

Inferência Estatística e Teste de Hipóteses

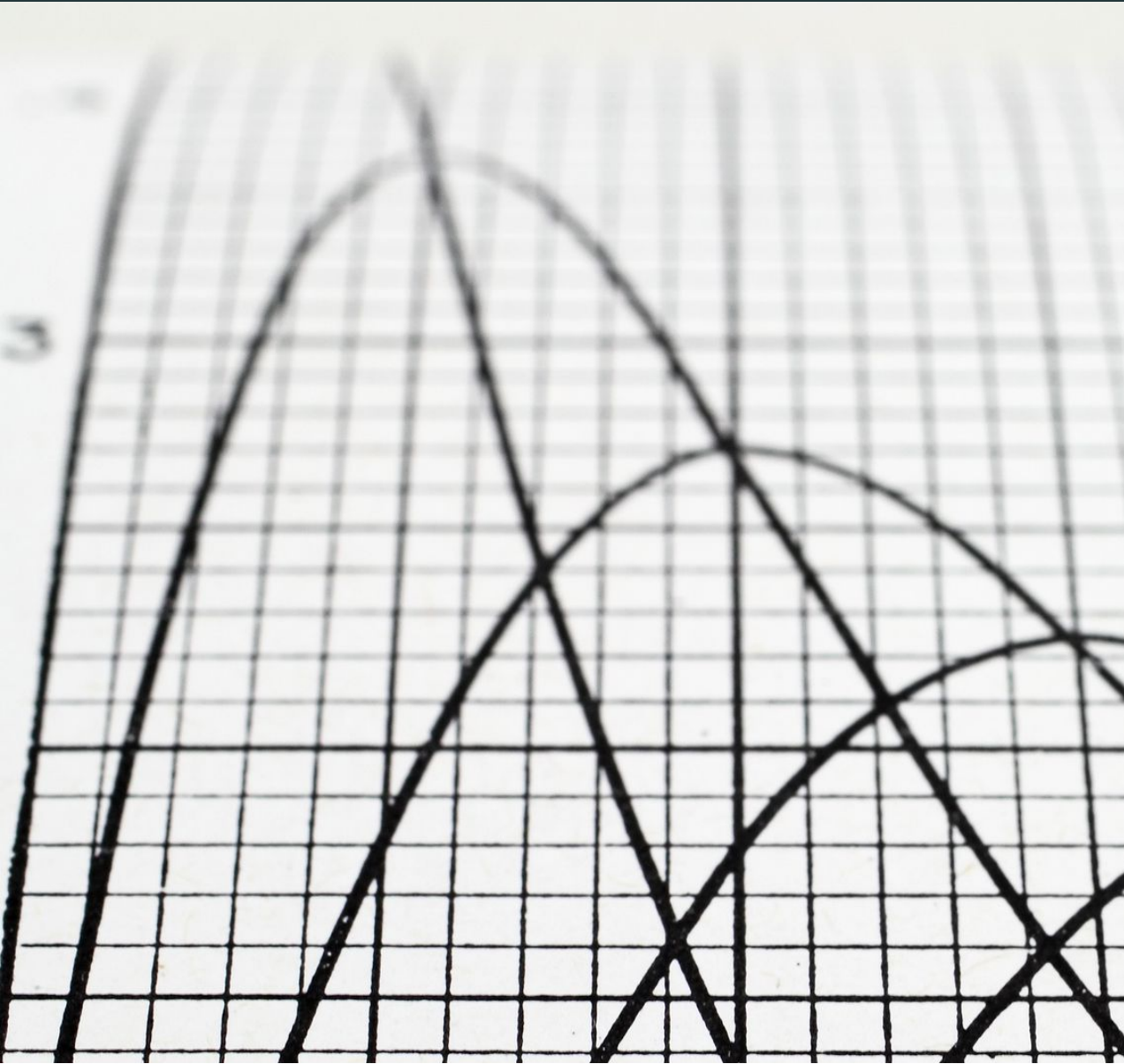
O modelo da distribuição normal

Fabio Cop (fabiocopf@gmail.com)

Instituto do Mar - UNIFESP

Última atualização em 26 de abril de 2022

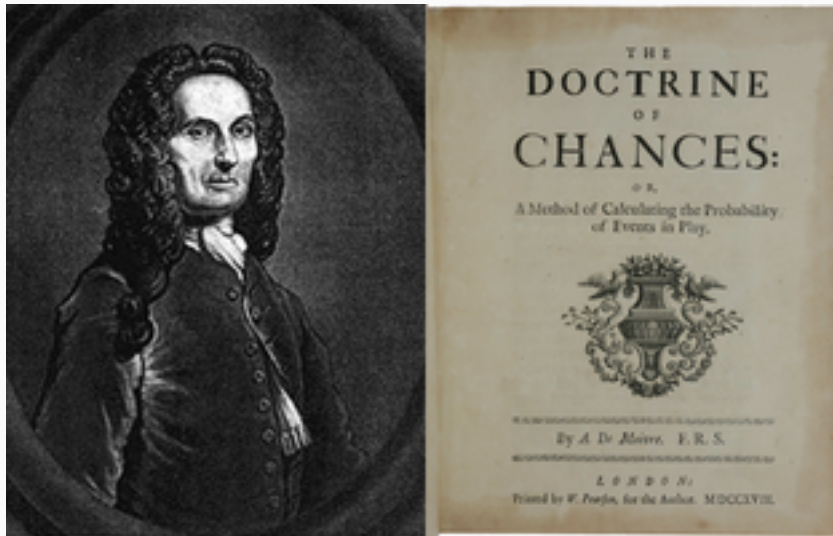
O modelo da distribuição normal



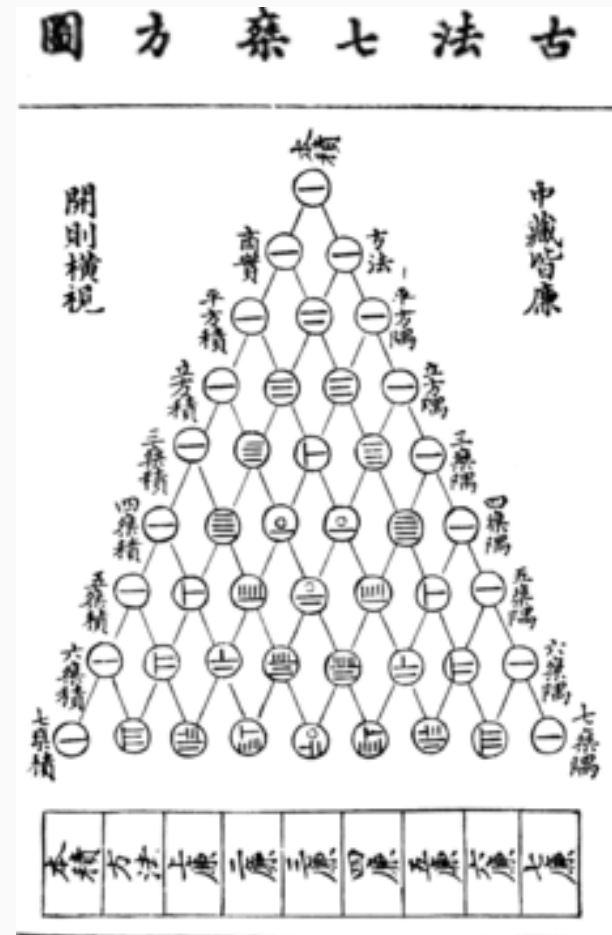
-
1. Um pouco de história
 2. A importância do modelo normal em estatística
 3. O modelo normal de probabilidades
 4. Entendendo a função normal de densidade de probabilidade
 5. Cálculo de probabilidade com a função normal de densidade
 6. A distribuição normal padronizada
 7. Probabilidades em uma distribuição normal padronizada
-

1. Um pouco de história

Alguns atribuem a proposição deste modelo a **Abraham de Moivre**, um matemático Francês que chegou a a distribuição normal como uma aproximação a **distribuição binomial** em seu livro **The Doctrine of Chances** em 1718.



Abraham de Moivre (1667 - 1754)



Triângulo de Yang Hui's de 1261

1. Um pouco de história

A distribuição normal também é conhecida como **distribuição gaussiana**. **Carl Friedrich Gauss** lidou com este modelo para lidar com a distribuição dos erros de observação de um fenômeno no contexto do **Método dos Mínimos Quadrados** em 1823.



1. Um pouco de história

Pierre-Simon Laplace fez contribuições importantes ao modelo e aplicação da distribuição normal. Em 1782 calculou a integral $\int e^{-t^2} = \sqrt{\pi}$ e, em 1810 apresentou o **Teorema Central do Limite** no contexto da agregação de várias observações independentes, contribuições fundamentais para o desenvolvimento da estatística experimental ao longo do século *XX*.



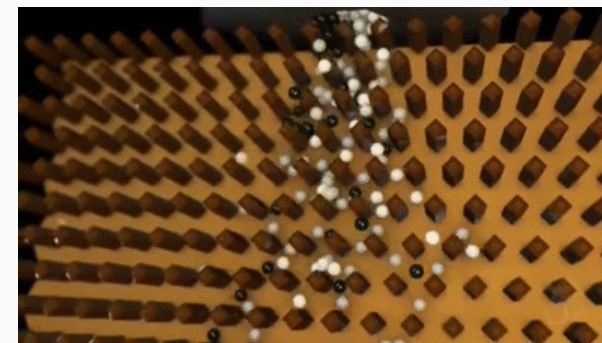
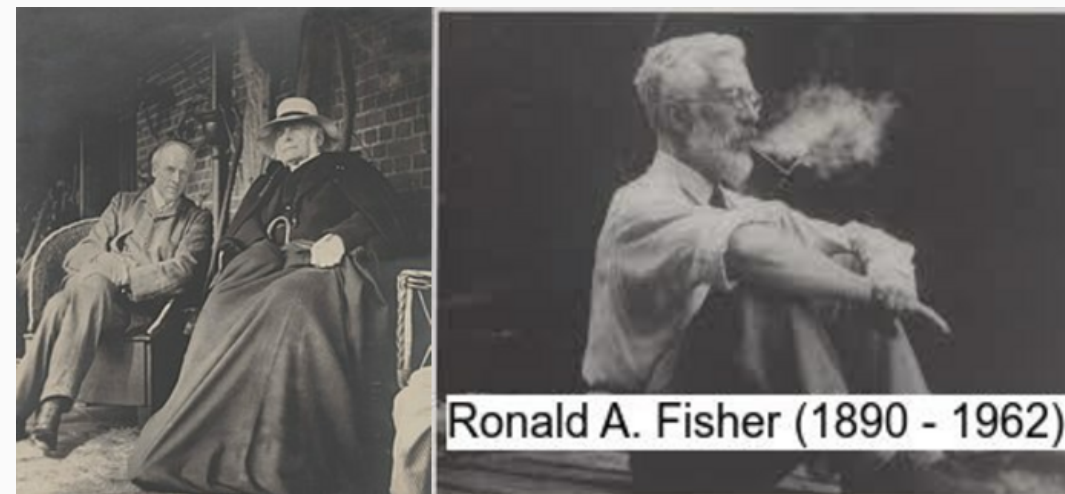
Pierre-Simon Laplace (1748 - 1827)

1. Um pouco de história

Até o final do início do século (XX) a curva era conhecida pelo termo **modelo de Gauss-Laplace**, quando começou a ser referida como **distribuição normal** por **Karl Pearson** e vários outros autores. Pearson foi também o primeiro a escrever o modelo em termos de seu **desvio padrão** σ . Logo em seguida, em 1915 **Ronald Fisher** adicionou o parâmetro de posição (m), isto é a média, expressando o modelo como fazemos atualmente.

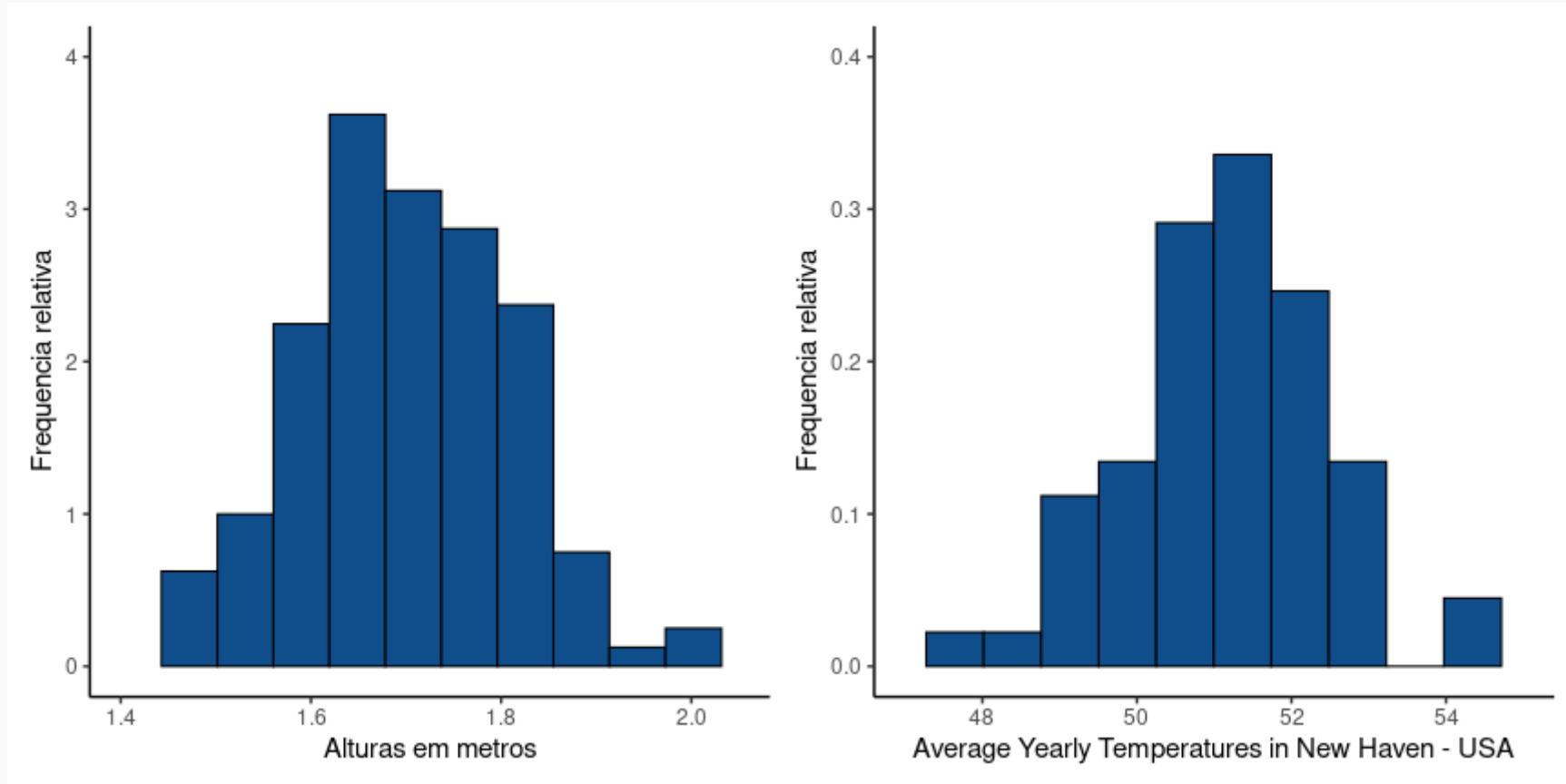
$$df = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

Finalmente, a **distribuição normal padrão**, com média (0) e desvio padrão (1) começou a aparecer nos livros texto por volta de 1950.



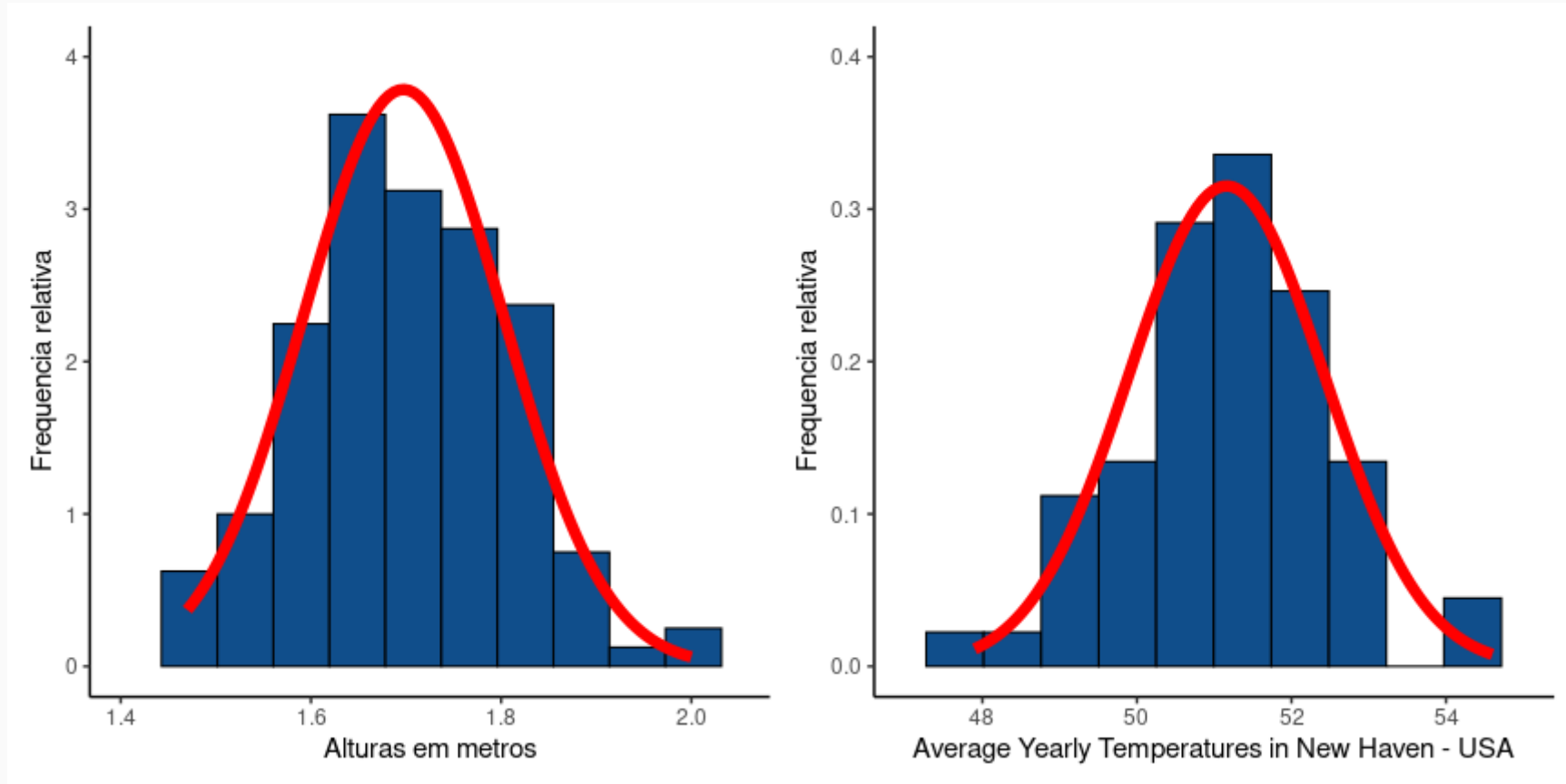
2. A importância do modelo normal em estatística

- A descrição de fenômenos e previsões com o modelo normal.



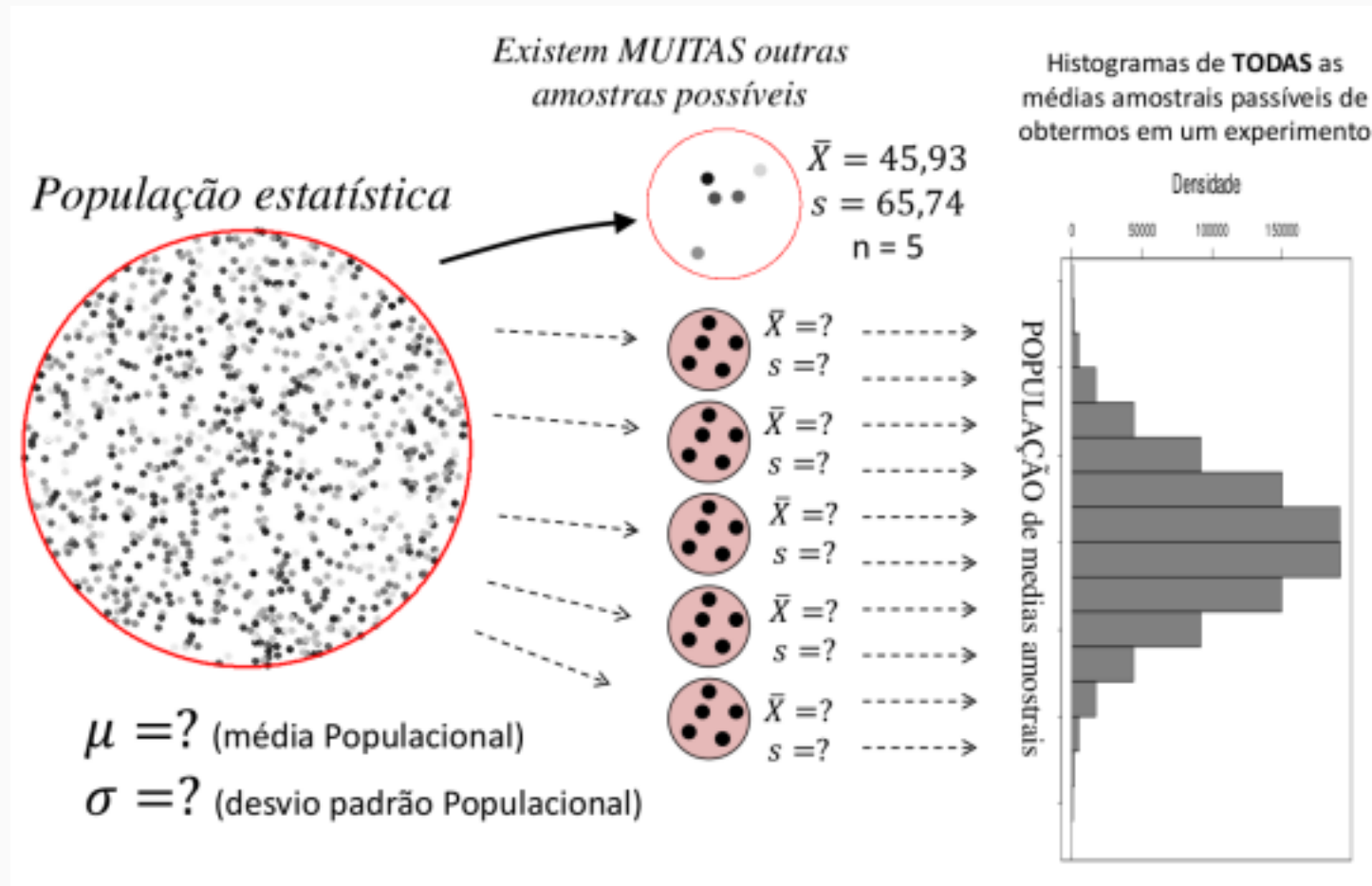
2. A importância do modelo normal em estatística

- A descrição de fenômenos e previsões com o modelo normal.



2. A importância do modelo normal em estatística

- O modelo normal com resultado do processo de amostragem



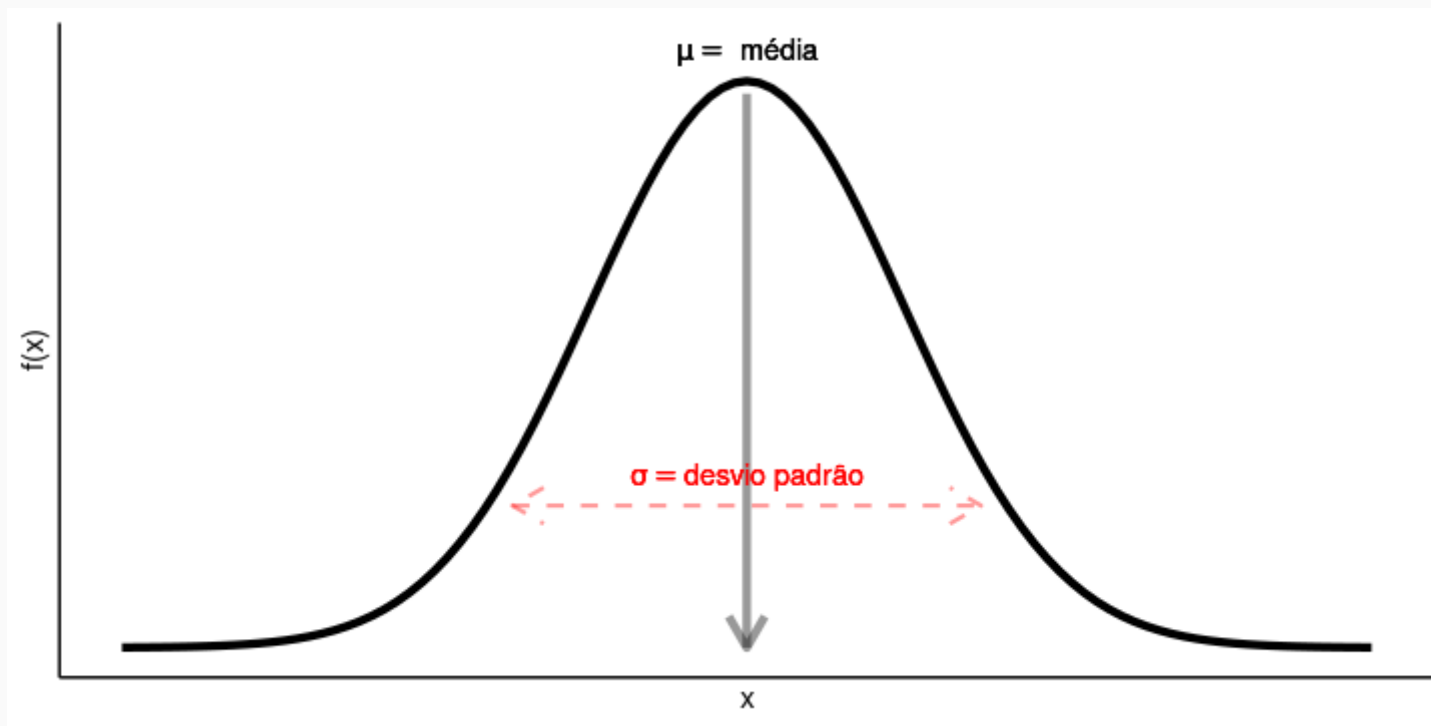
3. O modelo normal de probabilidades

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

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

Lê-se:

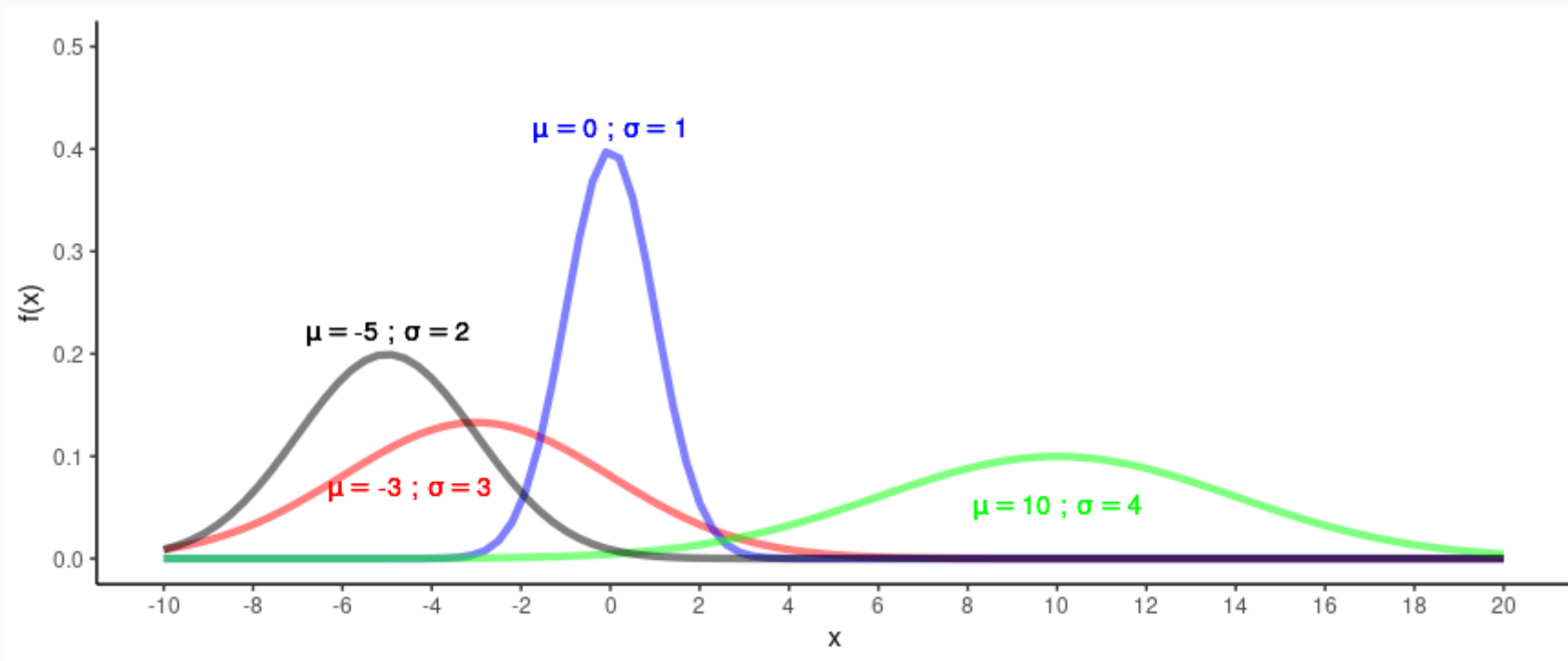
X é uma variável aleatória com distribuição normal, média μ e desvio padrão σ .



3. O modelo normal de probabilidades

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

Podemos alterar o formato da distribuição normal alterando seu parâmetros de posição (média - μ) e de dispersão (desvio padrão - σ).



4. Entendendo a função normal de densidade de probabilidade

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

Encontrando $f(x)$ para pontos específicos.

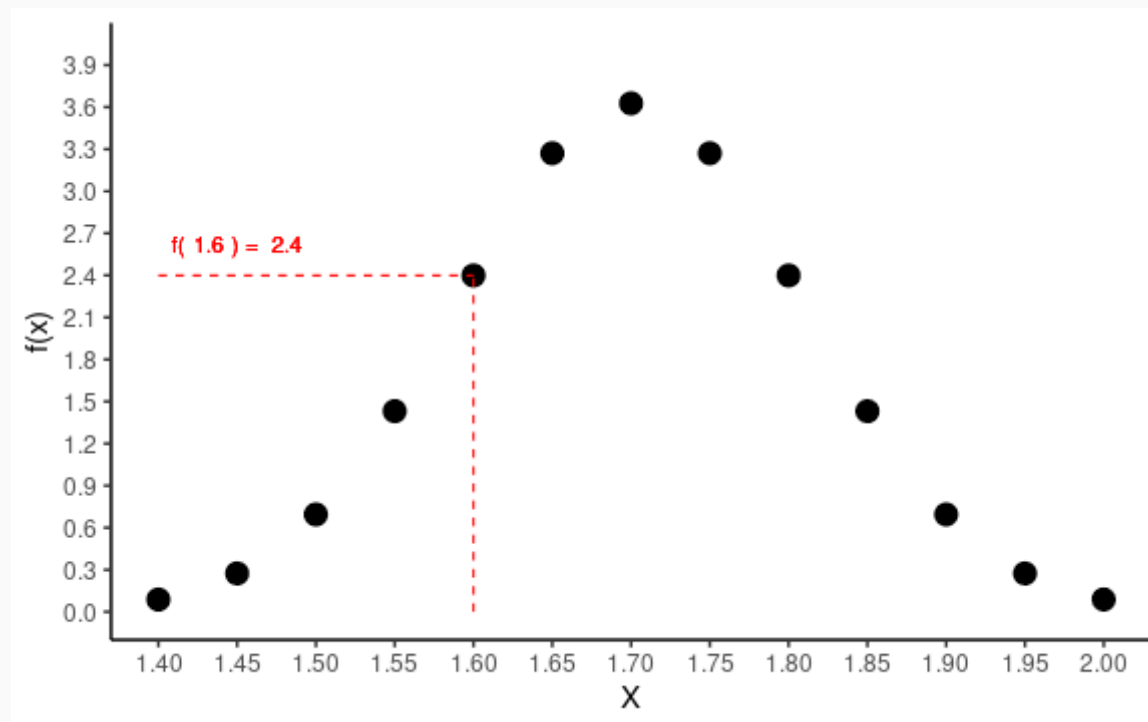
Seja:

$$\mu = 1.7; \sigma = 0.11$$

O ponto $x = 1.6$ tem:

$$f(x) = \frac{1}{\sqrt{(2\pi\sigma^2)}} e^{-\frac{1}{2}\left(\frac{1.6-\mu}{\sigma}\right)^2}$$

$$f(x) = 2.4$$



4. Entendendo a função normal de densidade de probabilidade

Calculo de $f(x)$ no R: a função `dnorm()`

Seja $\mu = 1.7$ e $\sigma = 0.11$

Para $x = 1.5$ fazemos:

```
mu = 1.7
dp = 0.11
x = 1.5
fx = dnorm(x, mean = mu, sd = dp)
fx
```

```
## [1] 0.6945048
```

$$f(x) = \frac{1}{\sqrt{(2\pi 0.11^2)}} e^{-\frac{1}{2} \left(\frac{1.5-1.7}{0.11} \right)^2} = 0.69$$

E para múltiplos valores de x :

```
mu = 1.7
dp = 0.11
x = c(1.4, 1.5, 1.6, 1.7)
fx = dnorm(x, mean = mu, sd = dp)
fx
```

```
## [1] 0.0879777 0.6945048 2.3991470 3.6267480
```

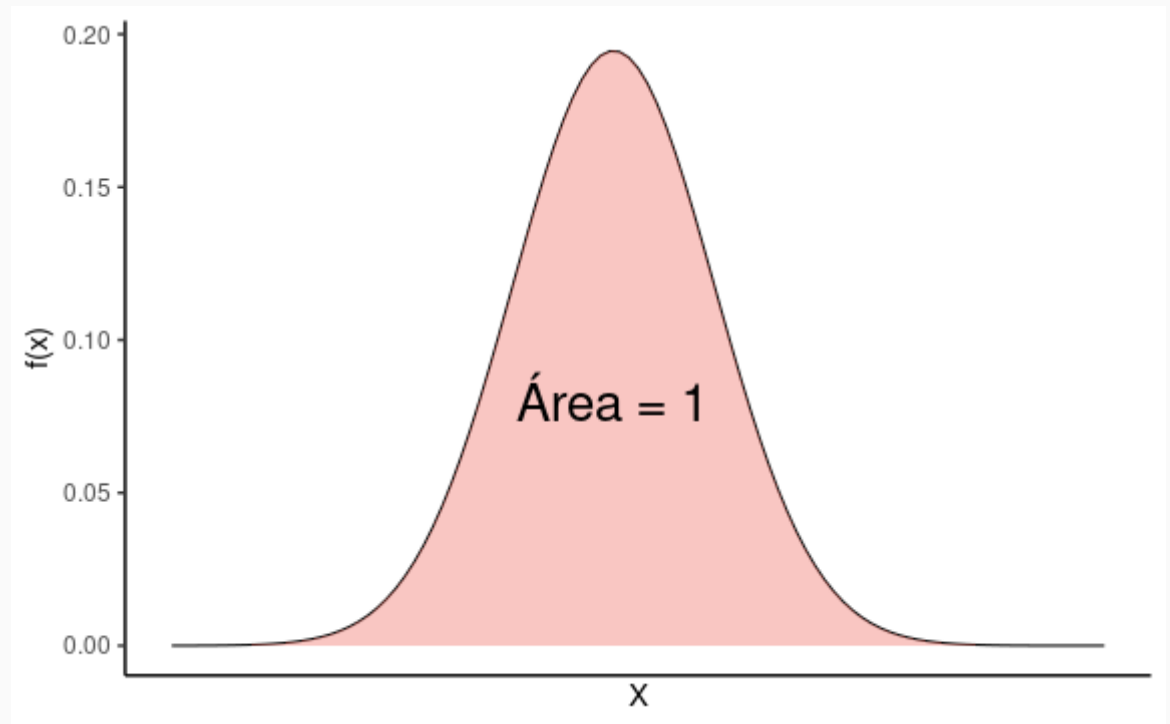

5. Cálculo de probabilidade com a função normal de densidade

A distribuição normal é uma função de **densidade de probabilidade**.

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

A **área** abaixo de $f(x)$ entre $-\infty$ e $+\infty$ é igual a **1**.

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

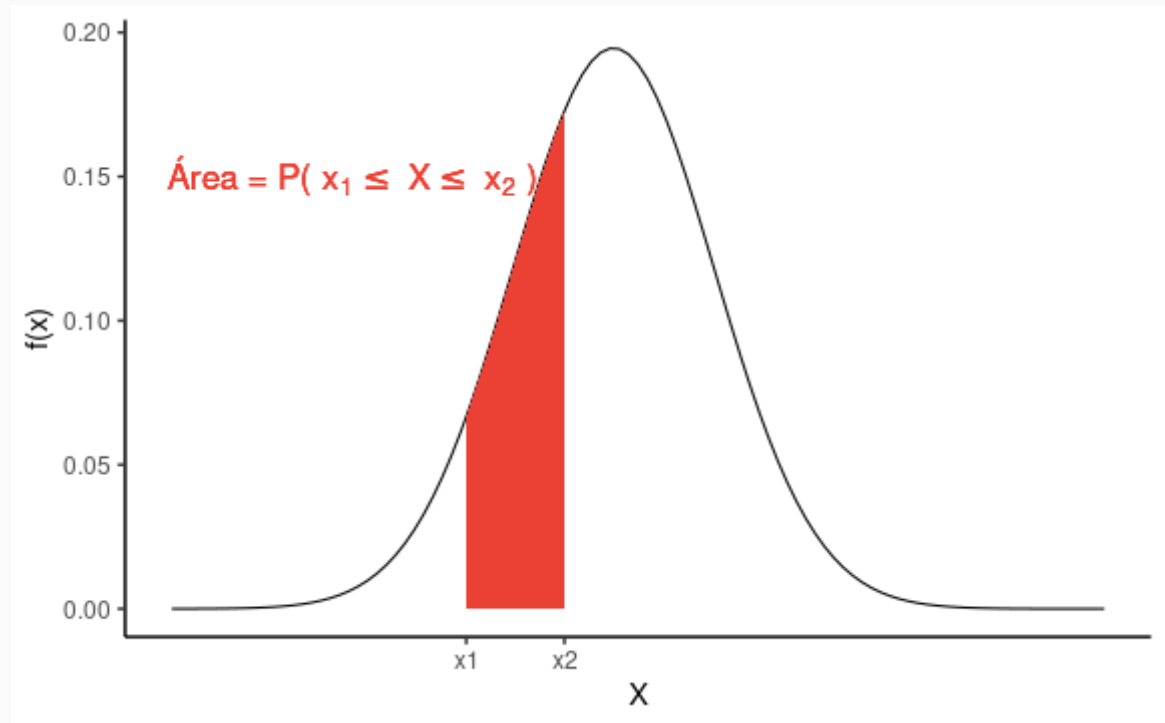


5. Cálculo de probabilidade com a função normal de densidade

A distribuição normal é uma função de **densidade de probabilidade**.

Utilizamos a distribuição normal para encontrar a probabilidade de uma variável aleatória X estar entre x_1 e x_2 , o que corresponde a encontrar a **área** abaixo de $f(x)$.

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x) dx$$



5. Cálculo de probabilidade com a função normal de densidade

Calculando de $P(X)$ no R: a função `pnorm()`

Seja $\mu = 1.7$ e $\sigma = 0.11$

$P(X < 1.5)$:

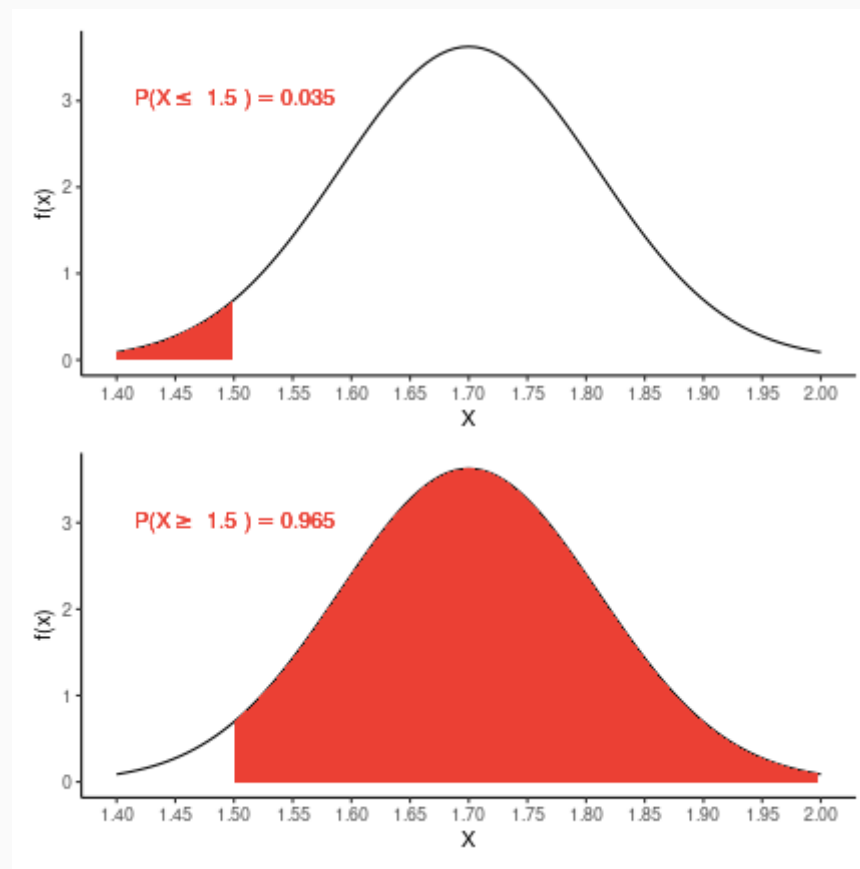
```
mu = 1.7  
dp = 0.11  
pnorm(q = 1.5, mean = mu, sd = dp, lower.tail = TRUE)
```

```
## [1] 0.03451817
```

$P(X > 1.5)$ --> `lower.tail = FALSE`:

```
mu = 1.7  
dp = 0.11  
pnorm(q = 1.5, mean = mu, sd = dp, lower.tail = FALSE)
```

```
## [1] 0.9654818
```



5. Cálculo de probabilidade com a função normal de densidade

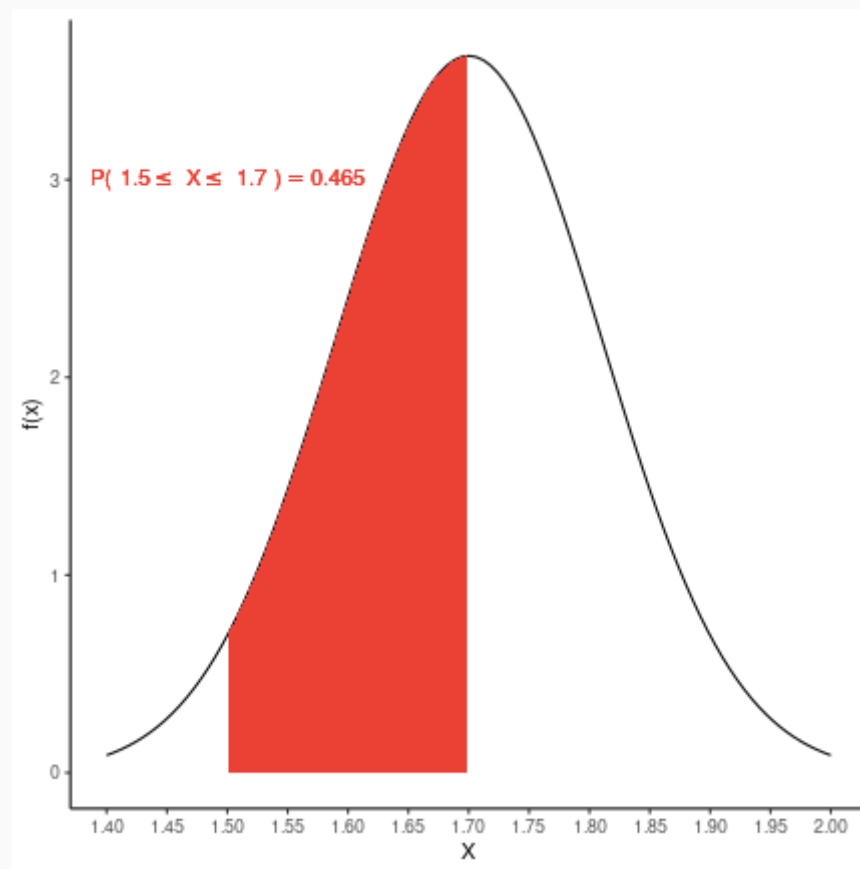
Calculando de $P(X)$ no R: a função `pnorm()`

Seja $\mu = 1.7$ e $\sigma = 0.11$

$$P(1.5 \leq X \leq 1.7) = P(X \leq 1.7) - P(X \leq 1.5)$$

```
P1 = pnorm(q = 1.5, mean = mu, sd = dp)
P2 = pnorm(q = 1.7, mean = mu, sd = dp)
Pintervalo = P2 - P1
Pintervalo
```

```
## [1] 0.4654818
```



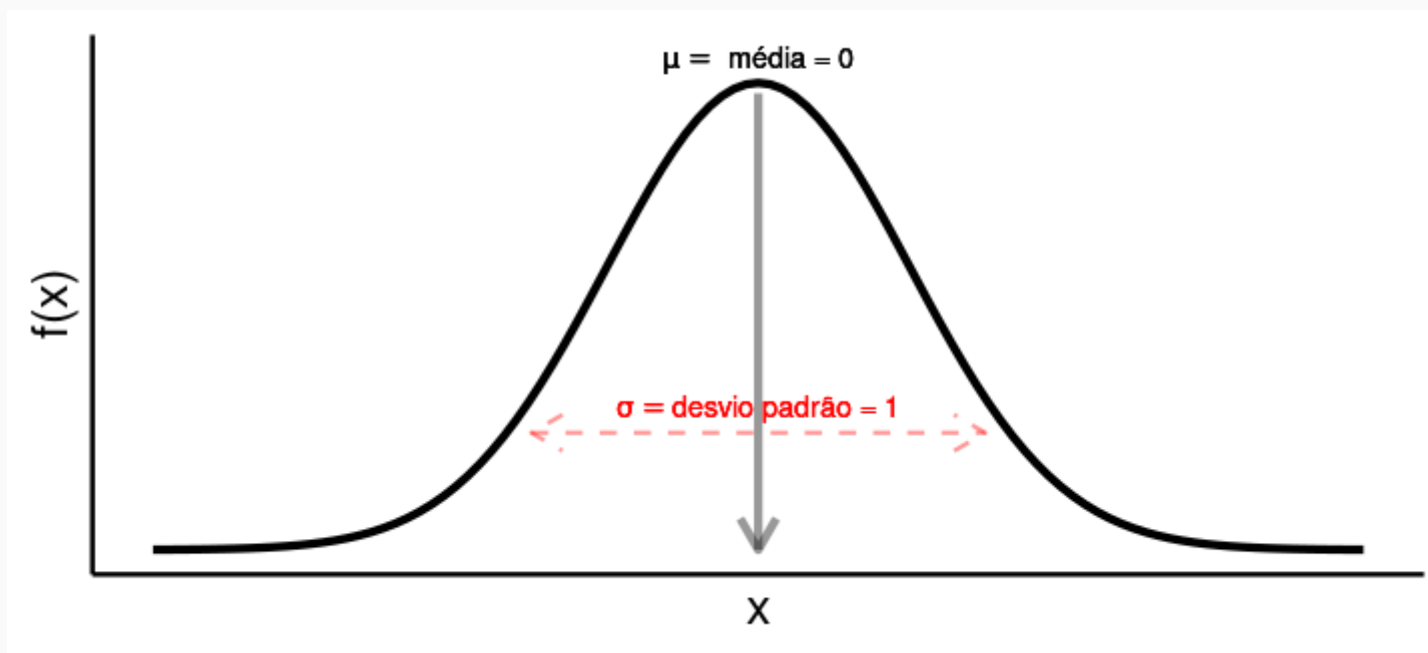
6. A distribuição normal padronizada

Historicamente, o cálculo da área na distribuição normal foi conduzido com o uso da distribuição normal **padronizada**.

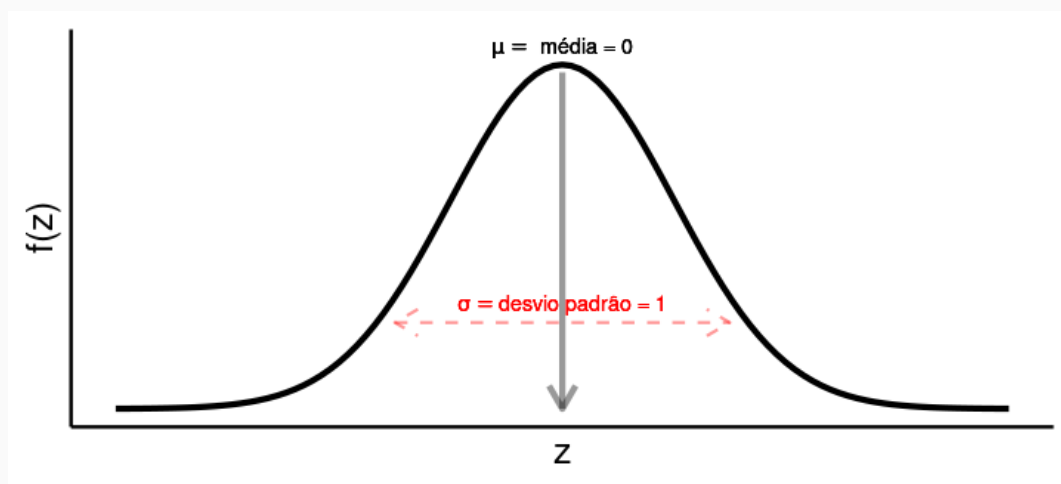
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

$$X \sim \mathcal{N}(0, 1)$$

X é uma variável aleatória com distribuição normal **padronizada** se, média $\mu = 0$ e desvio padrão $\sigma = 1$.



6. A distribuição normal padronizada: a tabela Z



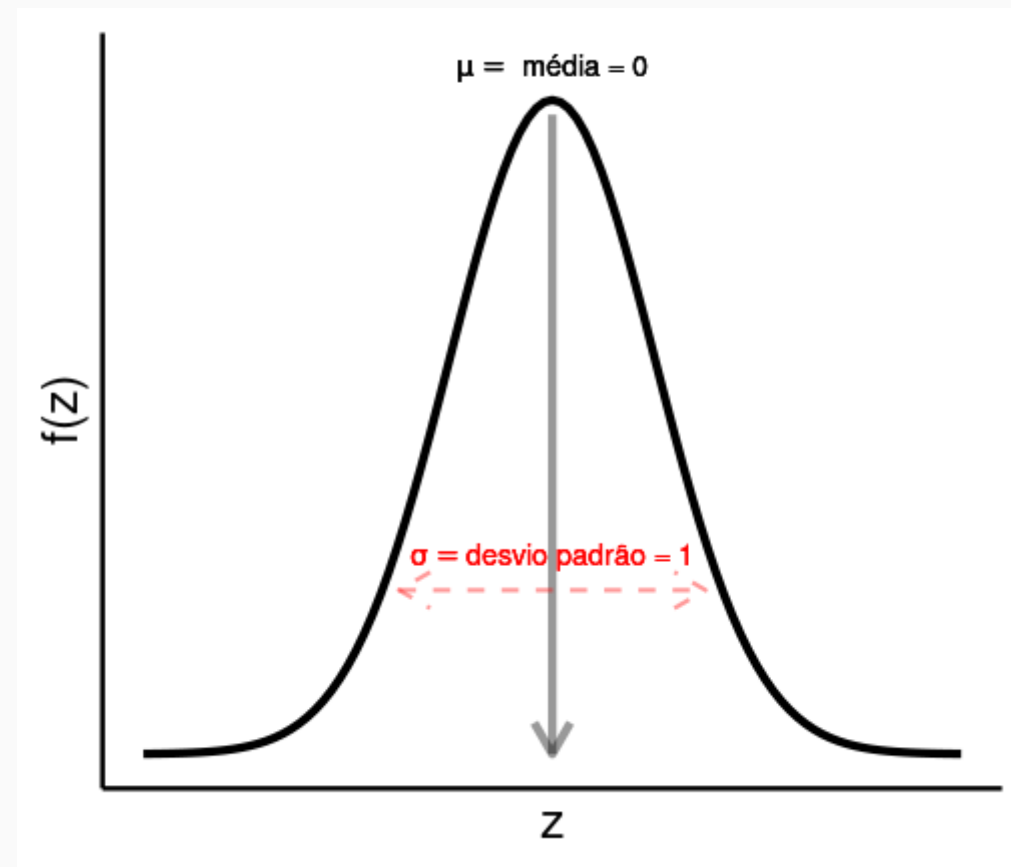
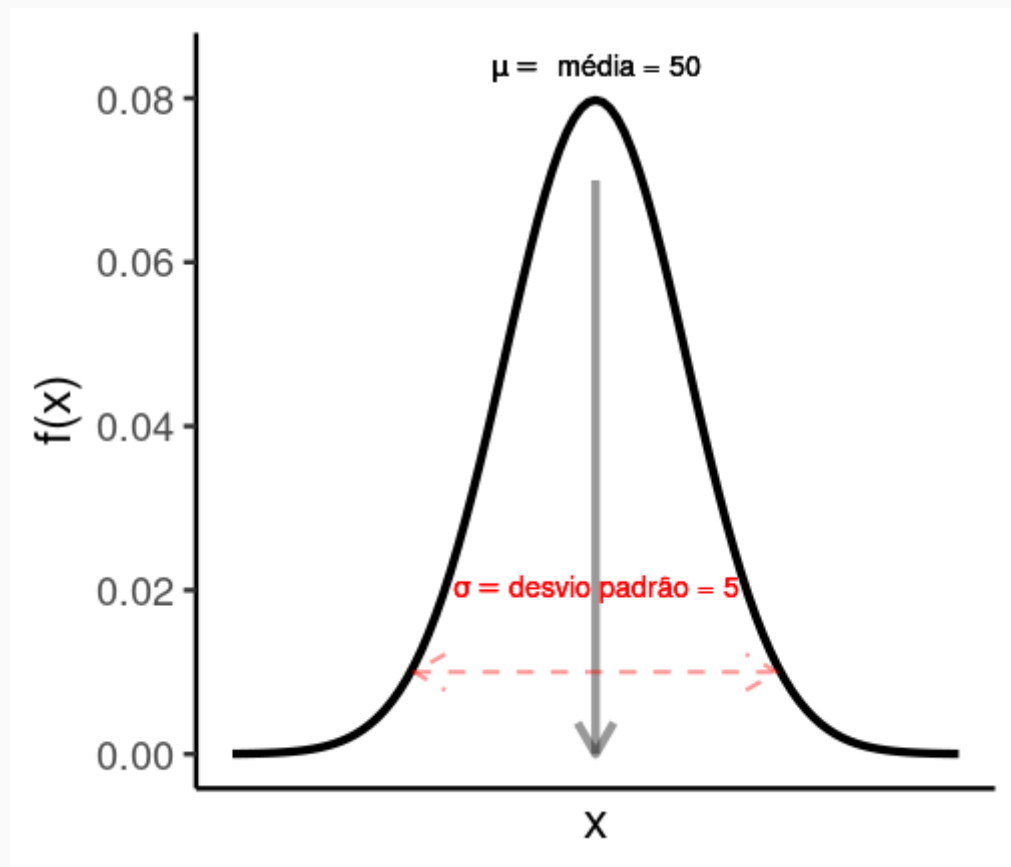
Distribuição Normal Padronizada (Z): área acumulada à esquerda, escores Negativos

Segunda casa decimal

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.5	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002
-3.4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002
-3.3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
-3.2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3.1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3.0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2.9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2.8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2.7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2.6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2.5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2.4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2.3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2.2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2.1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2.0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1.9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1.8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1.7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1.6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455

6. A distribuição normal padronizada: a transformação Z

$$z_i = \frac{x_i - \mu}{\sigma}$$



6. A distribuição normal padronizada: a transformação Z

$$z_i = \frac{x_i - \mu}{\sigma}$$

Seja $\mu = 1.7$ e $\sigma = 0.11$

A observação $x = 1.5$ corresponde a:

$$z_i = \frac{x_i - \mu}{\sigma} = \frac{1.5 - 1.7}{0.11} = -1.82$$

Em uma distribuição normal com $\mu = 1.7$ e $\sigma = 0.11$, a observação $x = 1.5$ está **1.82** desvios padrões abaixo da média.

Seja $\mu = 1.7$ e $\sigma = 0.11$

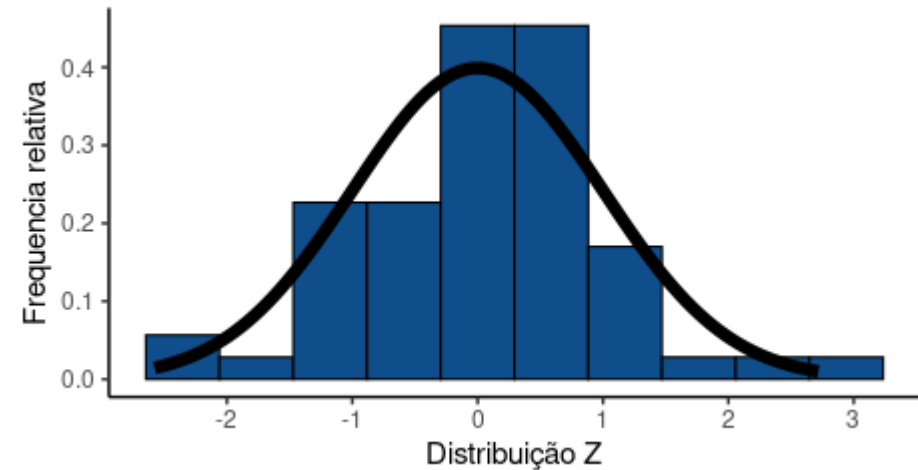
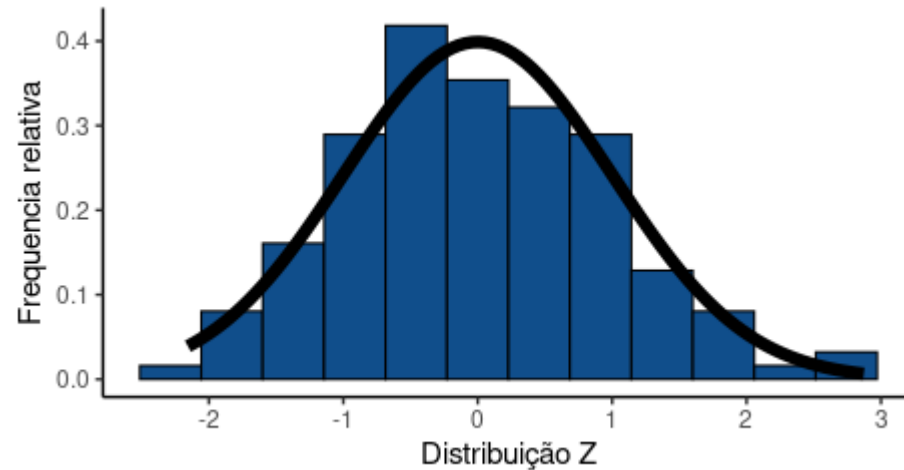
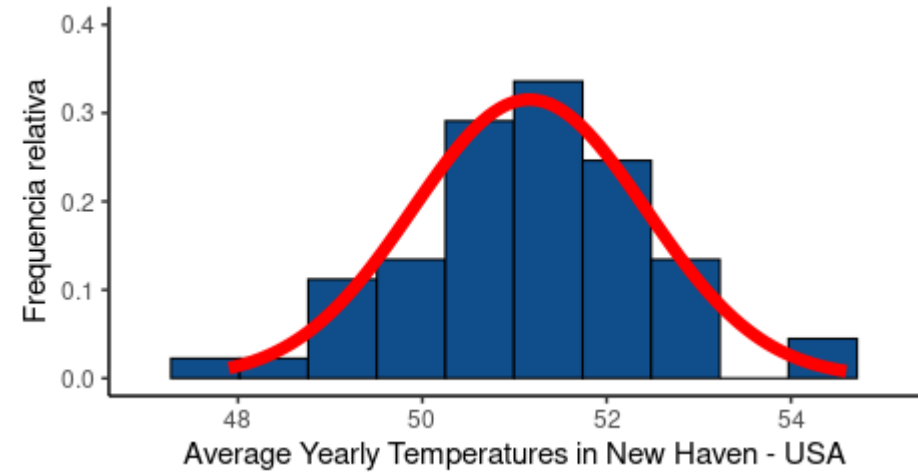
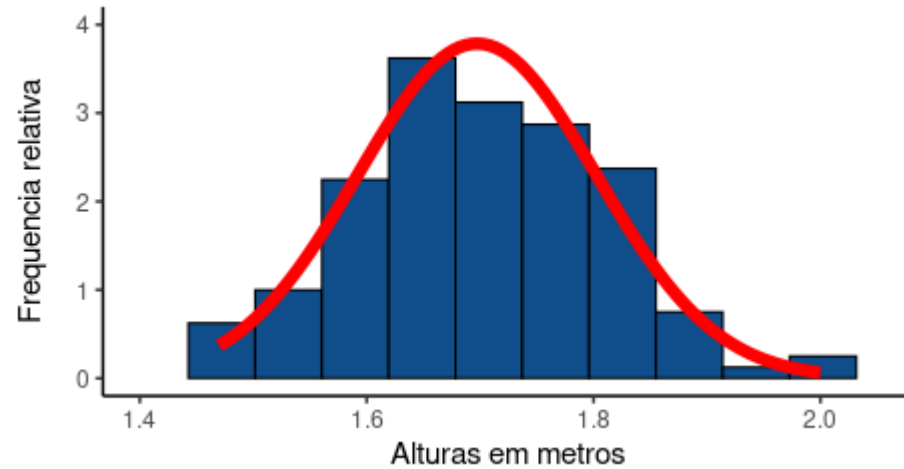
A observação $x = 1.9$ corresponde a:

$$z_i = \frac{x_i - \mu}{\sigma} = \frac{1.9 - 1.7}{0.11} = 1.82$$

Em uma distribuição normal com $\mu = 1.7$ e $\sigma = 0.11$, a observação $x = 1.9$ está **1.82** desvios padrões acima da média.

6. A distribuição normal padronizada

Após a transformação Z nos exemplos sobre altura dos alunos e chuva mensal temos:



6. Probabilidades em uma distribuição normal padronizada

Limites conhecidos de probabilidade

