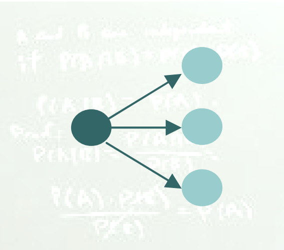
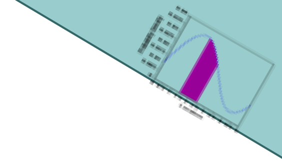
Probabilidade

Distribuição Normal



Distribuição Normal



- Uma variável aleatória contínua tem uma distribuição normal se sua distribuição é:
  - simétrica
  - apresenta (num gráfico) forma de um sino

Função Densidade da Distribuição Normal

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

Distribuição Normal

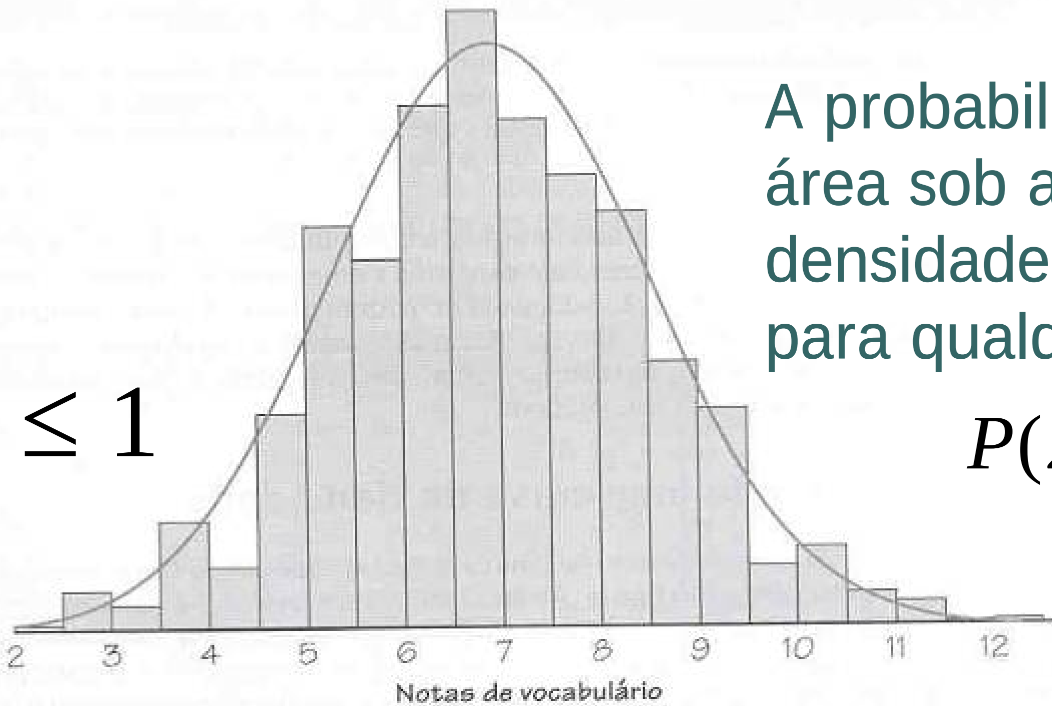
- Quando uma distribuição é contínua, o gráfico de distribuição é uma linha contínua
- Não se visualiza as barras de um histograma, mas freqüências de ocorrências de cada valor de *x* em intervalos infinitesimais
- Forma uma <u>Curva de Densidade de Probabilidade</u> (função *pdf – Probability Density Function*)

Função Densidade da Distribuição Normal

A função densidade da normal (e de qualquer outra variável aleatória contínua) pode ser compreendida como uma extensão natural de um histograma

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

$$0 \leq P(x) \leq 1$$



A probabilidade é a área sob a curva de densidade. Portanto, para qualquer P(x):

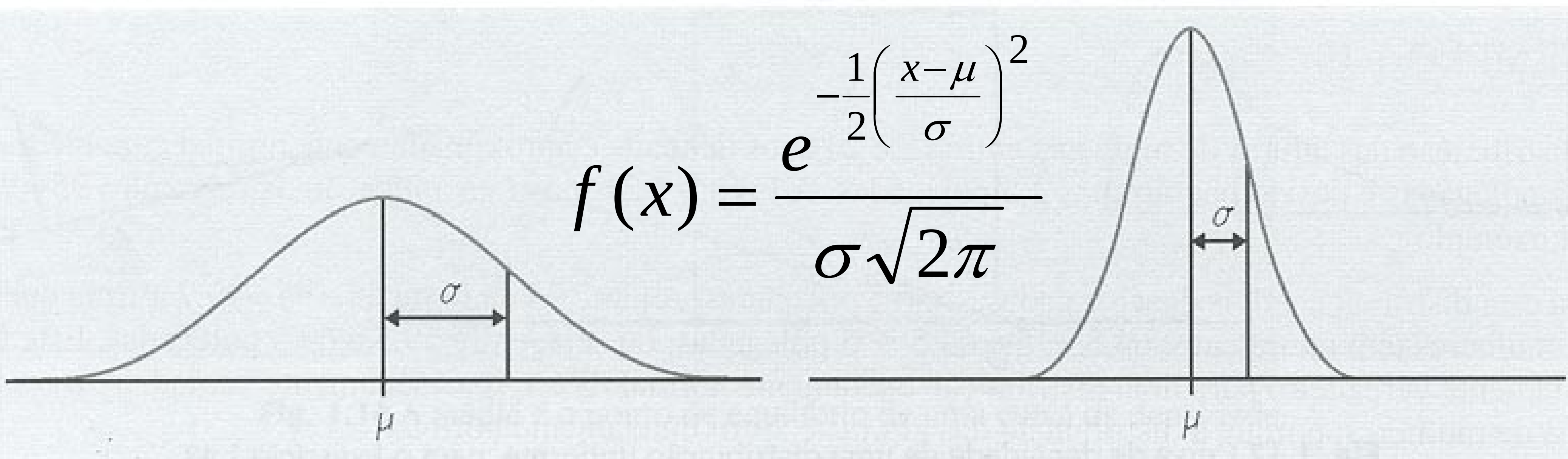
$$P(x) \geq 0$$

Distribuição Normal

Note que a distribuição normal é especificada por dois parâmetros:

μ representa a média populacional, e

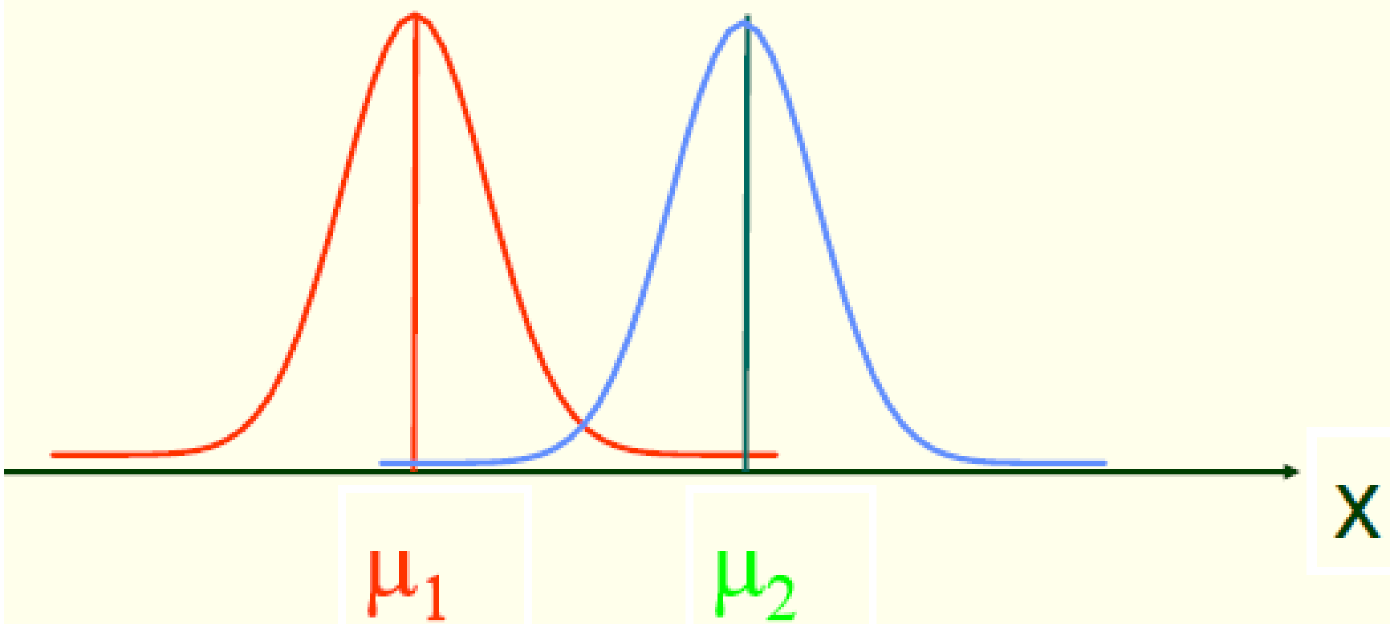
σ representa o desvio-padrão populacional



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

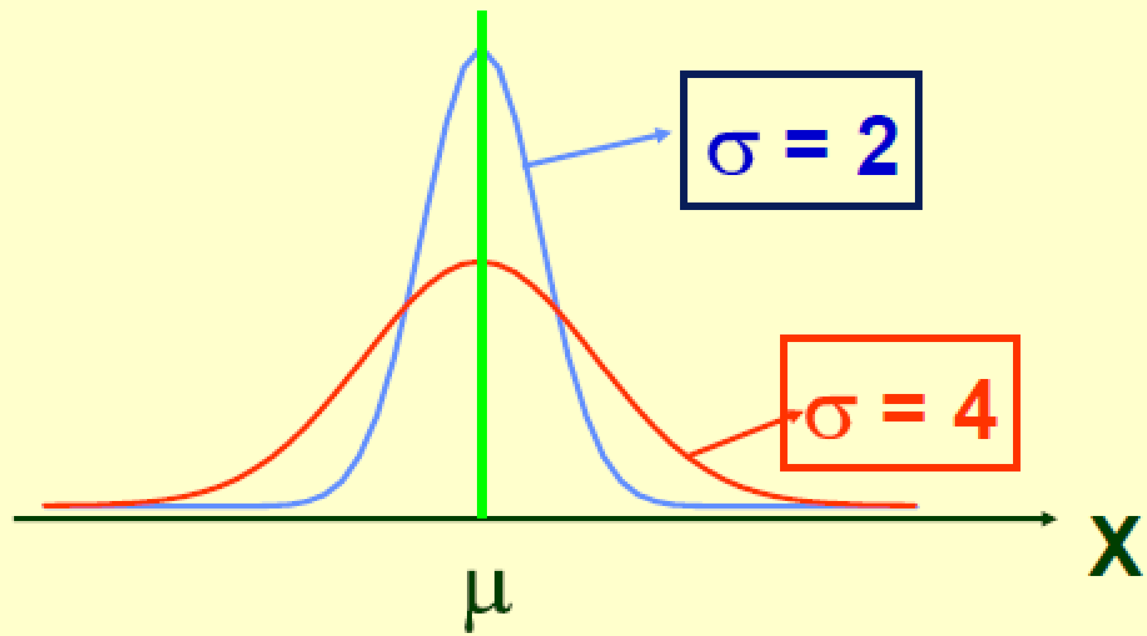
Média e Desvio Padrão

mesmo σ e diferentes μ



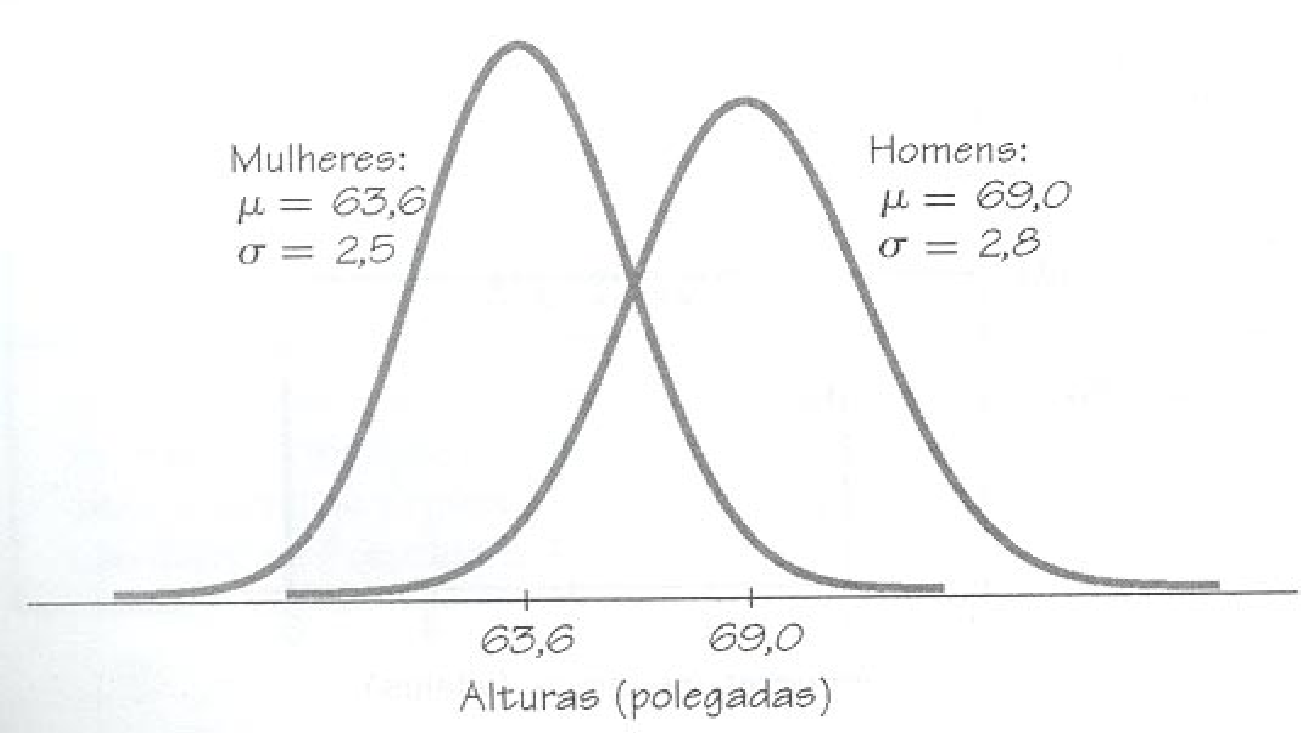
Média e Desvio Padrão

mesmo μ e diferentes σ



Distribuição Normal Padronizada

- Cada par de parâmetros (μ, σ) define uma distribuição normal distinta!
- A figura mostra as curvas de densidade para alturas de mulheres e homens adultos nos EUA



Distribuição Normal Padronizada

- A distribuição normal padronizada tem média e desvio padrão iguais a:

$$\boxed{\mu = 0} \qquad \boxed{\sigma = 1}$$

- **A distribuição normal padronizada facilita os cálculos de probabilidade, evitando o uso da fórmula e projetando qualquer análise mediante utilização de ESCORES (Z)**

$$\boxed{z = \frac{x - \mu}{\sigma}}$$

Distribuição Normal Padronizada

- Se x é uma observação de uma distribuição que tem média μ e desvio-padrão σ, o valor padronizado de x é

$$z = \frac{x - \mu}{\sigma}$$

- Note que o valor padronizado representa o número de desvios-padrão pelo qual um valor x dista da média (*para mais ou para menos*)

Distribuição Normal Padronizada

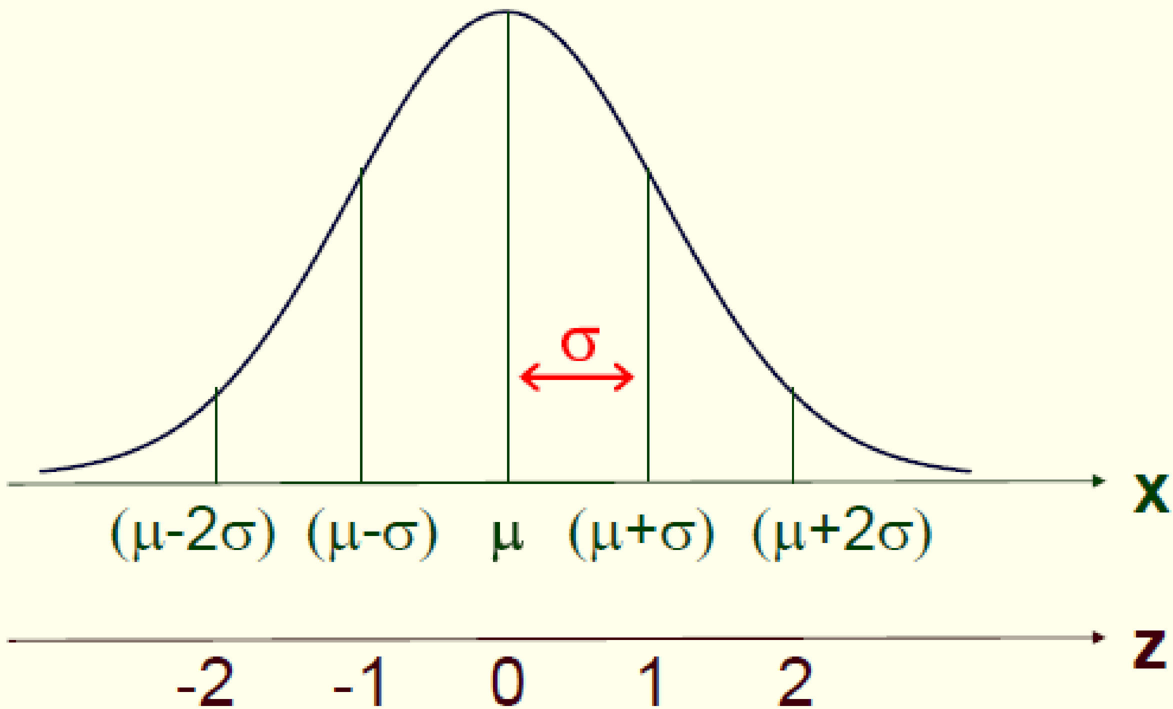
- Ou seja, como a distribuição normal padronizada é aquela que tem média 0 e desvio-padrão 1, ou seja N(0, 1)

- Se uma variável aleatória x tem distribuição normal qualquer N(μ, σ), então a variável padronizada

$$z = \frac{x - \mu}{\sigma}$$

tem distribuição normal

Normal Padronizada



Exemplo

- Um professor de cálculo aplica dois testes diferentes a duas turmas do seu curso. Os resultados foram:

Turma 1:	média = 75	desvio = 14
Turma 2:	média = 40	desvio = 8

- Que nota é relativamente melhor: 82 na turma 1, ou 46 na turma 2?

Distribuição Normal Padronizada

- A estimativa de probabilidades associadas a variáveis aleatórias contínuas envolve o cálculo de áreas sob a curva da densidade.

- O uso da distribuição normal padronizada nos permite calcular áreas sob a curva de uma distribuição normal qualquer, pois as áreas associadas com a normal padronizadas são tabeladas.

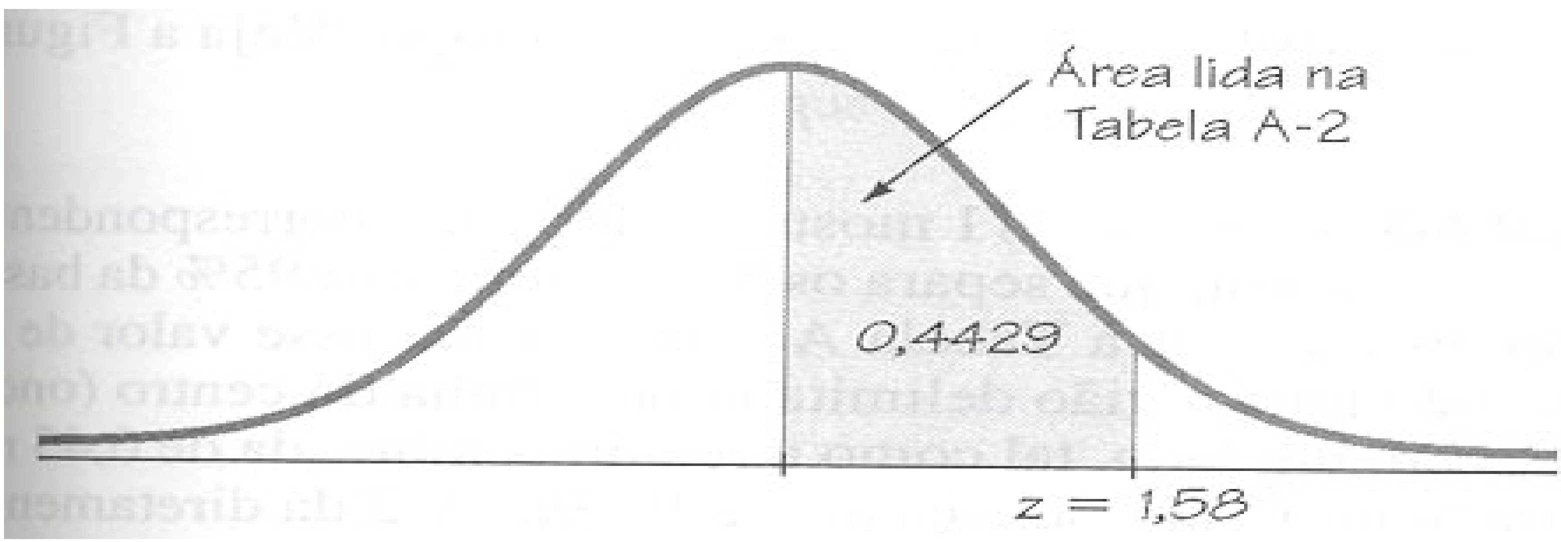
- A Tabela A-2 será usada para os cálculos de probabilidade envolvendo distribuições normais.

Exemplo

- Uma empresa fabrica termômetros que devem acusar a leitura de 0 °C no ponto de congelamento da água. Testes feitos em uma grande amostra desses termômetros revelaram que alguns acusavam valores inferiores a 0 °C e alguns acusavam valores superiores.
- Supondo que a leitura média seja 0°C e que o desvio-padrão das leituras seja 1,00 °C, qual a probabilidade de que, no ponto de congelamento, um termômetro escolhido aleatoriamente marque entre 0 e 1,58 °C ?
- Admita que a freqüência de erros se assemelhe a uma distribuição normal.

Exemplo

- A distribuição de probabilidade das leituras é uma normal padronizada porque as leituras têm $\mu = 0$ e $\sigma = 1$.
- A área da região sombreada, delimitada pela média 0 e pelo número positivo z, pode ser lida na Tabela A-2



Distribuição Normal Padronizada

EXEMPLO

- Portanto, a probabilidade de se escolher aleatoriamente um termômetro com erro entre 0 e 1,58 °C é 44,29 %

- Outra maneira de interpretar este resultado é concluir que 44,29% dos termômetros terão erros entre 0 e 1,58 °C

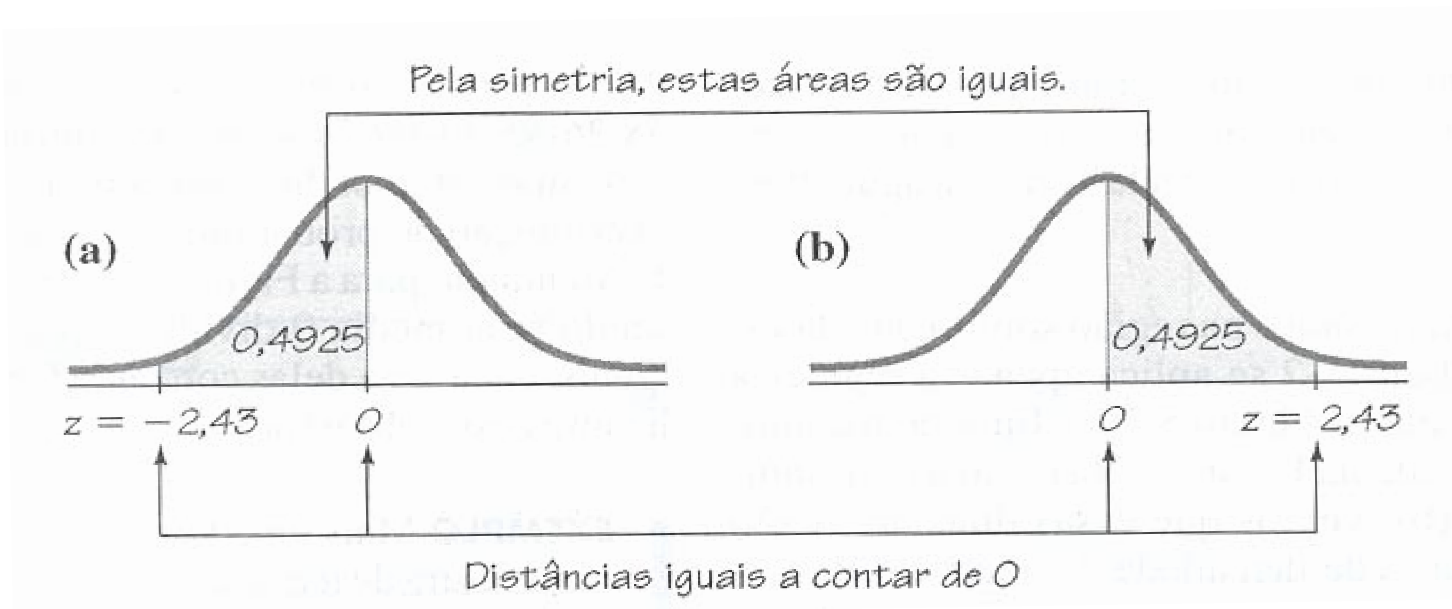
Distribuição Normal Padronizada

EXEMPLO

Com os termômetros do exemplo anterior, determine a probabilidade de se selecionar aleatoriamente um termômetro que acuse (*no ponto de congelamento da água*), uma leitura entre -2,43 °C e 0 °C?

Distribuição Normal Padronizada

- Estamos interessados na região sombreada da Figura (a), mas a Tabela A-2 se aplica apenas a regiões à direita da média (0), como a da Figura (b)
- Podemos ver que ambas as áreas são idênticas porque a curva de densidade é simétrica !

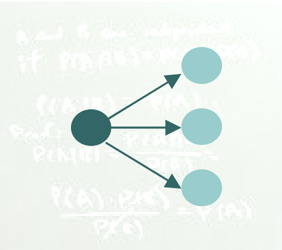


Distribuição Normal Padronizada

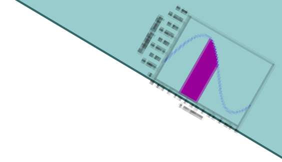
EXEMPLO

- Portanto, a probabilidade de se escolher aleatoriamente um termômetro com erro entre -2,43°C e 0°C é 49,25 %

- Em outras palavras, 49,25% dos termômetros terão erros entre -2,43 °C e 0 °C



Distribuição Normal Padronizada



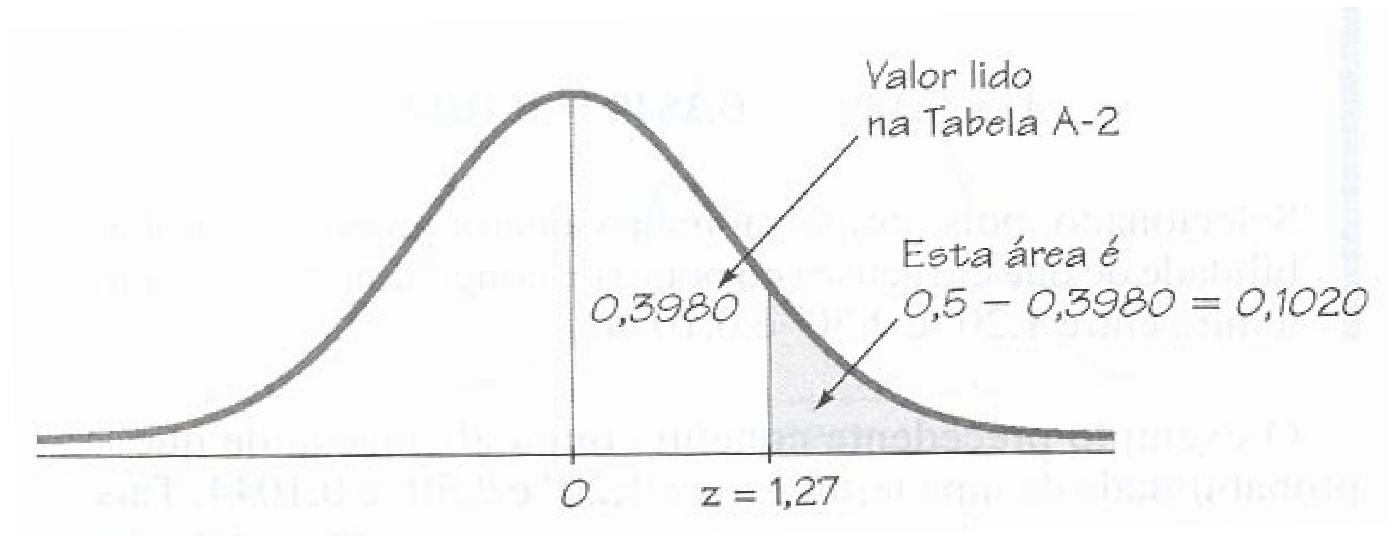
EXEMPLO

Mais uma vez, faremos uma escolha aleatória da mesma amostra de termômetros.

Qual a probabilidade de que o termômetro escolhido acuse (*no ponto de congelamento da água*), uma leitura superior a +1,27 °C ?

Distribuição Normal Padronizada

- A probabilidade de escolher um termômetro que acuse leitura superior a 1,27 °C corresponde à área sombreada da figura
- Se a área total sob a curva da densidade é igual a 1, a área à direita de zero vale metade, isto é, 0,5. Assim, podemos calcular facilmente a área sombreada !



Distribuição Normal Padronizada

EXEMPLO

- Podemos concluir que há uma probabilidade de 10,20% de escolher aleatoriamente um termômetro com leitura superior a +1,27 °C.

- Podemos dizer, ainda, que, em um grande lote de termômetros escolhidos aleatoriamente e testados, 10,20% deles acusarão leitura superior a +1,27 °C

Distribuição Normal Padronizada

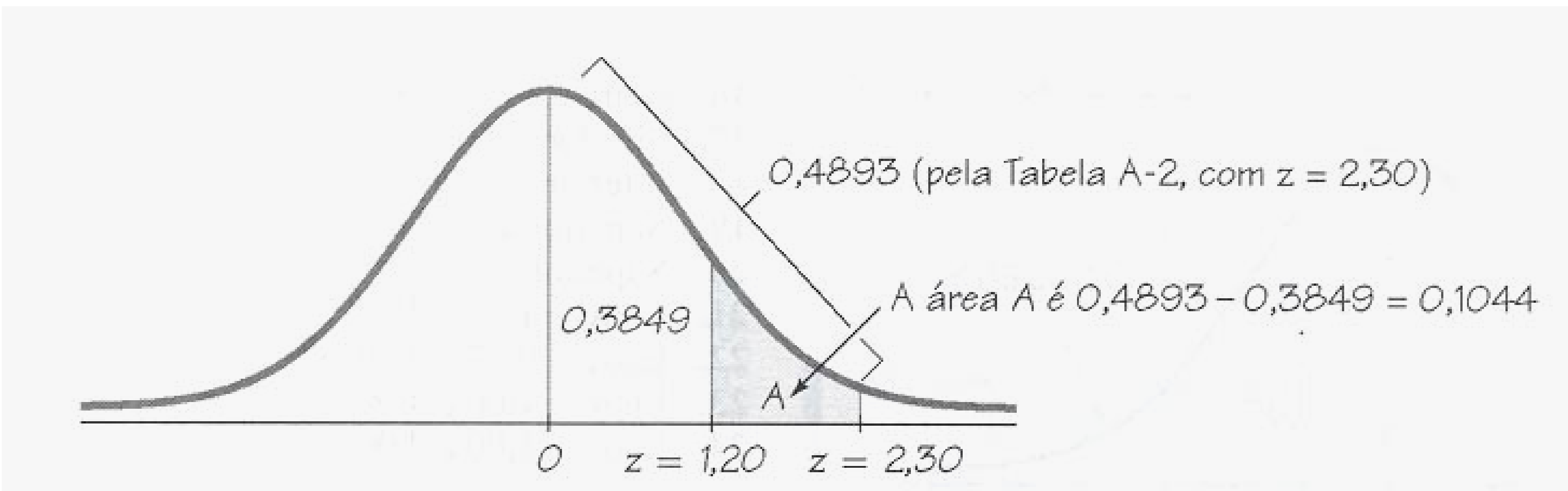
EXEMPLO

De novo, faremos uma escolha aleatória da mesma amostra de termômetros.

Qual a probabilidade de que o termômetro escolhido acuse (*no ponto de congelamento da água*), uma leitura entre 1,20 e 2,30 °C ?

Distribuição Normal Padronizada

- A probabilidade de escolher um termômetro que acuse leitura entre 1,20 e 2,30 °C corresponde à área ombreada da figura
- É fácil perceber que podemos calcular esta área, subtraindo-se a área de 0 até o maior valor (2,30), da área de 0 até o menor valor (1,20), que são lidas na Tabela A-2 !



Distribuição Normal Padronizada

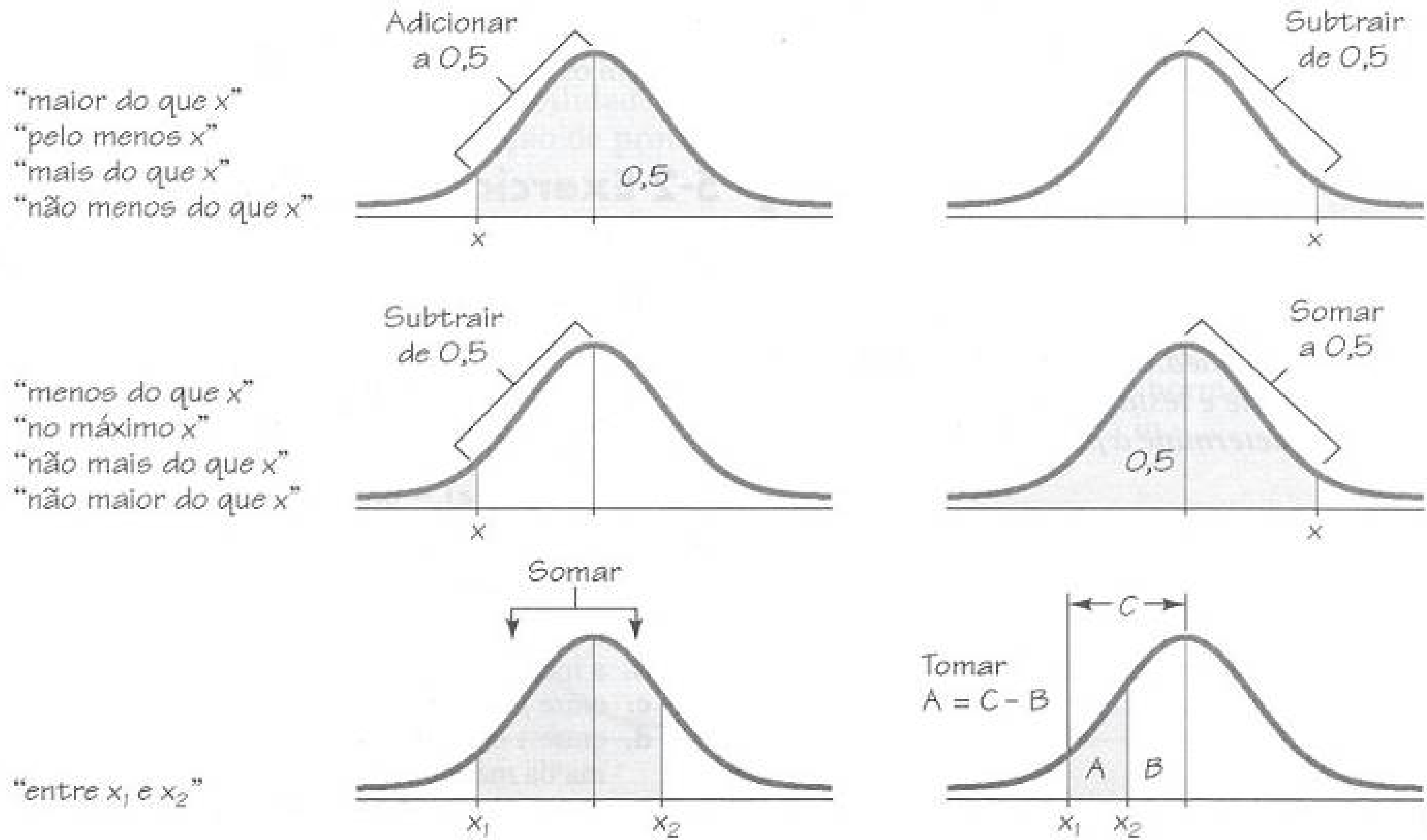
Dos exemplos anteriores, podemos expressar as probabilidades calculadas com a notação seguinte:

$P\ (a < z < b)$ denota a probabilidade de o valor de z estar entre a e b

$P\ (z > a)$ denota a probabilidade de o valor de z ser maior do que a

$P\ (z < a)$ denota a probabilidade de o valor de z ser menor do que a

As figuras abaixo ajudam na interpretação das expressões mais comuns no cálculo de probabilidades:

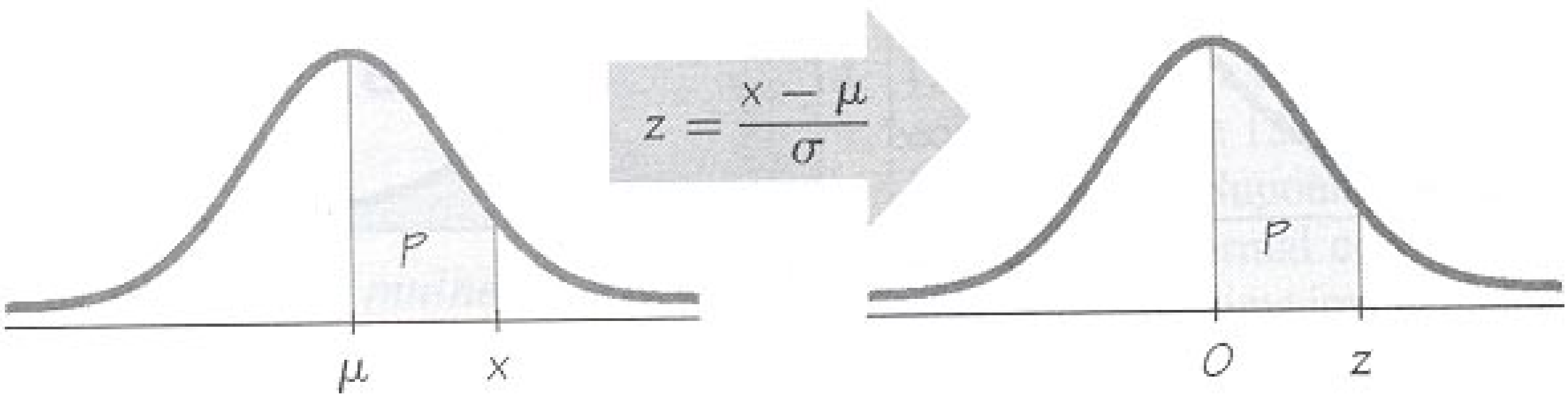


Distribuição Normal Não Padronizada

- Os exemplos feitos com o termômetro não são muito realistas porque a maioria das populações distribuídas normalmente têm média diferente de 0, desvio diferente de 1, ou ambos.
- Como proceder, então, para calcular probabilidades de distribuições normais não-padronizadas ?

Distribuição Normal Não Padronizada

A idéia é utilizar a fórmula dos valores padronizados e *TRANSFORMAR* qualquer distribuição normal dada na normal padronizada, como mostrado abaixo.



Distribuição Normal Não Padronizada

EXEMPLO

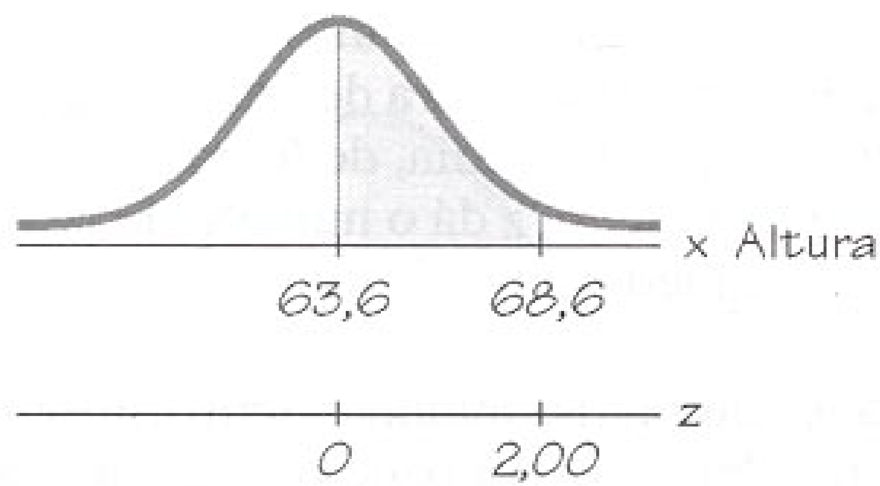
As alturas das mulheres americanas segue uma distribuição normal com média de 63,6” e desvio-padrão de 2,5”.

Selecionada uma mulher americana ao acaso, qual a probabilidade da sua altura estar entre 63,6 e 68,6 polegadas ?

Distribuição Normal Não Padronizada

- Devemos proceder da maneira descrita a seguir:
  - **Trace uma curva normal, assinale a média e outros valores de interesse, e sombreie a região que representa a probabilidade desejada**
  - **Para cada valor x da fronteira da região sombreada, aplique a fórmula para achar o valor padronizado z**
  - **Utilize a Tabela A-2 para**

    achar a área da região sombreada



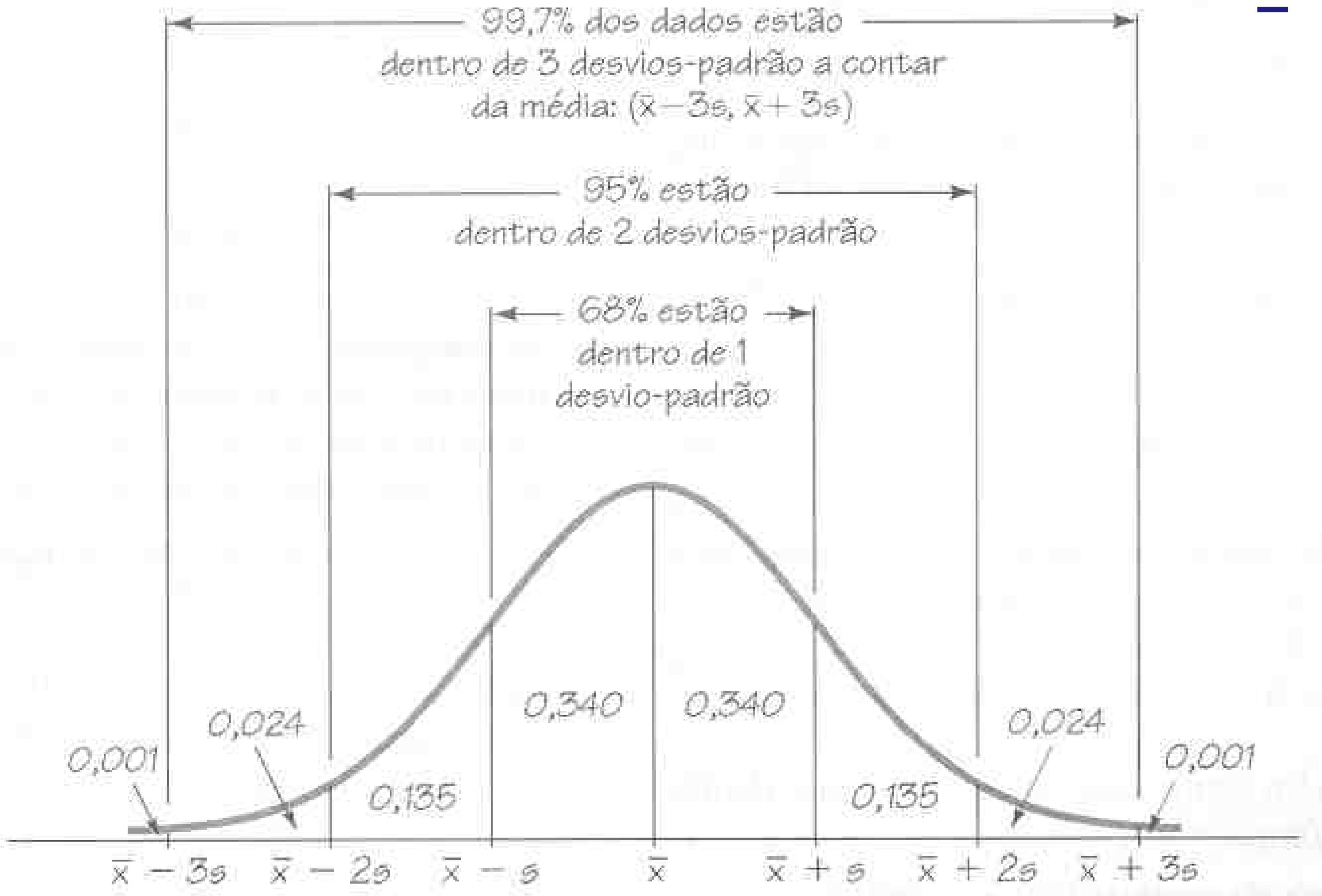
Distribuição Normal Não Padronizada

<u>EXEMPLO</u>

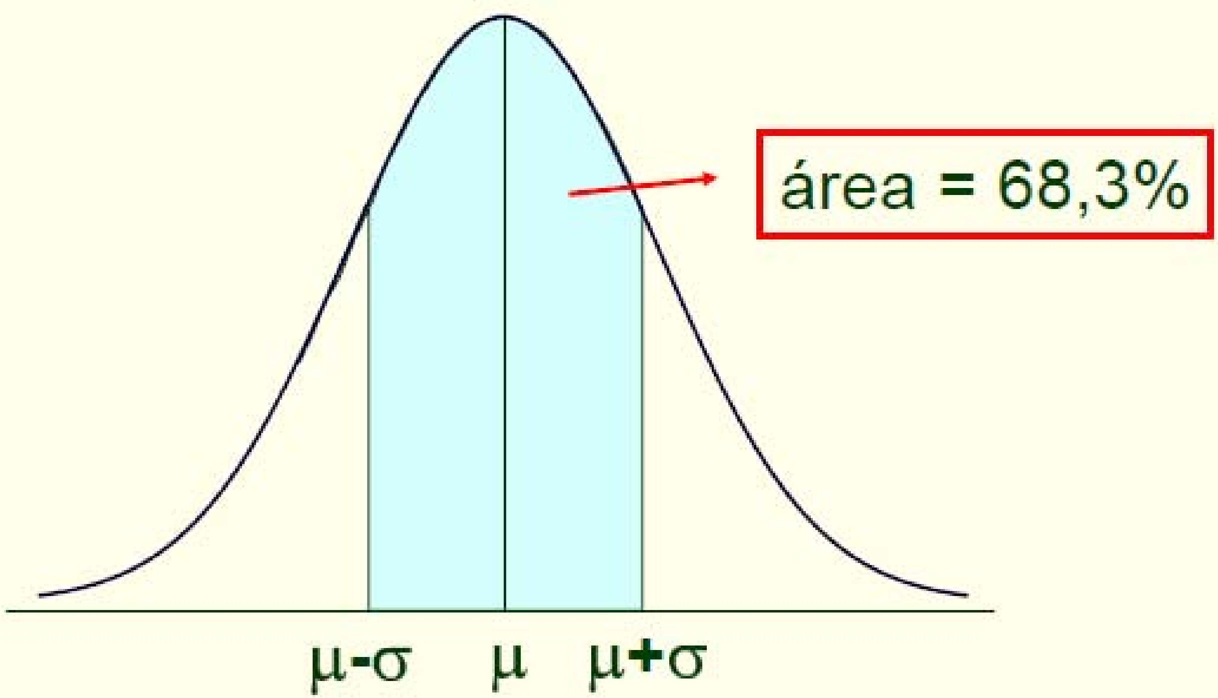
- Há, portanto, uma probabilidade de 0,4772 de escolher uma mulher com altura entre 63,6 pol. e 68,6 pol. Usando a notação, teríamos:
  - $P\,(63{,}6 < x < 68{,}6) \;=\; P\,(0 < z < 2{,}00) \;=$ 47,72%
- Outra forma de interpretar este resultado consiste em concluir que 47,72% das mulheres americanas têm altura entre 63,6 pol. e 68,6 pol.

Regra 68-95-99,7

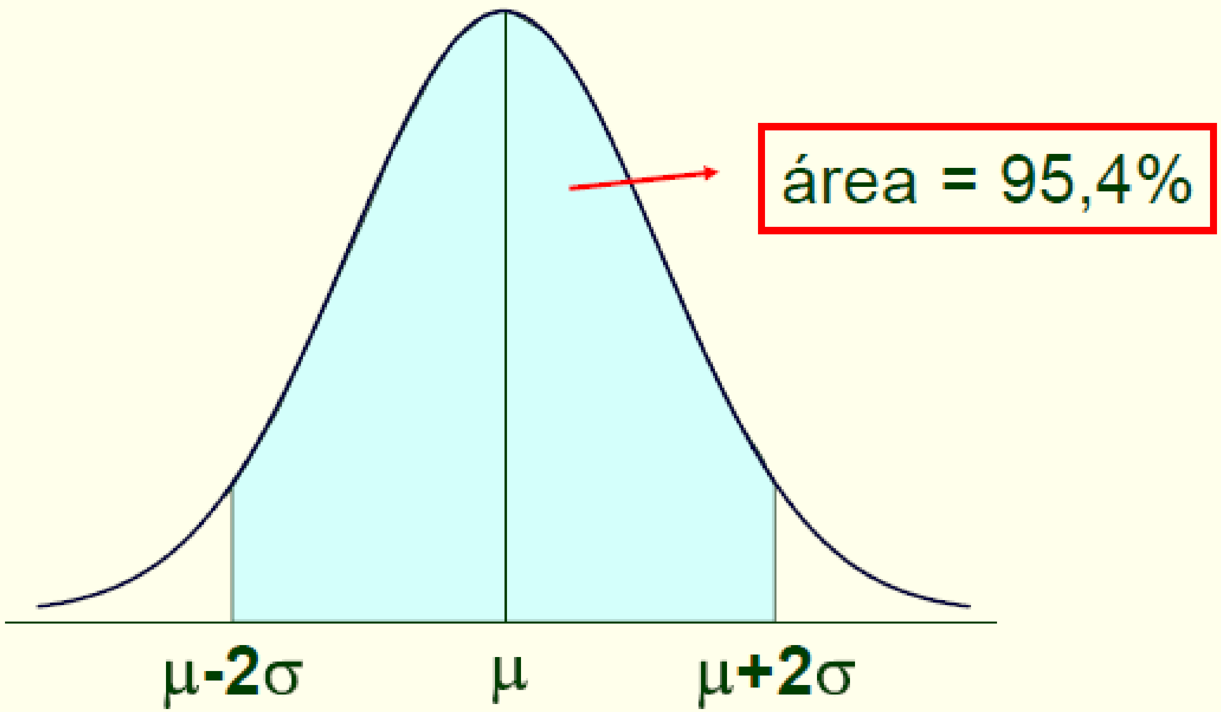
Numa distribuição normal N(μ, σ) ou N($\bar{x}$, s)



Exemplo



Exemplo



Exemplo