ACH3584 - Estatística II

Aulas 1 e 2: Variáveis aleatórias contínuas e a distribuição Normal

Alexandre Ribeiro Leichsenring
alexandre.leichsenring@usp.br

Organização

1. Variáveis Aleatórias Contínuas

2. A Distribuição Normal

Organização

1. Variáveis Aleatórias Contínuas
2. A Distribuição Normal

Variáveis Aleatórias Contínuas

- Variáveis aleatórias contínuas são aquelas que assumem valores no conjunto dos números reais.
- São variáveis como salário, tempo de utilização de um equipamento, peso, preço de produtos.

Exemplos

Variável Aleatória	Valores possíveis
Peso	74,5 ; 89,1 ; 54,8 ; ...
Horas	1,5 ; 3,5 ; 2 ; ...
Dinheiro	23,75 ; 43,90 ; 21,10 ; ...

Lembrando do caso discreto

Uma variável aleatória X, do tipo **discreto**, está bem caracterizada se indicarmos os possíveis valores que ela pode assumir:

$$x_1, x_2, \ldots, x_n, \ldots$$

e as respectivas probabilidades:

$$p(x_1), p(x_2), \ldots, p(x_n), \ldots$$

Ou seja, se conhecermos sua função de probabilidade:

$$(x, p(x))$$

Também usamos a notação:

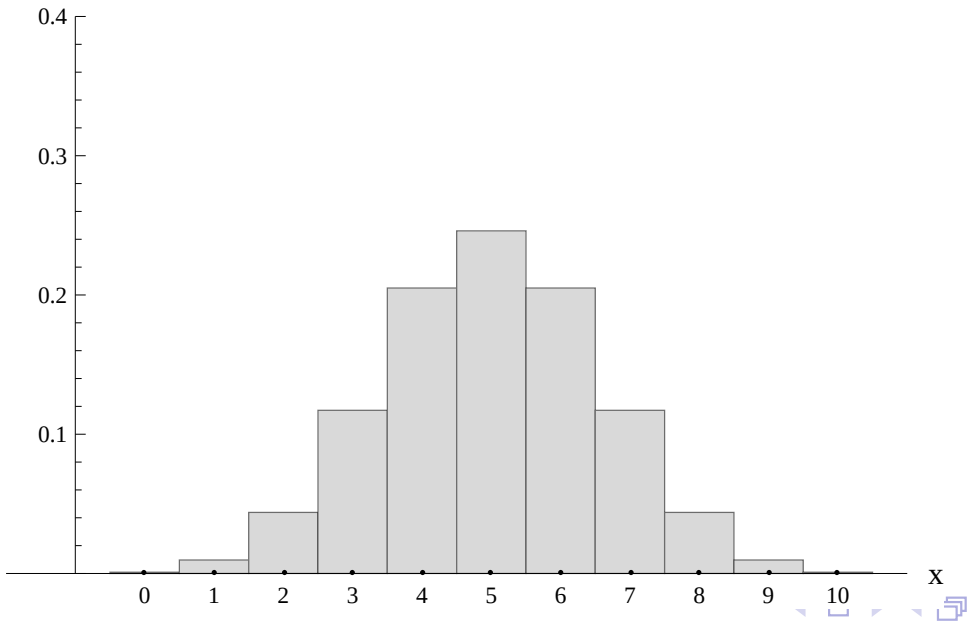
$$p(x) = \mathbf{P}(X = x)$$

Dizemos que $p(x)$ caracteriza a *distribuição de probabilidade* da v.a. X.

Distribuição de uma Variável Aleatória Contínua

Função Densidade de Probabilidade

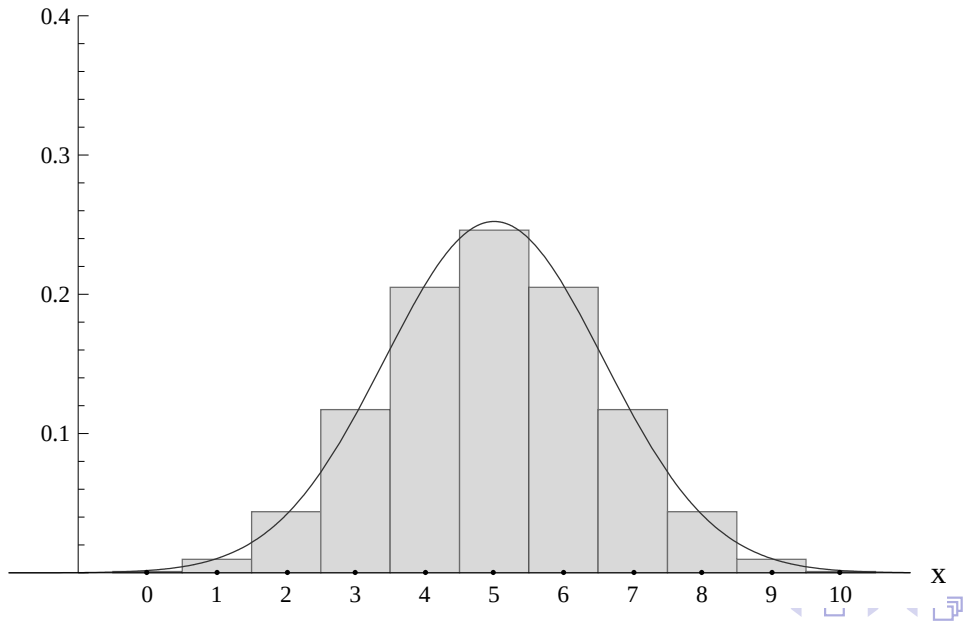
- A **função densidade de probabilidade** de uma variável aleatória contínua caracteriza sua distribuição de probabilidade;
- É o equivalente para o caso contínuo da função de probabilidade das variáveis aleatórias discretas;
- É o modelo teórico para as frequências relativas dos possíveis valores da variável.



Distribuição de uma Variável Aleatória Contínua

Função Densidade de Probabilidade

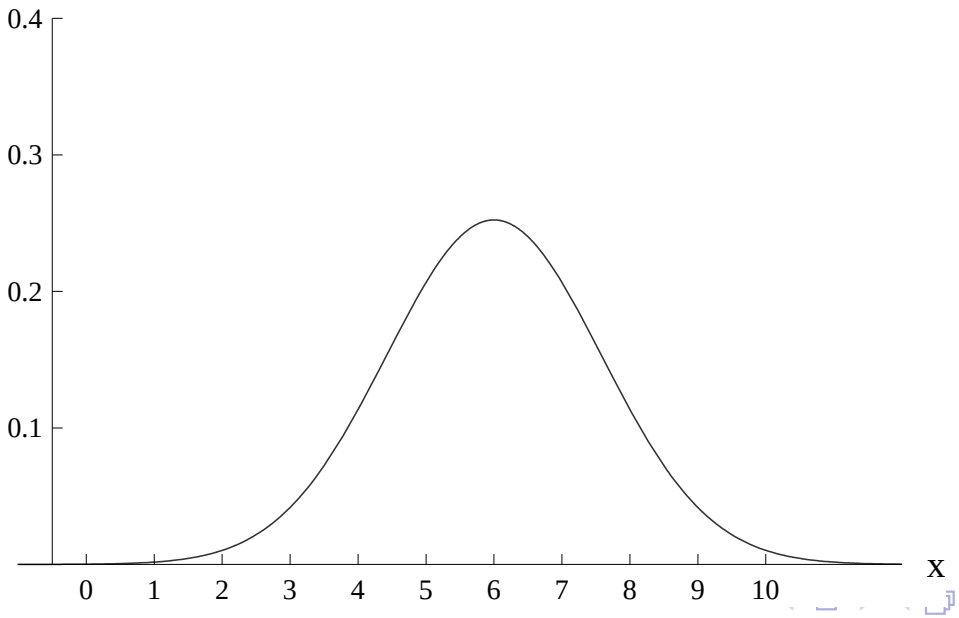
- A **função densidade de probabilidade** de uma variável aleatória contínua caracteriza sua distribuição de probabilidade;
- É o equivalente para o caso contínuo da função de probabilidade das variáveis aleatórias discretas;
- É o modelo teórico para as frequências relativas dos possíveis valores da variável.



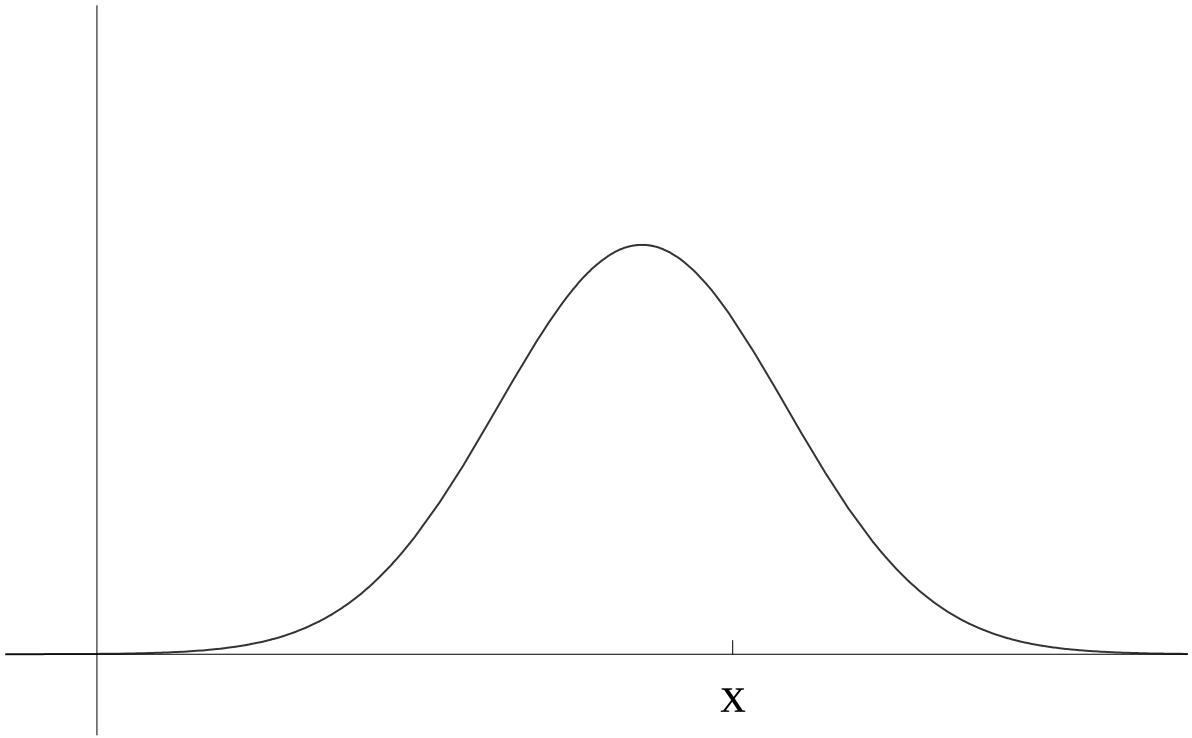
Distribuição de uma Variável Aleatória Contínua

Função Densidade de Probabilidade

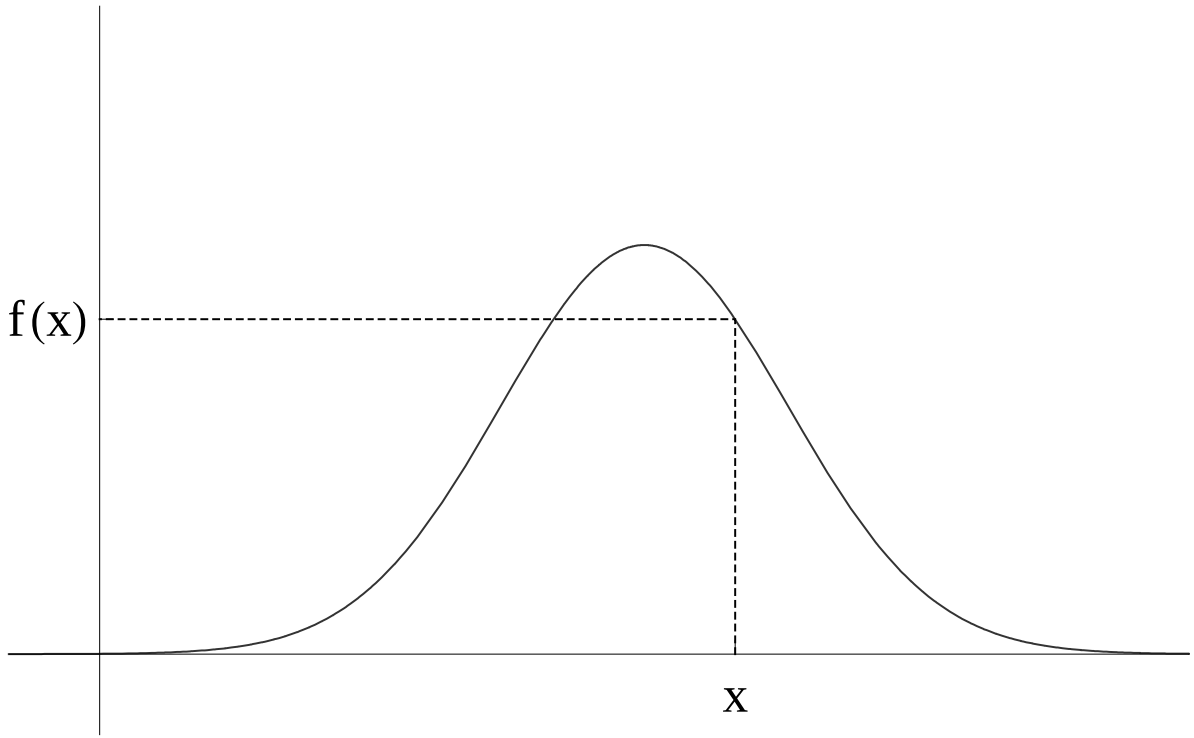
- A **função densidade de probabilidade** de uma variável aleatória contínua caracteriza sua distribuição de probabilidade;
- É o equivalente para o caso contínuo da função de probabilidade das variáveis aleatórias discretas;
- É o modelo teórico para as frequências relativas dos possíveis valores da variável.



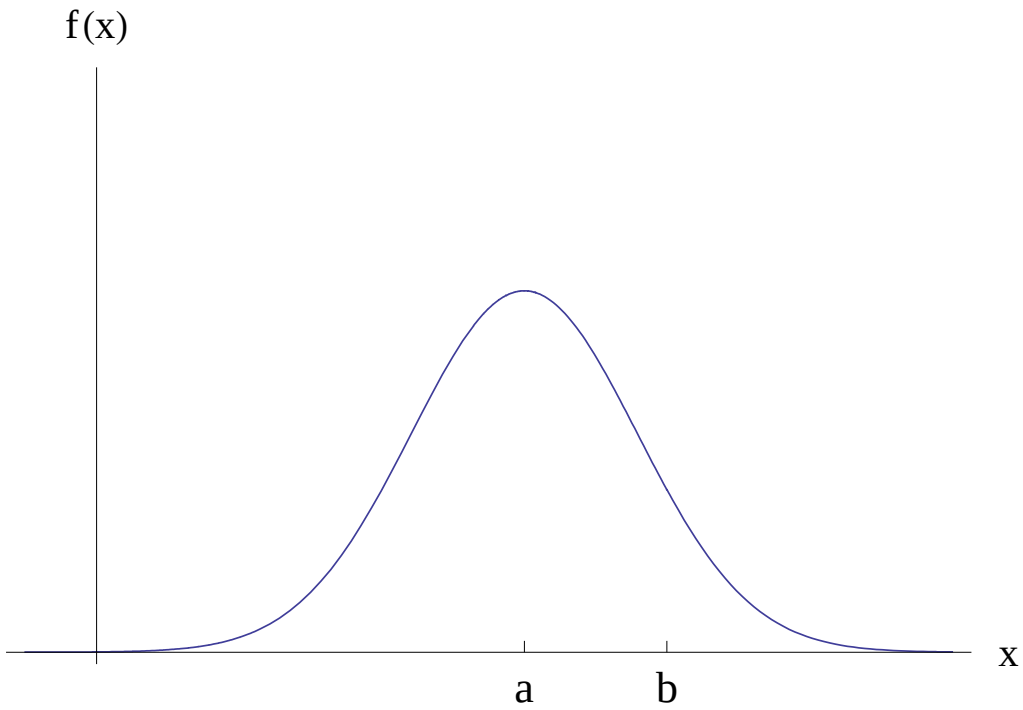
Seja $f(x)$ a função densidade de uma varável aleatória contínua X, então a probabilidade de que X assuma um valor "próximo" a x é proporcional a $f(x)$.



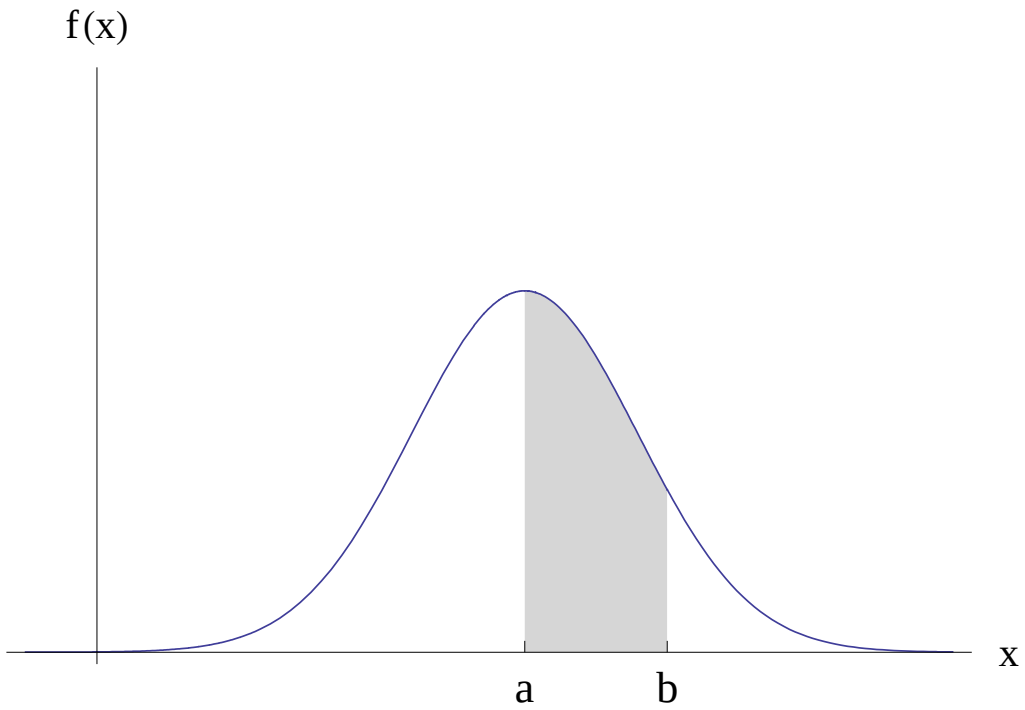
Seja $f(x)$ a função densidade de uma varável aleatória contínua X, então a probabilidade de que X assuma um valor "próximo" a x é proporcional a $f(x)$.



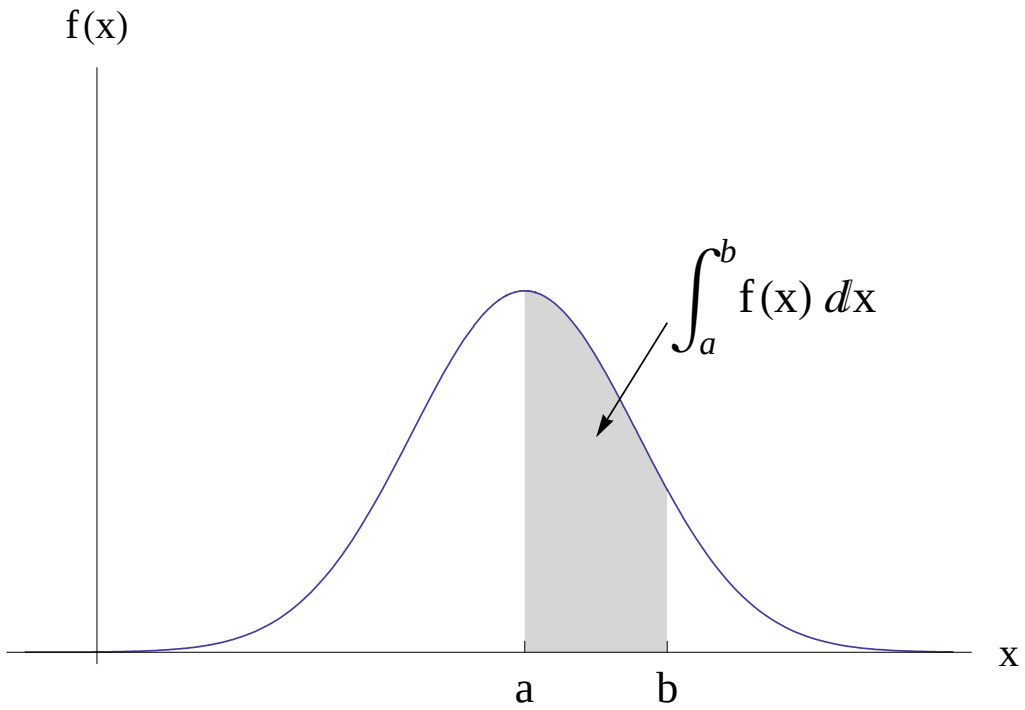
A probabilidade de que X assuma valores num intervalo $[a, b]$ é dada pela área da sua função densidade embaixo do intervalo $[a, b]$.



A probabilidade de que X assuma valores num intervalo $[a, b]$ é dada pela área da sua função densidade embaixo do intervalo $[a, b]$.



A probabilidade de que X assuma valores num intervalo $[a, b]$ é dada pela área da sua função densidade embaixo do intervalo $[a, b]$.



Propriedades da função densidade

Propriedades da função densidade

A função densidade deve satisfazer duas condições análogas às condições exigidas para a função discreta de probabilidade:

i) $f(x) \geq 0$, para todo x;

ii) A área sob o gráfico de $f(x)$ é igual a 1.

Organização

1 Variáveis Aleatórias Contínuas

2 A Distribuição Normal

Distribuição Normal

A distribuição normal é essencialmente importante na estatística porquê:

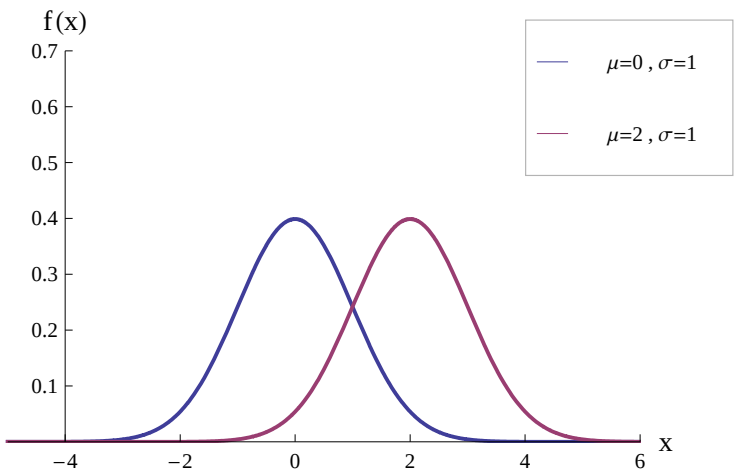
- Inúmeros fenômenos contínuos parecem seguir a dsitribuição normal ou podem ser aproximados por meio dela;
- Podemos utilizá-la para aproximar várias distribuições discretas;
- Oferece a base para a *inferência estatística clássica* devido à sua relação com o *teorema do limite central (TLC)*.

Função Densidade da Normal

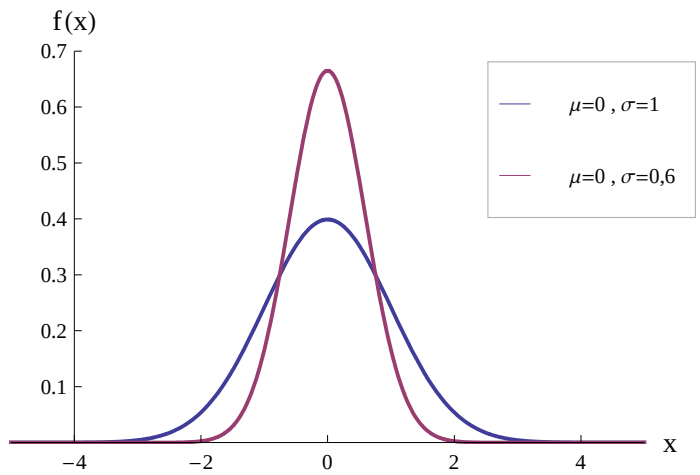
Dizemos que uma v.a. X tem distribuição normal com parâmetros μ e σ^2 se sua função densidade de probabilidade é dada por:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ para } -\infty < x < \infty. \qquad (1)$$

μ: parâmetro de posição



σ : parâmetro de dispersão



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

Propriedades da distribuição Normal

A esperança e a variância da distribuição da normal são dadas por:

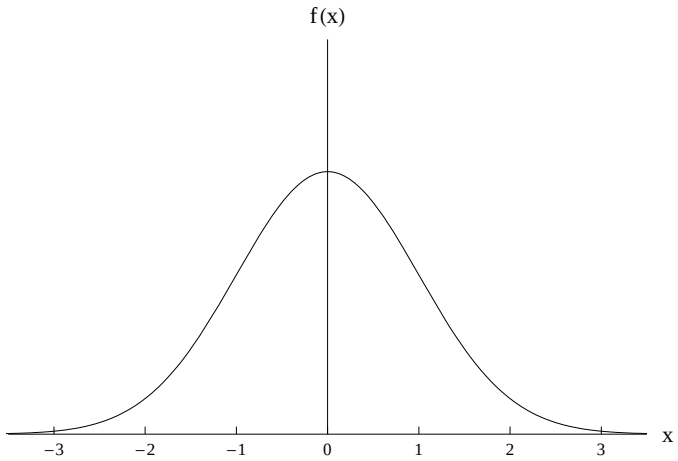
- $\mathbf{E}[X] = \mu$; $\text{Var}[X] = \sigma^2$.

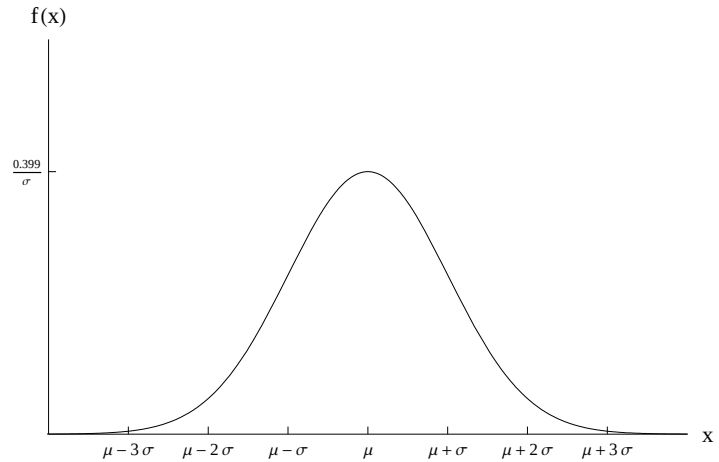
Outras propriedades

(i) $f(x)$ é simétrica em relação a μ;

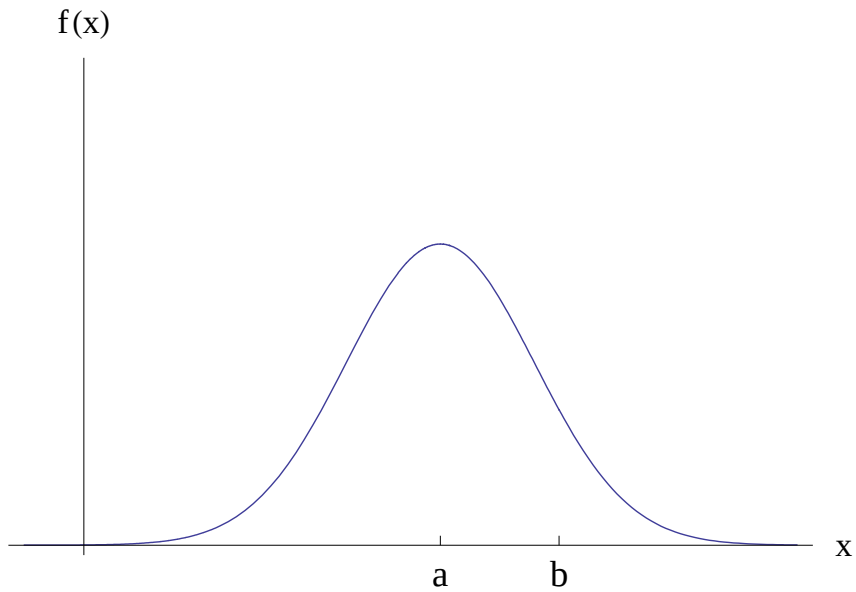
(ii) $f(x) \to 0$ quando $x \to \pm\infty$;

(iii) O valor máximo de $f(x)$ se dá em $x = \mu$.



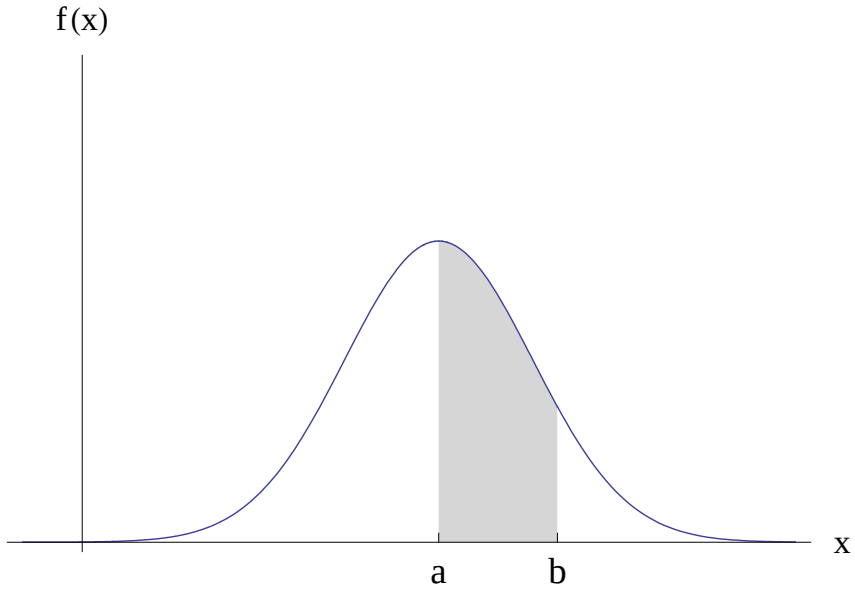


A probabilidade de que X assuma valores num intervalo $[a, b]$ é dada pela área da sua função densidade embaixo do intervalo $[a, b]$.



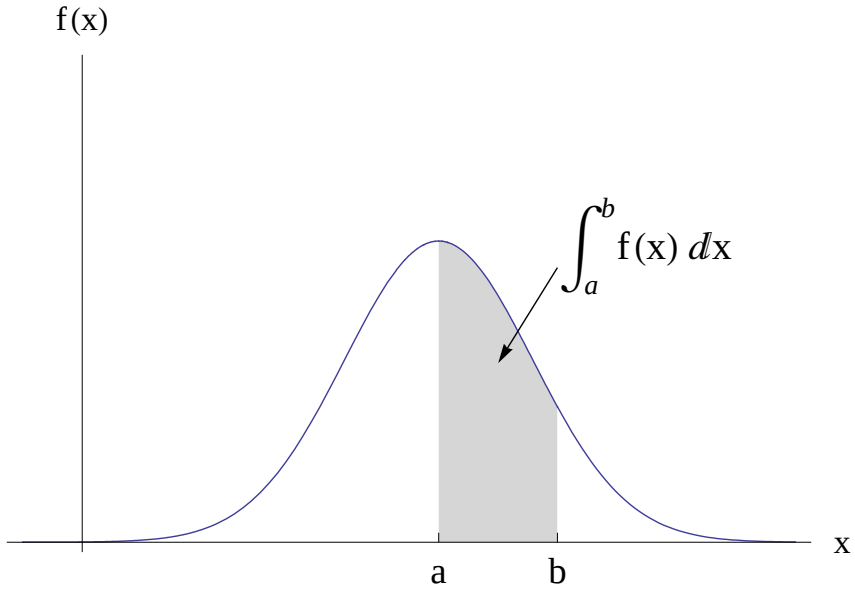
$$\mathbf{P}(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

A probabilidade de que X assuma valores num intervalo $[a, b]$ é dada pela área da sua função densidade embaixo do intervalo $[a, b]$.



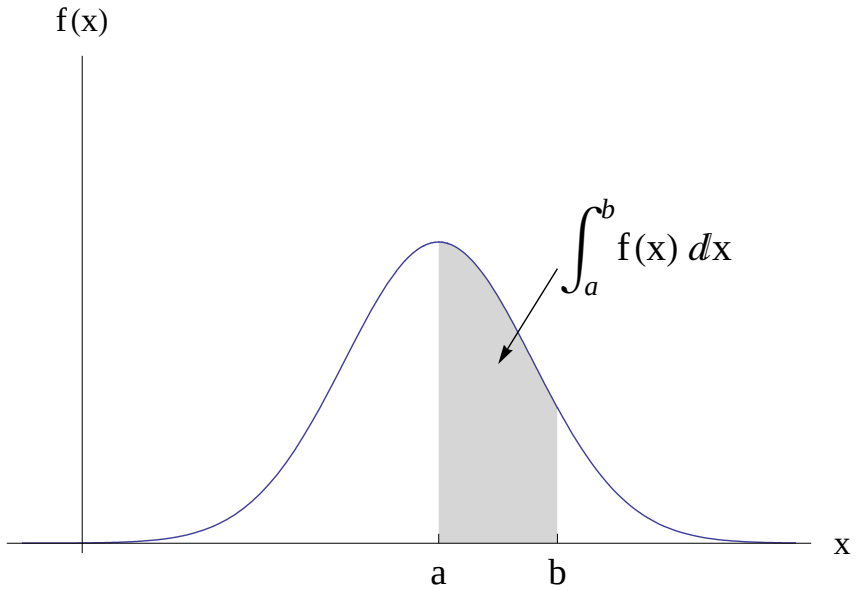
$$\mathbf{P}(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

A probabilidade de que X assuma valores num intervalo $[a, b]$ é dada pela área da sua função densidade embaixo do intervalo $[a, b]$.



$$\mathbf{P}(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

A probabilidade de que X assuma valores num intervalo $[a, b]$ é dada pela área da sua função densidade embaixo do intervalo $[a, b]$.



$$\mathbf{P}(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$\int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad \Rightarrow Que?????$$

- Os cálculos dessas integrais são razoavelmente difíceis!
- Felizmente, esses valores podem ser facilmente obtidos a partir dos *softwares* estatísticos ou mesmo no Excel.
- Também há tabelas que apresentam os cálculos dessas áreas para a **normal padrão**, isto é, a normal com média 0 e variância 1 ($X \sim N(0,1)$).

Exemplo

Seja X uma variável aleatória com distribuição normal padrão. Calcule

a) $\mathbf{P}(X \leq 1,96)$

b) $\mathbf{P}(X \geq 0,5)$

c) $\mathbf{P}(X < -0,5)$

d) $\mathbf{P}(-0,3 < X < 1)$

e) $\mathbf{P}(X > 3,5)$

f) O valor de x para o qual $\mathbf{P}(X < x) = 0,99$

Solução

$\Rightarrow$ Tabela

Exemplo

Seja X uma variável aleatória com distribuição normal padrão. Calcule

a) $\mathbf{P}(X \leq 1,96)$

b) $\mathbf{P}(X \geq 0,5)$

c) $\mathbf{P}(X < -0,5)$

d) $\mathbf{P}(-0,3 < X < 1)$

e) $\mathbf{P}(X > 3,5)$

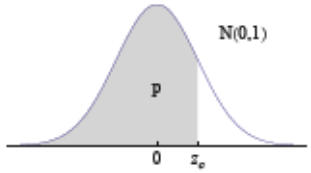
f) O valor de x para o qual $\mathbf{P}(X < x) = 0,99$

Solução

$\Rightarrow$ Tabela

B.1 Normal Padrão

Valores de p tais que $\mathrm{P}(Z < z_c) = p$.



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,10	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,20	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,30	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,40	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,50	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,60	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,70	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,80	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,90	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,00	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,10	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,20	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,30	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,40	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,50	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,60	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,70	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,80	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,90	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,00	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,10	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,20	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,30	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,40	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,50	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,60	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,70	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,80	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,90	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,00	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,10	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,20	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,30	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,40	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

Padronizando a Normal

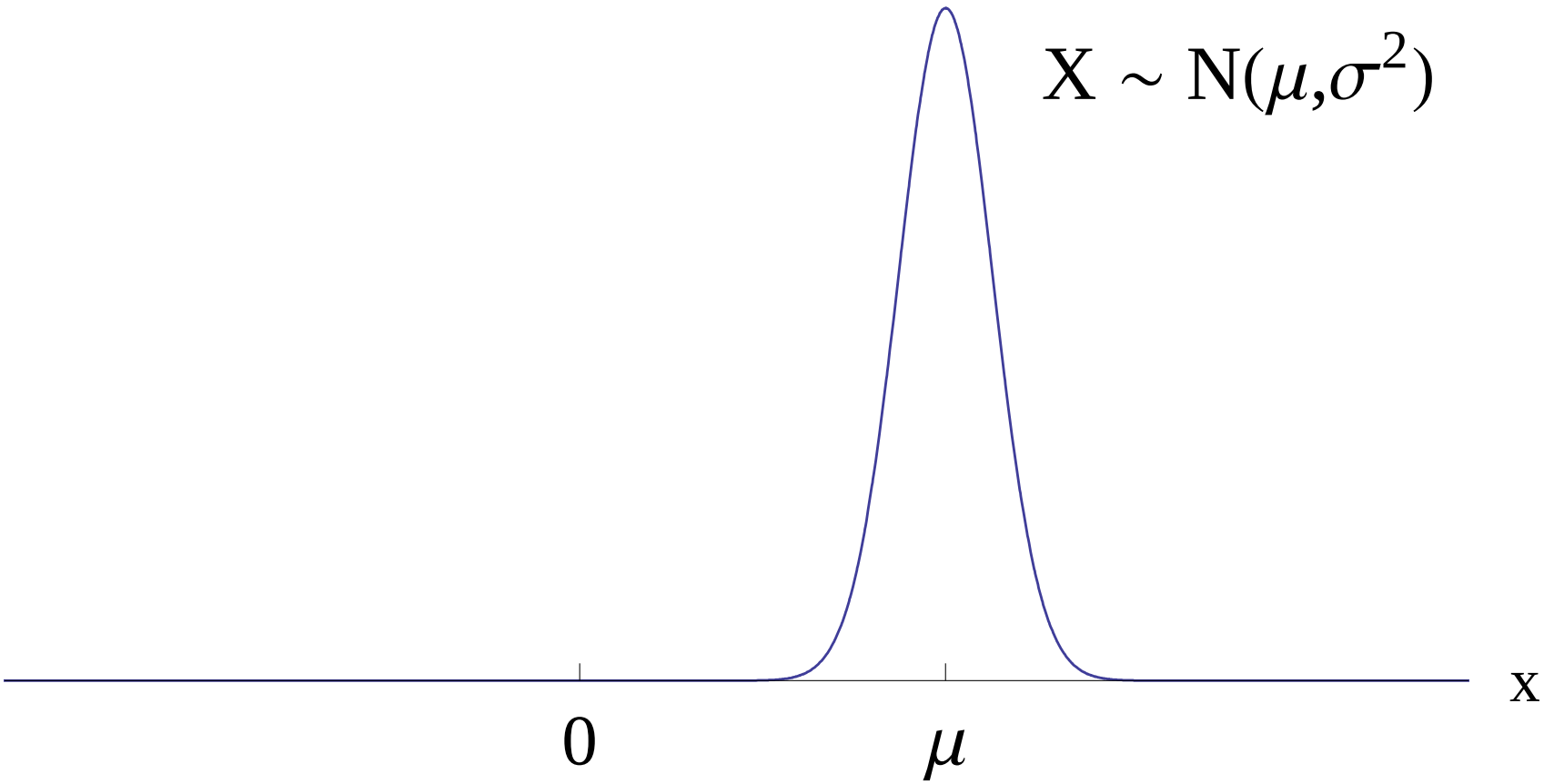
Existe uma relação muito interessante entre uma normal com média μ e variância σ^2 e a normal padrão.

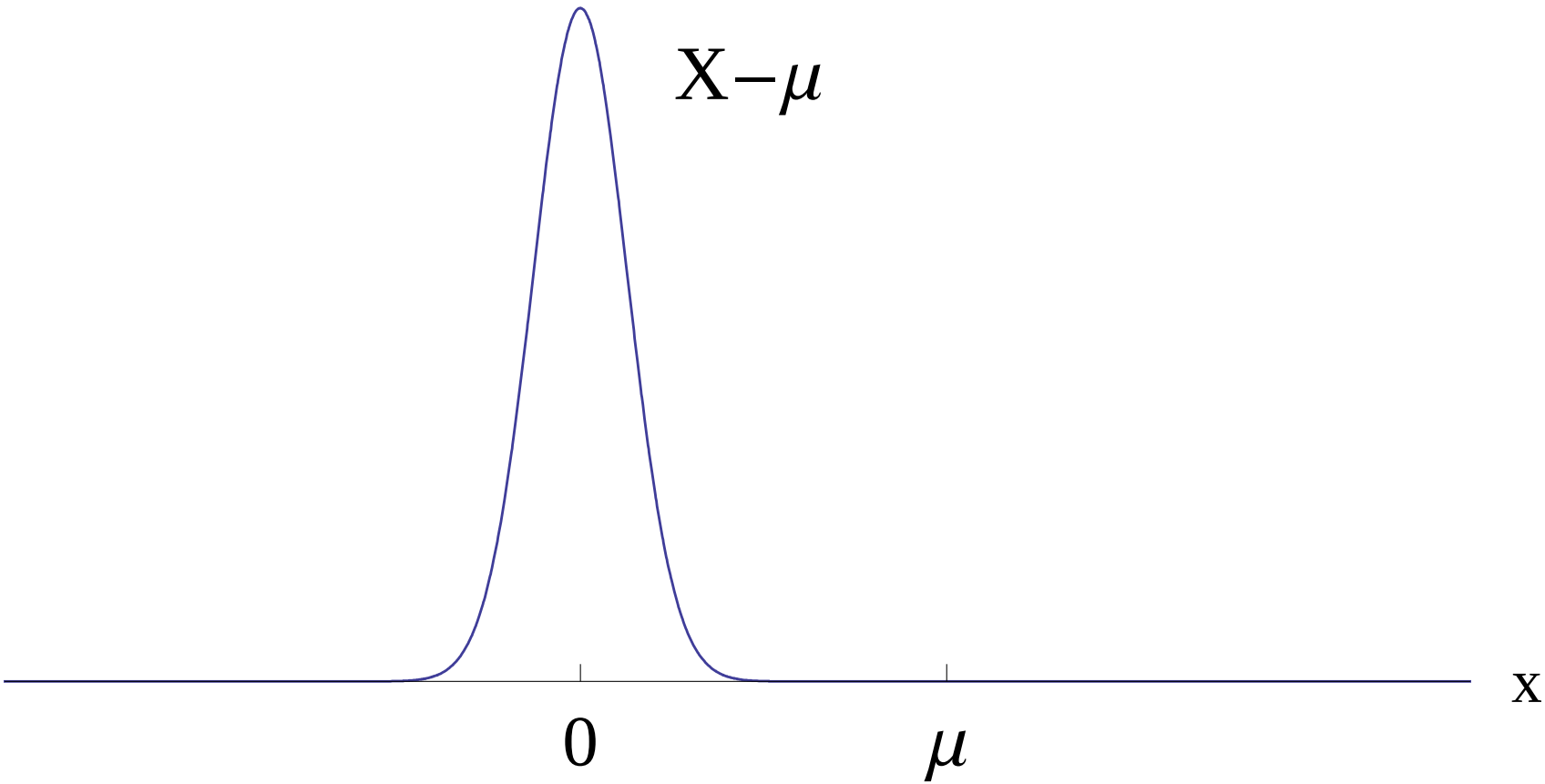
Seja $X \sim N(\mu, \sigma^2)$ e $Z = \frac{(X-\mu)}{\sigma}$. Então

$$\begin{aligned} \mathbf{E}[Z] &= 0 \\ \text{Var}[Z] &= 1 \end{aligned}$$

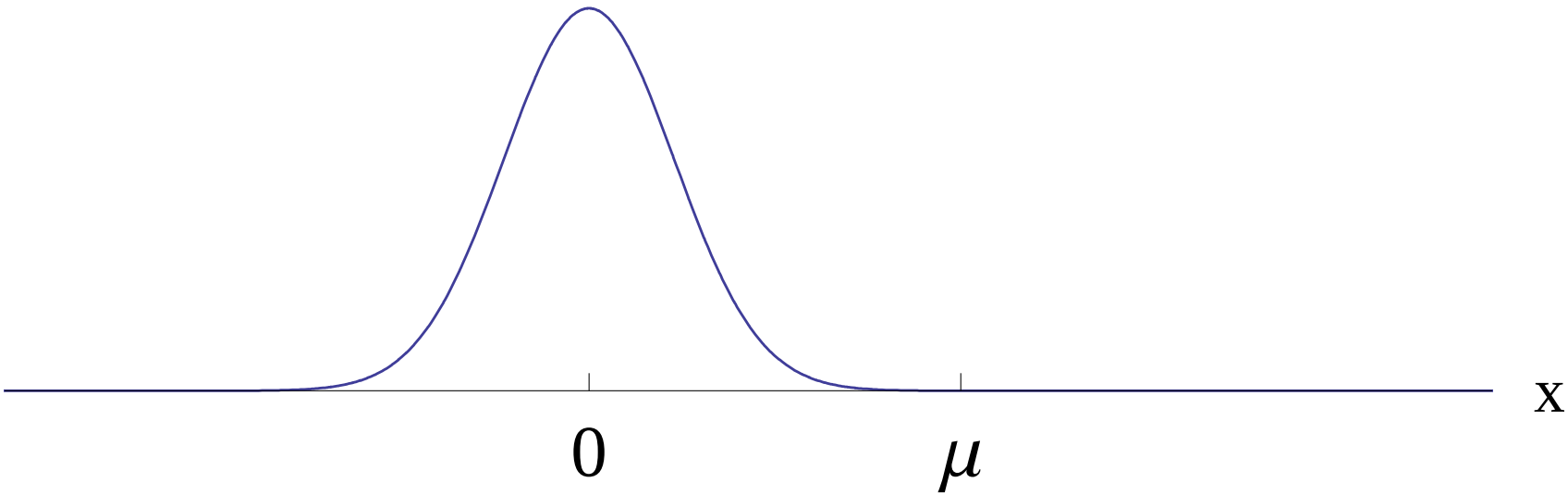
Essa transformação não afeta a normalidade:

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


$X \sim N(\mu,\sigma^2)$
x
0
μ


X−μ
x
0
μ

$$\frac{X - \mu}{\sigma} \sim \text{N}(0,1)$$



Cálculo de probabilidades envolvendo uma distribuição Normal qualquer

Suponha que $X \sim N(\mu, \sigma^2)$ e que queremos calcular $\mathbf{P}(X \leq a)$.

Observe que:

$$\{X \leq a\} \Leftrightarrow \{X - \mu \leq a - \mu\} \Leftrightarrow \left\{\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right\}$$

Logo,

$$\mathbf{P}\{X \leq a\} = \mathbf{P}\left\{\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right\}$$

Como

$$\frac{X-\mu}{\sigma} \sim N(0,1)$$

resulta que

$$\mathbf{P}\{X \leq a\} = \mathbf{P}\left\{Z \leq \frac{a-\mu}{\sigma}\right\}$$

Onde

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

Cálculo de probabilidades com a distribuição Normal

Suponha que $X \sim N(\mu, \sigma^2)$.

$$\begin{aligned}
\mathbf{P}\{X \leq a\} &= \mathbf{P}\left\{Z \leq \frac{a-\mu}{\sigma}\right\} \\
\mathbf{P}\{X \geq a\} &= \mathbf{P}\left\{Z \geq \frac{a-\mu}{\sigma}\right\} \\
\mathbf{P}\{a \leq X \leq b\} &= \mathbf{P}\left\{\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right\}
\end{aligned}$$

onde

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

Obs. Para fazer cálculo de probabilidades com a distribuição Normal, basta usar as fórmulas acima substituindo os valores de a, μ e σ pelos valores dados.

Exemplo

Seja X uma variável aleatória com distribuição normal com média 2 e variância 9, ou seja, $X \sim N(2; 3^2)$. Calcule

a) $\mathbf{P}(X > 2)$

b) $\mathbf{P}(X < 5)$

c) $\mathbf{P}(2 < X < 5)$

Solução

a) $X \sim N(2; 3^2) \Rightarrow \mu = 2$ e $\sigma = 3$.

$$
\begin{aligned}
\mathbf{P}\{X > a\} &= \mathbf{P}\left\{Z > \frac{a-\mu}{\sigma}\right\} \\
\mathbf{P}(X > 2) &= \mathbf{P}\left(Z > \frac{2-\mu}{\sigma}\right) \\
&= \mathbf{P}\left(Z > \frac{2-2}{3}\right) \\
&= \mathbf{P}(Z > 0) \\
&= \frac{1}{2}
\end{aligned}
$$

Exemplo

Seja X uma variável aleatória com distribuição normal com média 2 e variância 9, ou seja, $X \sim N(2; 3^2)$. Calcule

a) $\mathbf{P}(X > 2)$

b) $\mathbf{P}(X < 5)$

c) $\mathbf{P}(2 < X < 5)$

Solução

a) $X \sim N(2; 3^2) \Rightarrow \mu = 2$ e $\sigma = 3$.

$$
\begin{aligned}
\mathbf{P}\{X > a\} &= \mathbf{P}\left\{Z > \frac{a-\mu}{\sigma}\right\} \\
\mathbf{P}(X > 2) &= \mathbf{P}\left(Z > \frac{2-\mu}{\sigma}\right) \\
&= \mathbf{P}\left(Z > \frac{2-2}{3}\right) \\
&= \mathbf{P}(Z > 0) \\
&= \frac{1}{2}
\end{aligned}
$$

b) $\mu = 2$ e $\sigma = 3$.

$$\begin{aligned}
\mathbf{P}(X < 5) &= \mathbf{P}\left(Z < \frac{5-\mu}{\sigma}\right) \\
&= \mathbf{P}\left(Z < \frac{5-2}{3}\right) \\
&= \mathbf{P}(Z < 1) \\
&= 0,8413
\end{aligned}$$

c)

$$\begin{aligned}
\mathbf{P}(2 < X < 5) &= \mathbf{P}(X < 5) - \mathbf{P}(X < 2) \\
&= 0,8413 - 0,5 = 0,3413
\end{aligned}$$

Exercício

As alturas dos alunos de uma universidade têm distribuição aproximadamente normal com média 170cm e desvio padrão 15cm. Sorteado um aluno ao acaso, determine:

a) A probabilidade do aluno sorteado ter menos do que 140cm.

b) A probabilidade do aluno sorteado ter mais do que 190cm.

c) A probabilidade do aluno sorteado ter entre 140cm e 190cm.