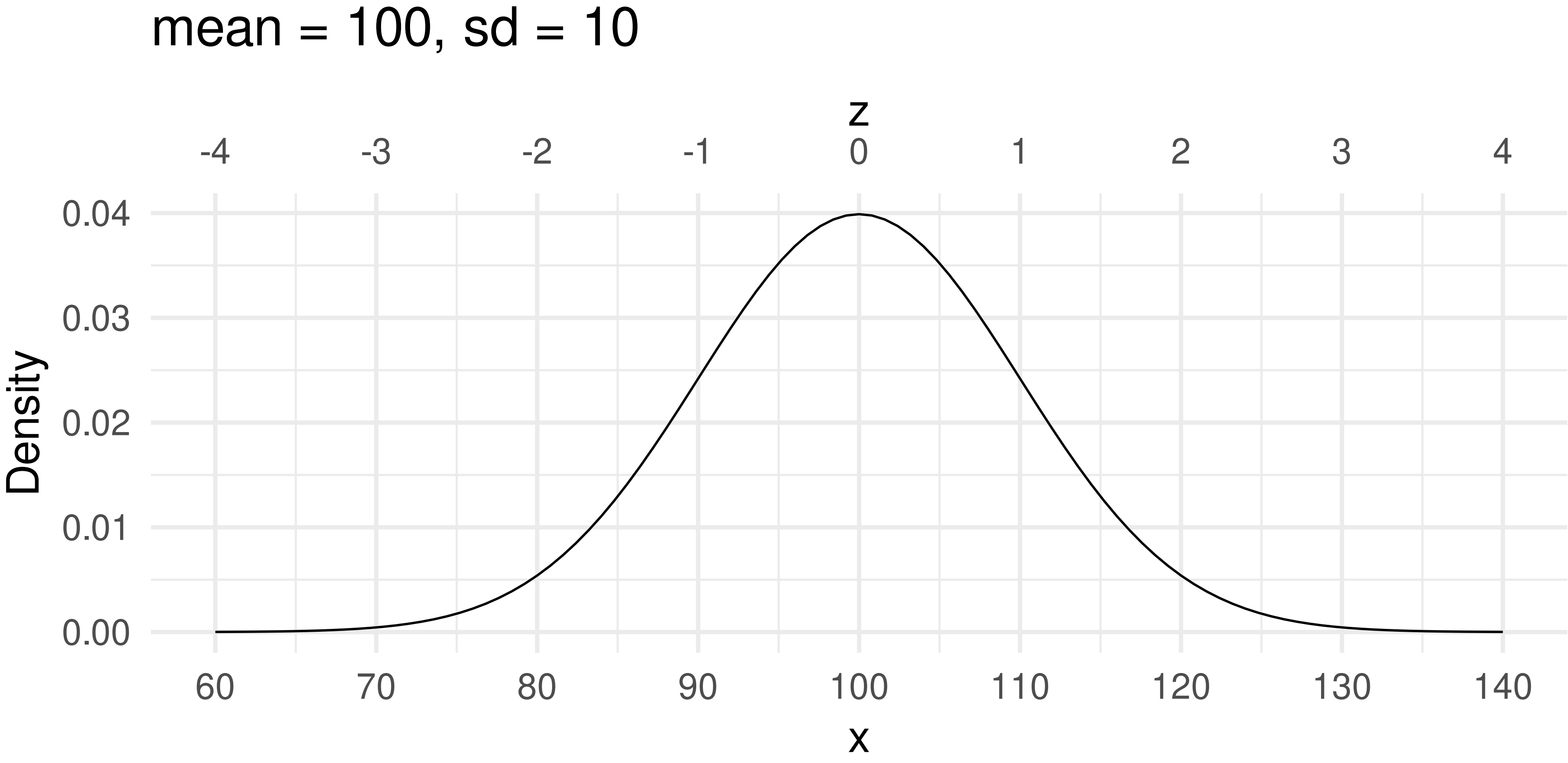
$$z \qquad \frac{x \qquad \mu}{\sigma}$$

Quindi μ semplicemente sposta la distribuzione e σ la riscala. Ogni x può anche essere definito come:

$$x \qquad \mu \qquad z\sigma$$

Normale standard

Se proviamo a visualizzarla il rapporto tra x e z è chiaro:



Normale standard

Quindi un punto z significa che siamo 3 deviazioni standard sotto la media. Nel caso di μ e σ quindi abbiamo che z significa x .

Ovviamente, sappiamo anche che dire z significa sapere che solo il 0.13% ha valori uguali o più piccoli.

```r
pnorm(-3) # di default, mean = 0 sd = 1
```

```
[1] 0.001349898
```

Lavorare in punti z è chiaramente più comodo. Ragioniamo in una nuova unità di misura ovvero la deviazione standard.

Normale standard

Nel caso di dati veri, la normalità può essere assunta ma comunque possiamo calcolare i punti z:

```r
# 5 numeri casuali
x <- rnorm(5, mean = 100, sd = 10)
x
```

```
[1]  98.61862 106.85311 115.01468  91.35710 118.94106
```

```r
# non esattamente i valori veri, solo 10 osservazioni
mean(x)
```

```
[1] 106.1569
```

```r
sd(x)
```

```
[1] 11.38244
```

```r
# punti z empirici, non teorici
z <- (x - mean(x)) / sd(x)
z
```

```
[1] -0.66227385  0.06116381  0.77819612 -1.30023265  1.12314656
```