

Capítulo 5 | Distribuições de probabilidade normal



Descrição do capítulo

- *5.1 Introdução à distribuição normal e distribuição normal padrão*
- *5.2 Distribuições normais: encontrando probabilidades*
- *5.3 Distribuições normais: encontrando valores*
- *5.4 Distribuições amostrais e o Teorema Central do Limite*
- *5.5 Aproximações normais para distribuições binomiais*

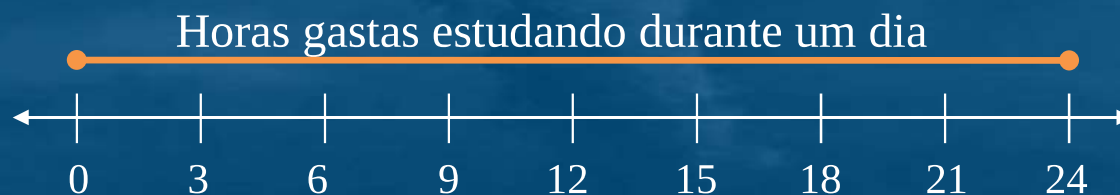
Seção 5.1

Introdução à distribuição normal e distribuição normal padrão

Propriedades de uma distribuição normal

Variável aleatória contínua

- *Tem um número infinito de valores possíveis que podem ser representados por um intervalo na reta numérica*



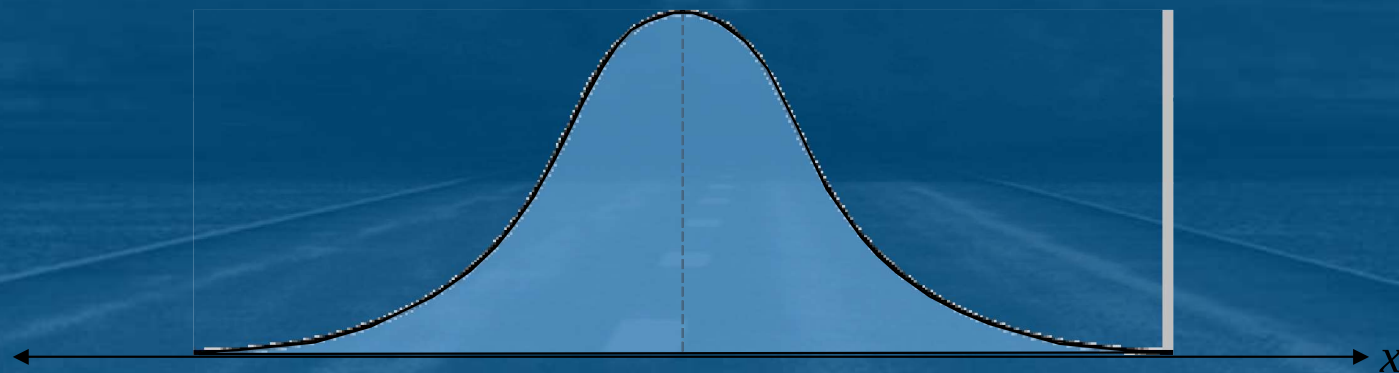
O tempo gasto estudando pode ser qualquer número entre 0 e 24.

Distribuição de probabilidade contínua

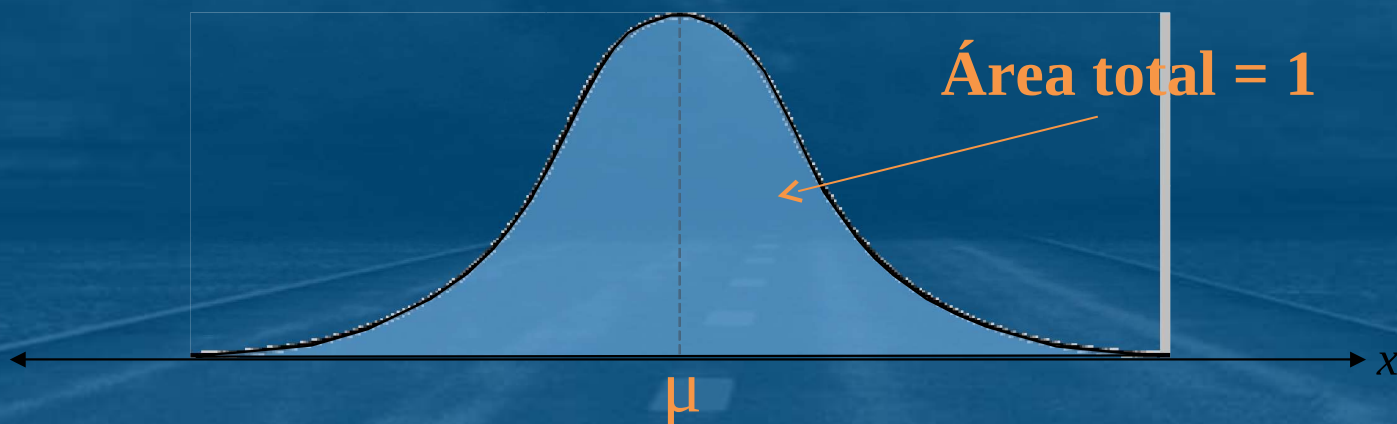
- *A distribuição de probabilidade de uma variável aleatória contínua*

Distribuição normal

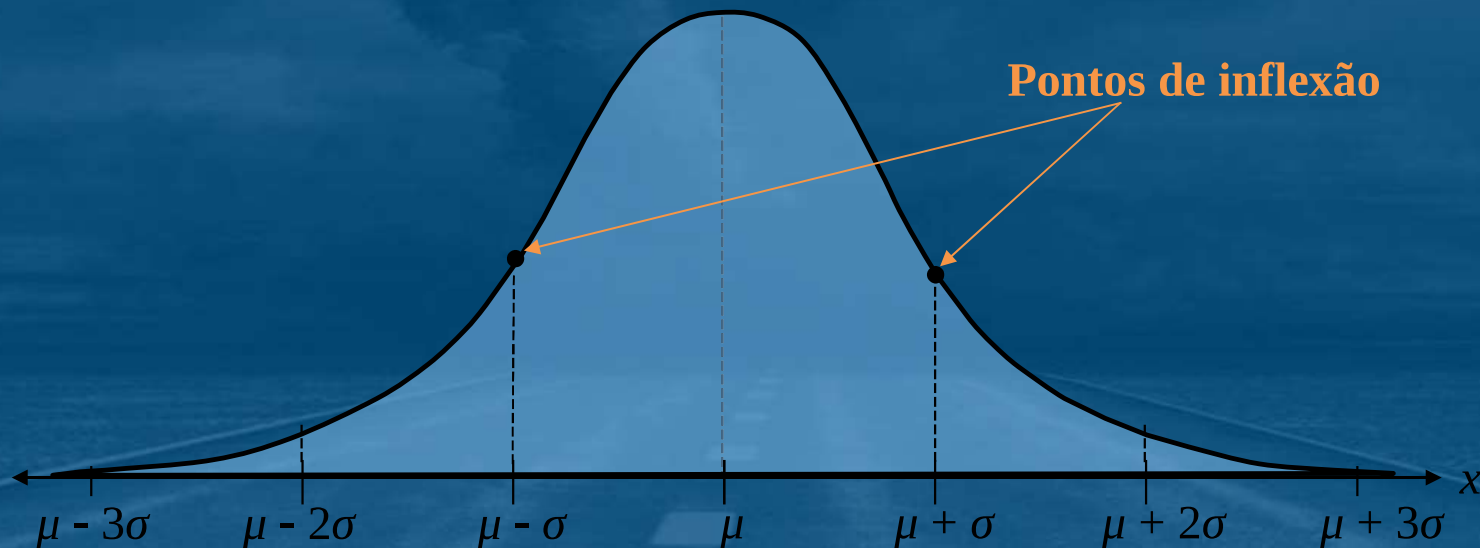
- *Uma probabilidade contínua para uma variável aleatória, x*
- *A mais importante probabilidade contínua na estatística*
- *O gráfico de uma distribuição normal é chamado de curva normal*



1. *A média, a mediana e a moda são iguais.*
2. *A curva normal tem formato de sino e é simétrica em relação à média.*
3. *A área total abaixo da curva é igual a um.*
4. *A curva normal se aproxima do eixo x , mas nunca o toca, conforme se afasta da média.*

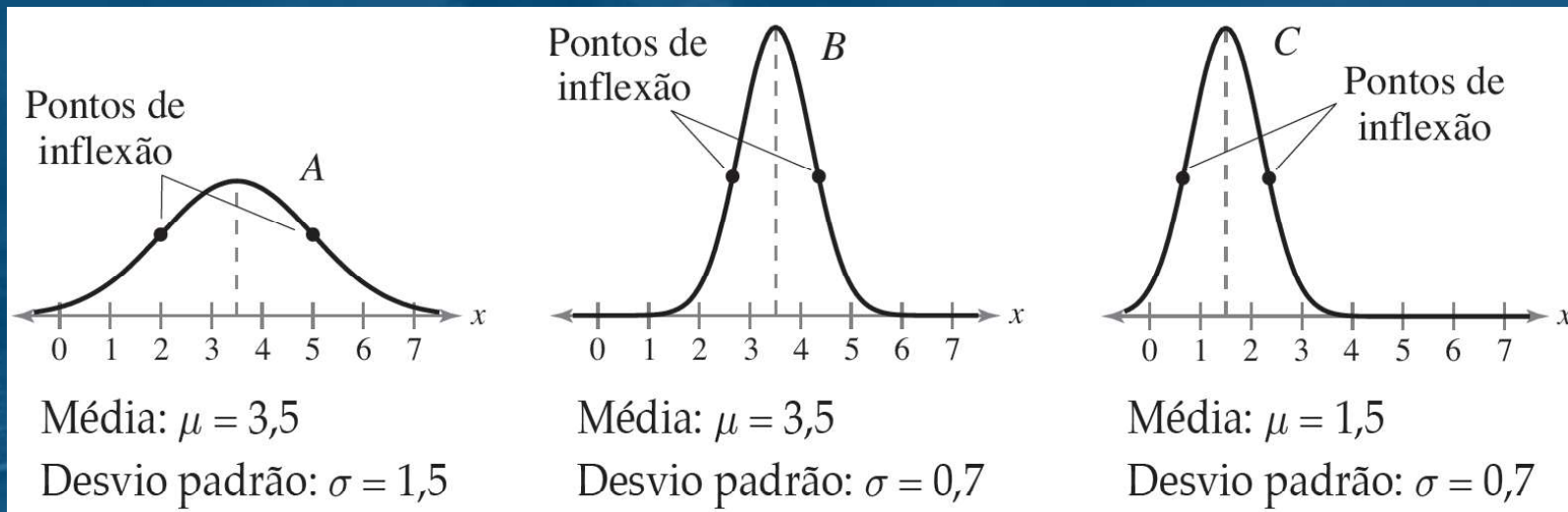


5. *Entre $\mu - \sigma$ e $\mu + \sigma$ (no centro da curva), o gráfico se curva para baixo. O gráfico se curva para cima à esquerda de $\mu - \sigma$ e à direita de $\mu + \sigma$. Os pontos nos quais a curva muda a sua trajetória para cima ou para baixo são chamados de pontos de inflexão.*



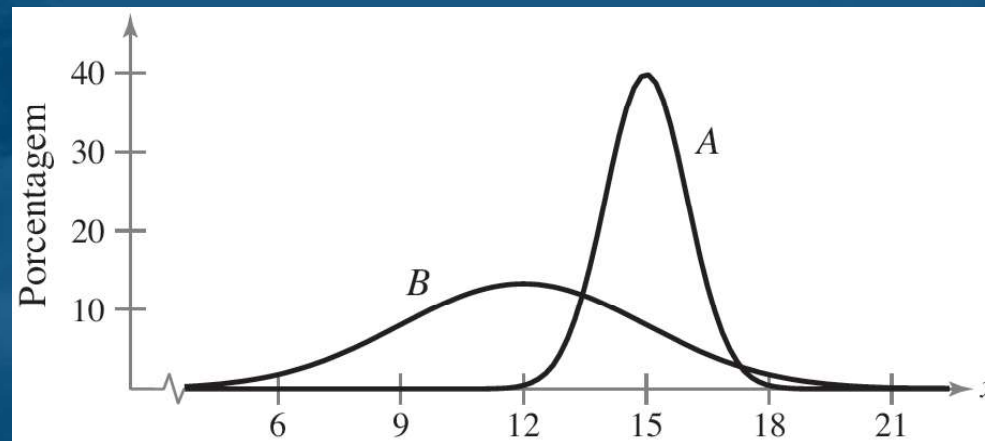
Médias e desvios padrão

- *Uma distribuição normal pode ter qualquer média e qualquer desvio padrão positivo*
- *A média dá a localização da linha de simetria*
- *O desvio padrão descreve a dispersão dos dados*



Exemplo: entendendo média e desvio padrão

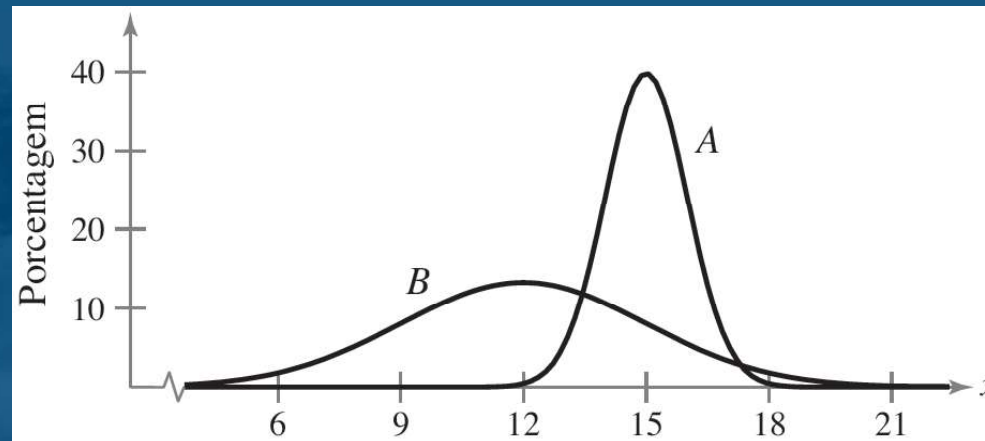
1. Qual curva tem a maior média?



Solução:

A curva A tem a maior média (a linha de simetria da curva A ocorre em $x = 15$. A linha de simetria da curva B ocorre em $x = 12$).

2. Qual curva tem o maior desvio padrão?



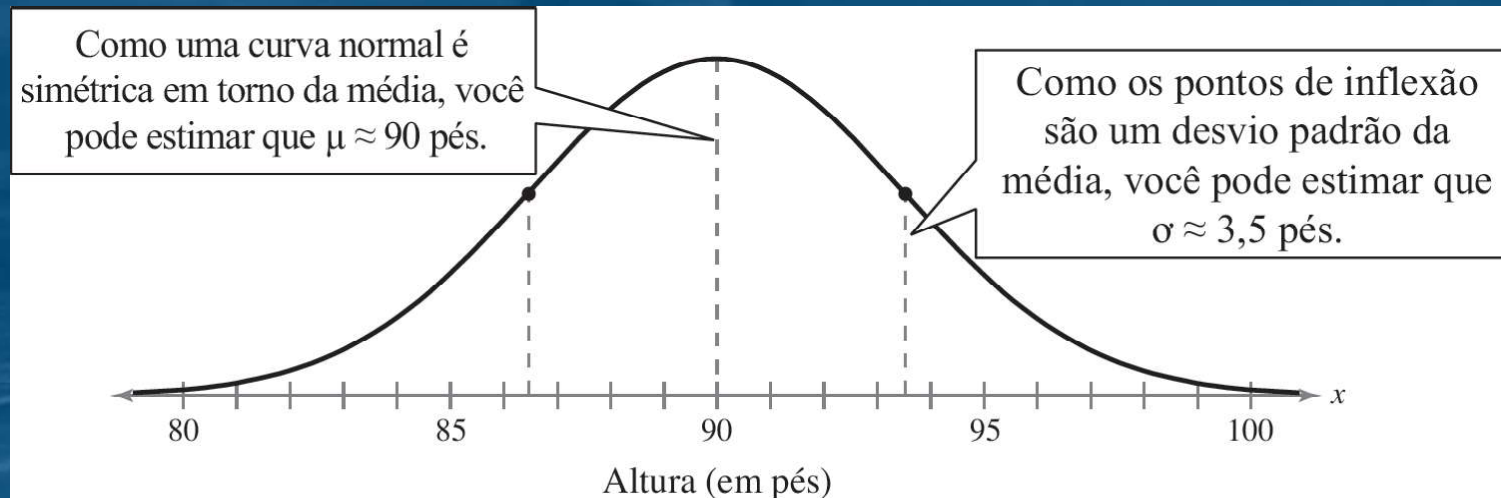
Solução:

A curva B tem o maior desvio padrão (a curva B é mais dispersa que a curva A).

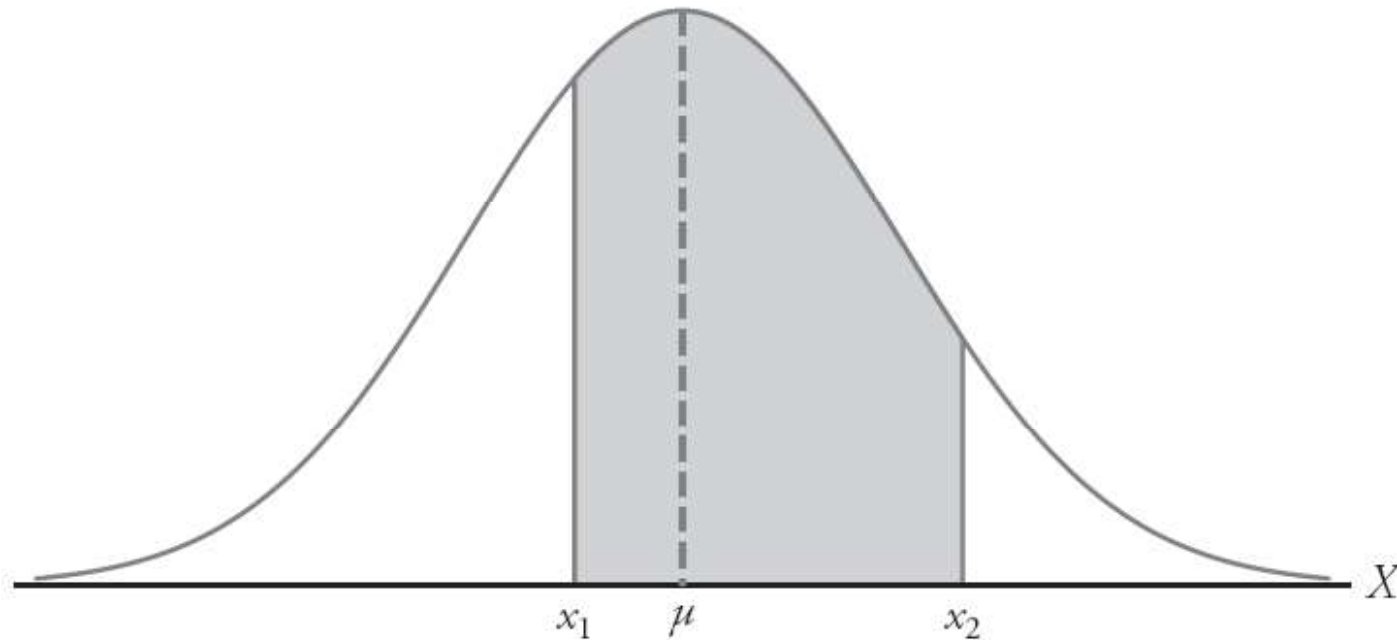
Exemplo: interpretando gráficos

As alturas (em pés) de árvores de carvalho adultas são normalmente distribuídas. A curva normal apresentada mostra essa distribuição. Qual é a média de altura de uma árvore de carvalho adulta? Estime o desvio padrão.

Solução:



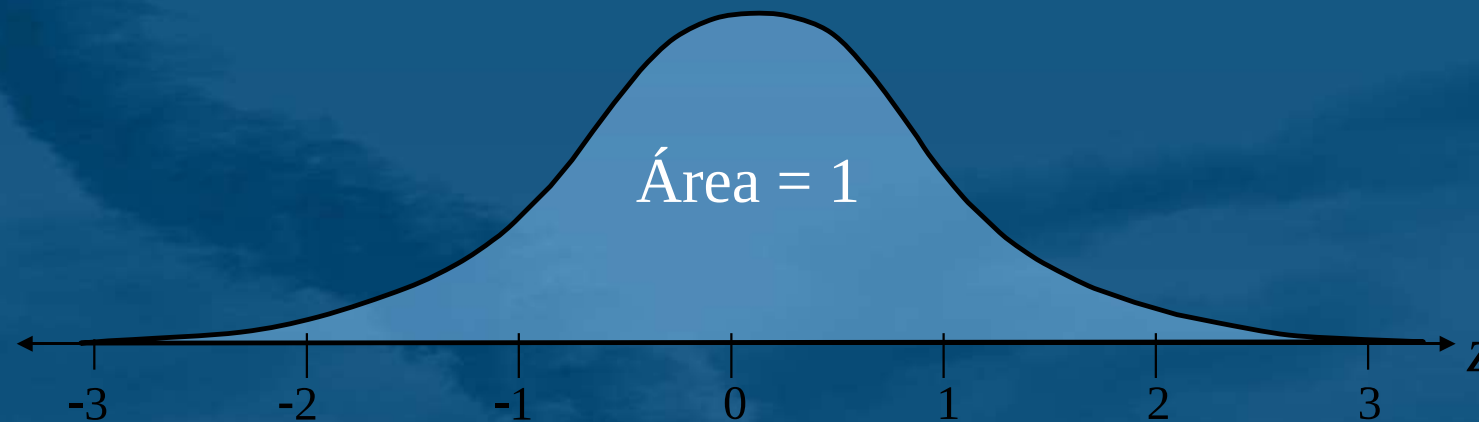
A área na região sombreada



...pode ser calculada como

$$\begin{aligned} P(x_1 < X < x_2) &= \int_{x_1}^{x_2} n(x; \mu, \sigma) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{x_1}^{x_2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx, \end{aligned}$$

A distribuição normal padrão



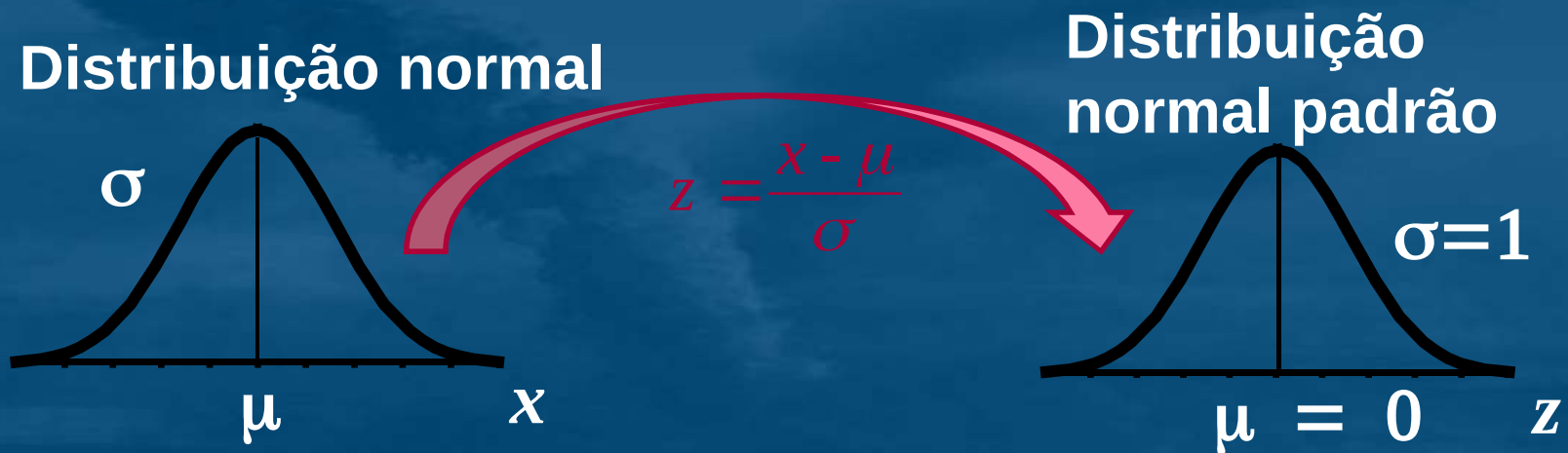
- Qualquer valor de x pode ser transformado em um escore z usando a fórmula

$$z = \frac{\text{Valor} - \text{Média}}{\text{Desvio padrão}}$$

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

Arredonde para o centésimo mais próximo.

- Se cada valor de dados de uma variável aleatória normalmente distribuída x for transformada em um escore z , o resultado será a distribuição normal padrão



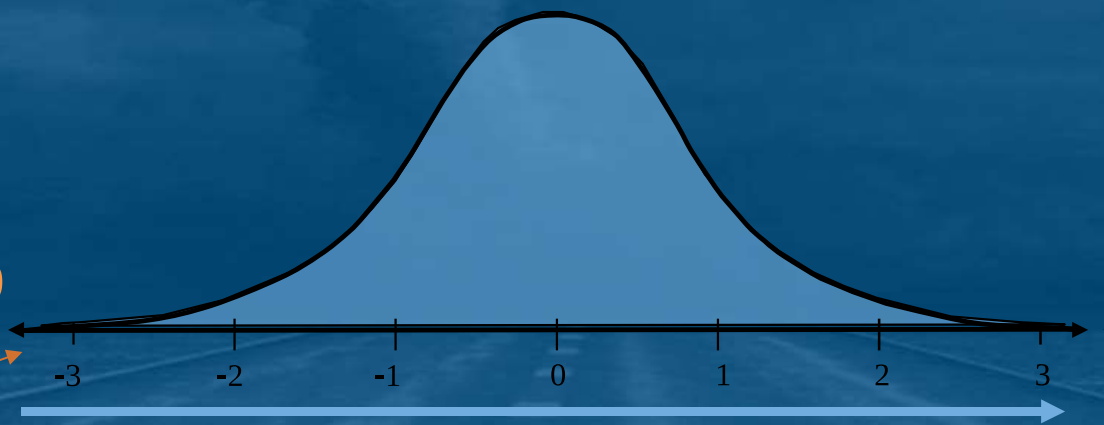
- Use a tabela normal padrão para encontrar a área cumulativa abaixo da curva normal padrão

Propriedades da distribuição normal padrão

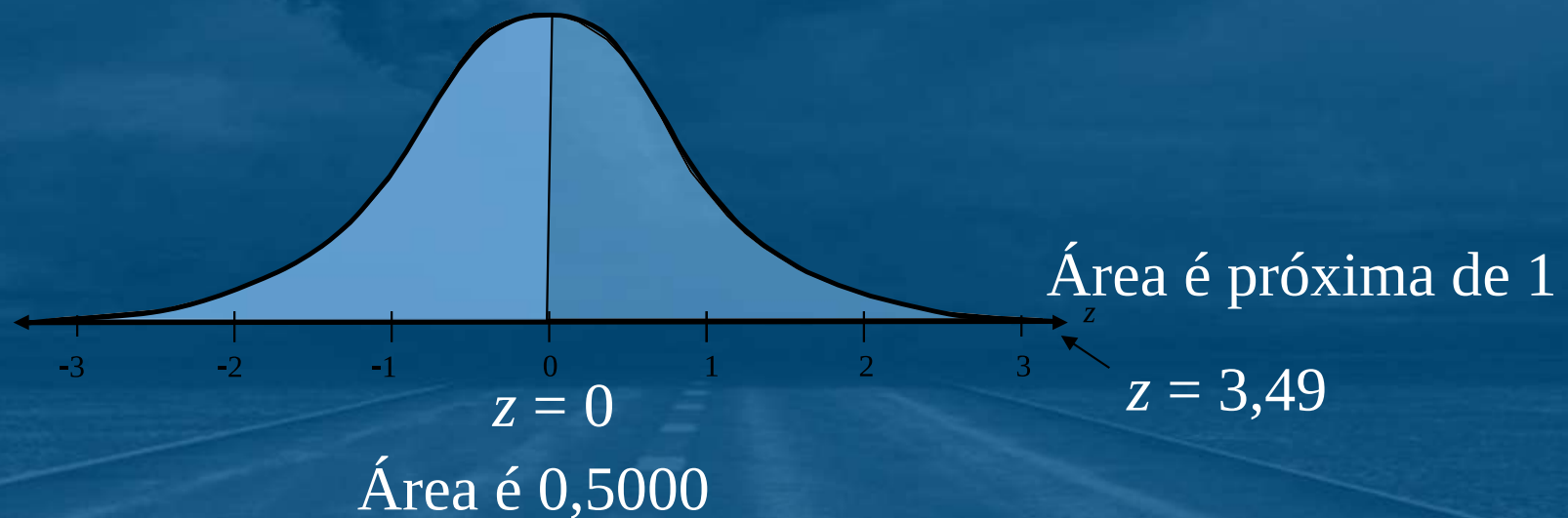
1. A área cumulativa é próxima de 0 para escore z próximos de $z = -3,49$.
2. A área cumulativa aumenta conforme o escore z aumenta.

Área é
próxima de 0

$z = -3,49$

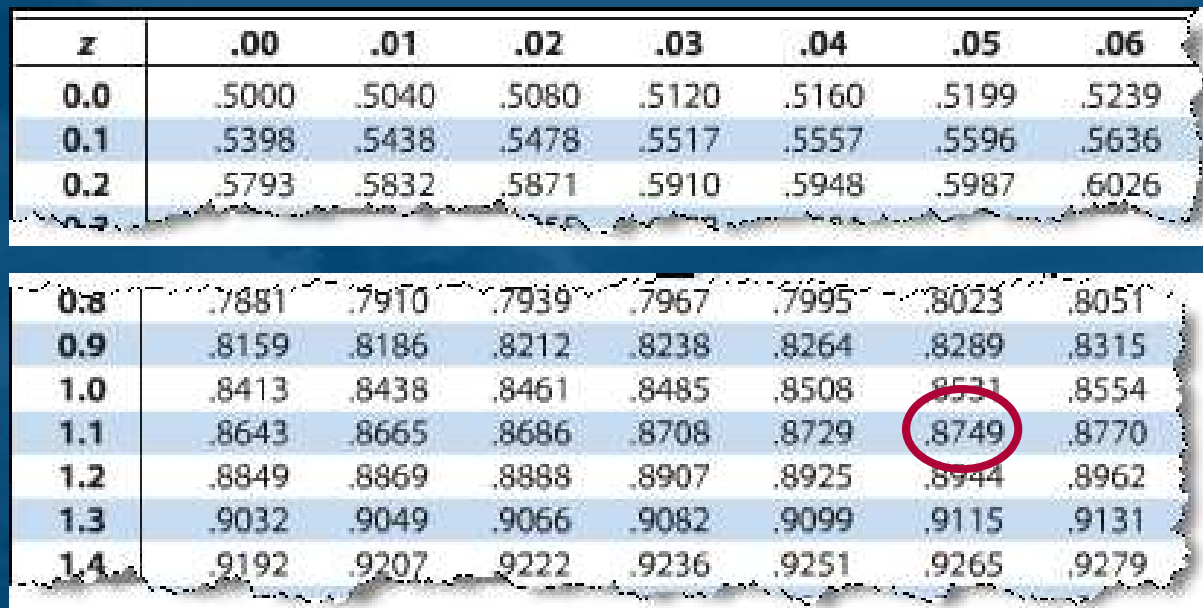


3. A área cumulativa para $z = 0$ é 0,5000.
4. A área cumulativa é próxima de 1 para escore z próximo de $z = 3,49$.



Exemplo: usando a tabela normal padrão

Encontre a área cumulativa que corresponda a um escore z de 1.15.



z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026

0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279


Solução:

Encontre 1.1 na coluna à esquerda.


Cruze a fileira para a coluna sob 0.05.

A área à esquerda de $z = 1,15$ é 0,8749.

Encontre a área cumulativa que corresponda a um escore z de $-0,24$.



z	.09	.08	.07	.06	.05	.04	.03
-3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
-3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004
-3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006
-3.1	.0006	.0007	.0007	.0008	.0008	.0008	.0008
-3.0	.0008	.0009	.0009	.0010	.0010	.0010	.0010
-2.9	.0010	.0011	.0011	.0012	.0012	.0012	.0012
-2.8	.0013	.0013	.0014	.0014	.0014	.0014	.0014
-2.7	.0015	.0015	.0016	.0016	.0016	.0016	.0016
-2.6	.0017	.0017	.0018	.0018	.0018	.0018	.0018
-2.5	.0019	.0019	.0020	.0020	.0020	.0020	.0020
-2.4	.0021	.0021	.0022	.0022	.0022	.0022	.0022
-2.3	.0023	.0023	.0024	.0024	.0024	.0024	.0024
-2.2	.0025	.0025	.0026	.0026	.0026	.0026	.0026
-2.1	.0027	.0027	.0028	.0028	.0028	.0028	.0028
-2.0	.0029	.0029	.0030	.0030	.0030	.0030	.0030
-1.9	.0031	.0031	.0032	.0032	.0032	.0032	.0032
-1.8	.0033	.0033	.0034	.0034	.0034	.0034	.0034
-1.7	.0035	.0035	.0036	.0036	.0036	.0036	.0036
-1.6	.0037	.0037	.0038	.0038	.0038	.0038	.0038
-1.5	.0039	.0039	.0040	.0040	.0040	.0040	.0040
-1.4	.0041	.0041	.0042	.0042	.0042	.0042	.0042
-1.3	.0043	.0043	.0044	.0044	.0044	.0044	.0044
-1.2	.0045	.0045	.0046	.0046	.0046	.0046	.0046
-1.1	.0047	.0047	.0048	.0048	.0048	.0048	.0048
-1.0	.0049	.0049	.0050	.0050	.0050	.0050	.0050
-0.9	.0051	.0051	.0052	.0052	.0052	.0052	.0052
-0.8	.0053	.0053	.0054	.0054	.0054	.0054	.0054
-0.7	.0055	.0055	.0056	.0056	.0056	.0056	.0056
-0.6	.0057	.0057	.0058	.0058	.0058	.0058	.0058
-0.5	.0059	.0059	.0060	.0060	.0060	.0060	.0060
-0.4	.0061	.0061	.0062	.0062	.0062	.0062	.0062
-0.3	.0064	.0064	.0065	.0065	.0065	.0065	.0065
-0.2	.0066	.0066	.0067	.0067	.0067	.0067	.0067
-0.1	.0069	.0069	.0070	.0070	.0070	.0070	.0070
0.0	.0071	.0071	.0072	.0072	.0072	.0072	.0072



Solução:

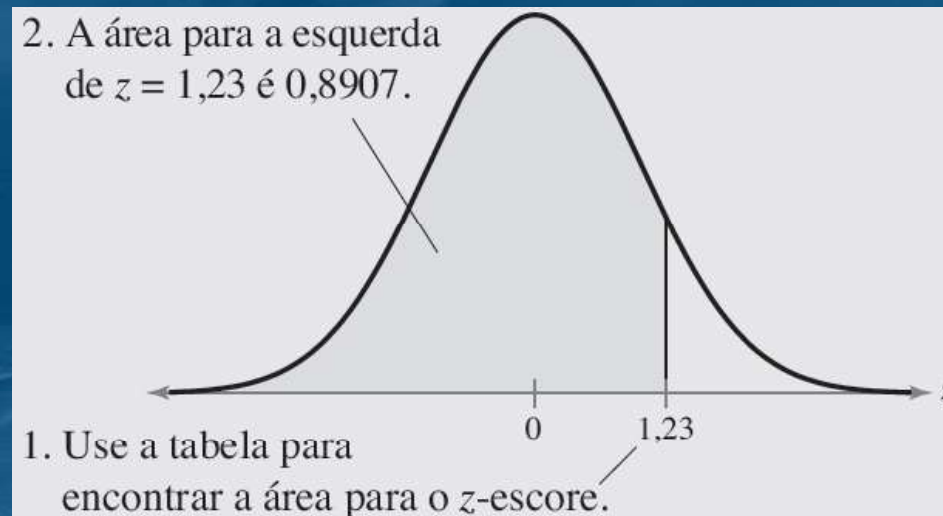
Encontre $-0,2$ na coluna à esquerda.

Cruze a fileira para a coluna sob 0.04 .

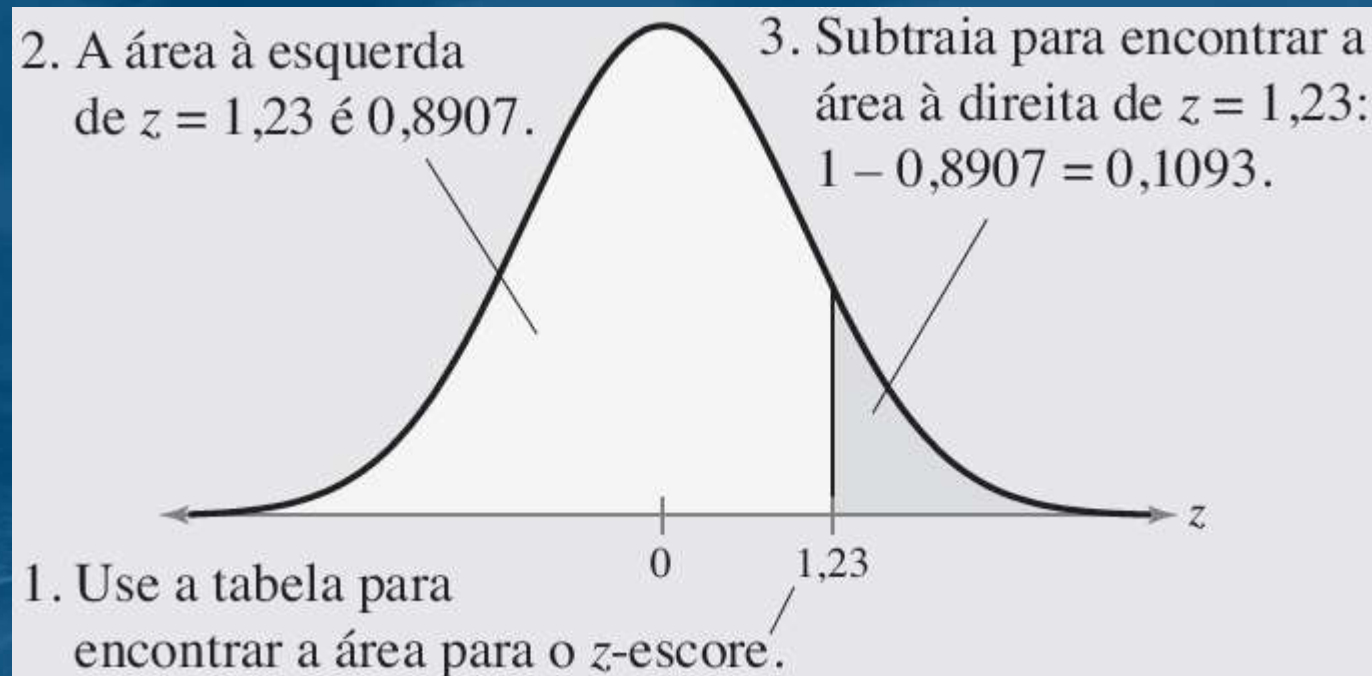
A área à esquerda de $z = -0,24$ é $0,4052$.

Encontrando áreas sob a curva normal padrão

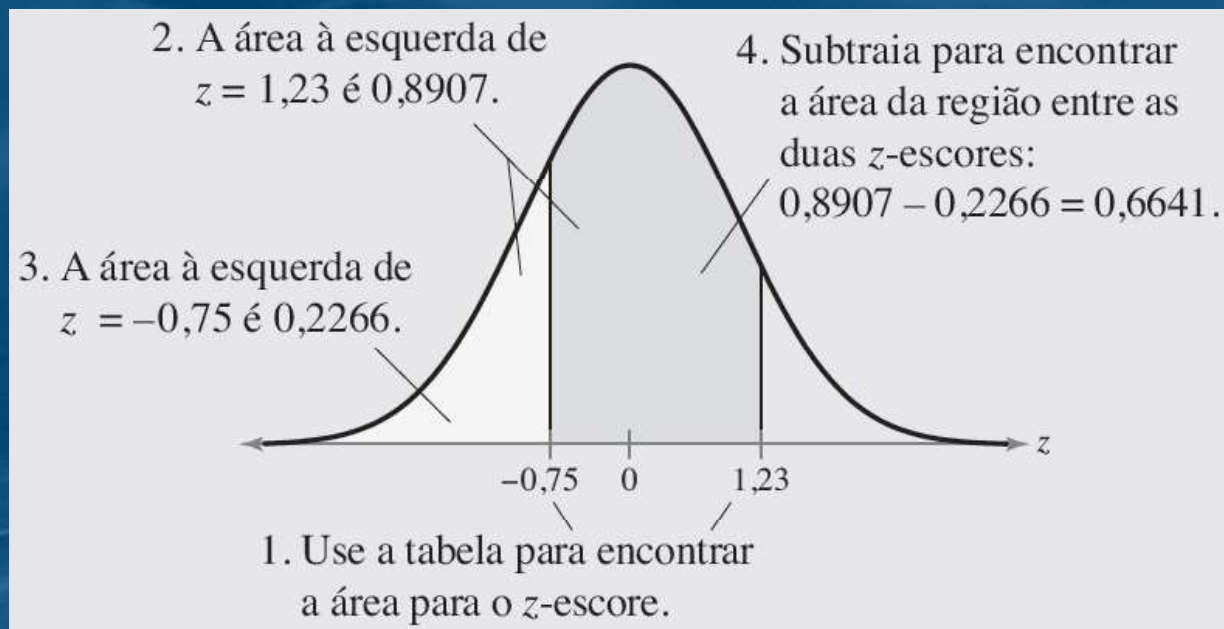
1. *Esboce a curva normal padrão e preencha a área apropriada abaixo da curva.*
2. *Encontre a área seguindo as direções para cada caso.*
 - a. Para encontrar a área à esquerda de z , encontre a área que corresponda a z na tabela normal padrão.



- b. Para encontrar a área à *direita* de z , use a tabela normal padrão para encontrar a área correspondente a z . Então subtraia a área de 1.



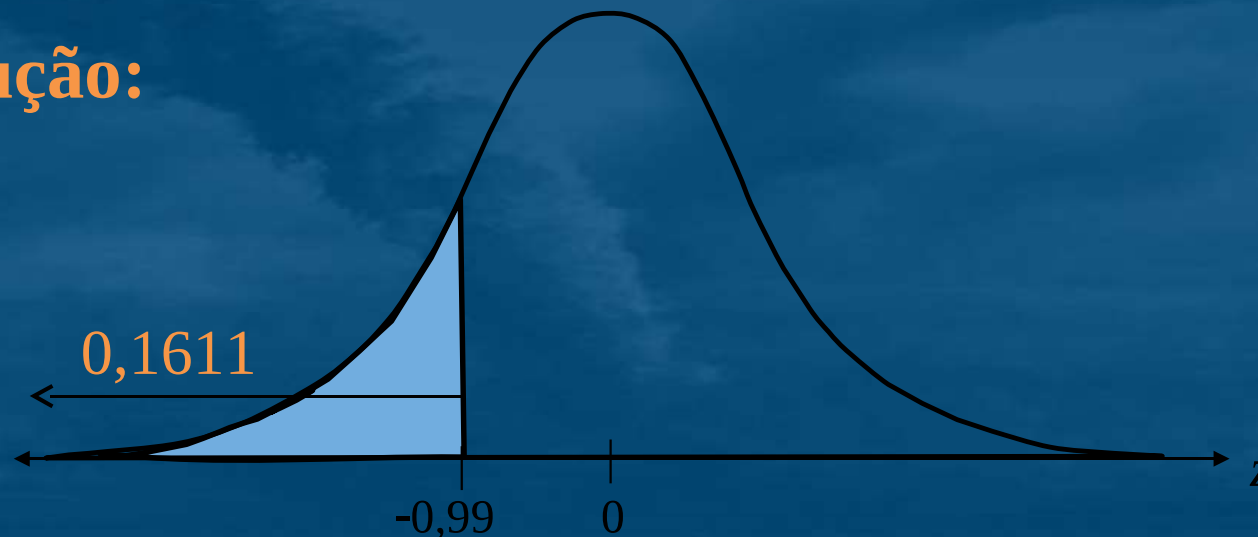
- c. Para encontrar a área *entre* dois escores z , encontre a área correspondente a cada escore z na tabela normal padrão. Então subtraia a área menor da área maior.



Exemplo: encontrando a área sob a curva normal padrão

Encontre a área sob a curva normal padrão à esquerda de $z = -0,99$.

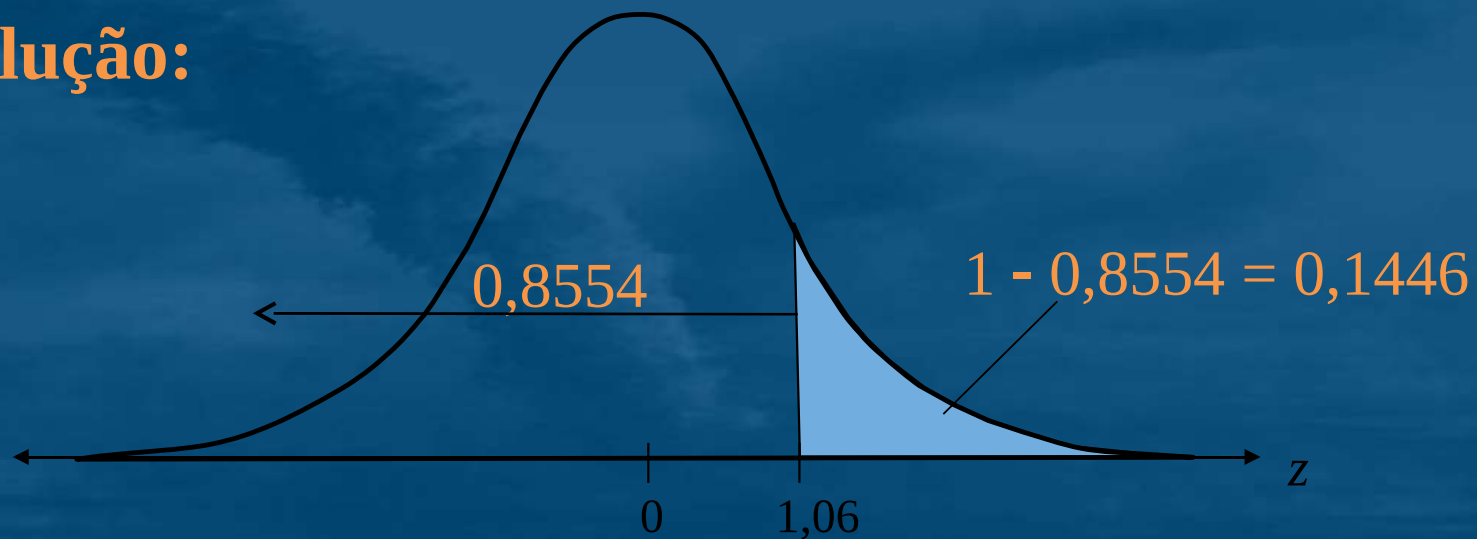
Solução:



Pela tabela normal padrão, a área é igual a $0,1611$.

Encontre a área sob a curva normal padrão à esquerda de $z = 1,06$.

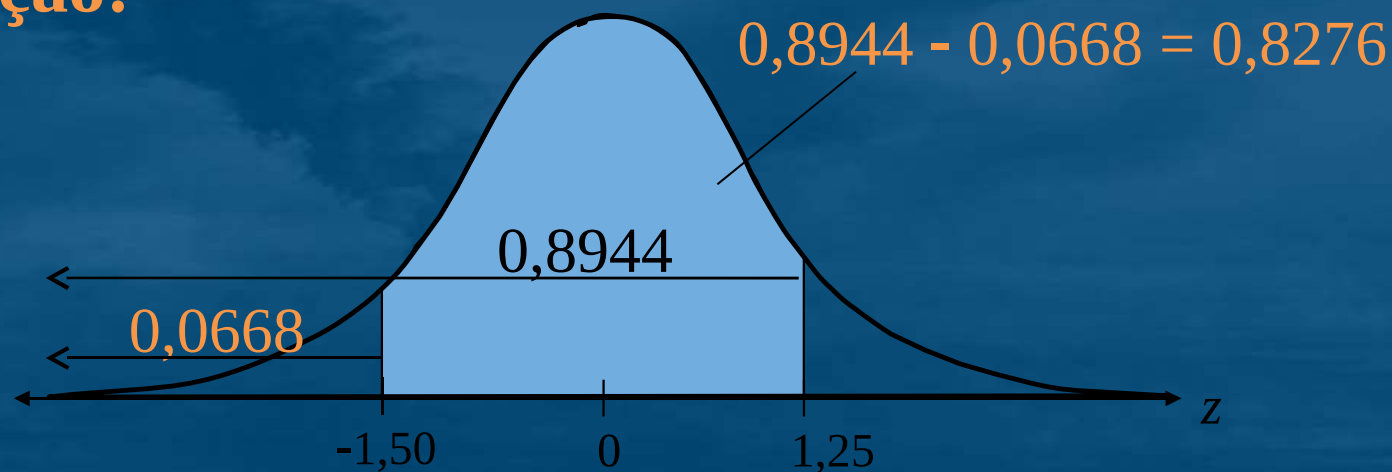
Solução:



Pela tabela normal padrão, a área é igual a 0,1446.

Encontre a área sob a curva normal padrão à esquerda de $z = -1,5$ e $z = 1,25$.

Solução:



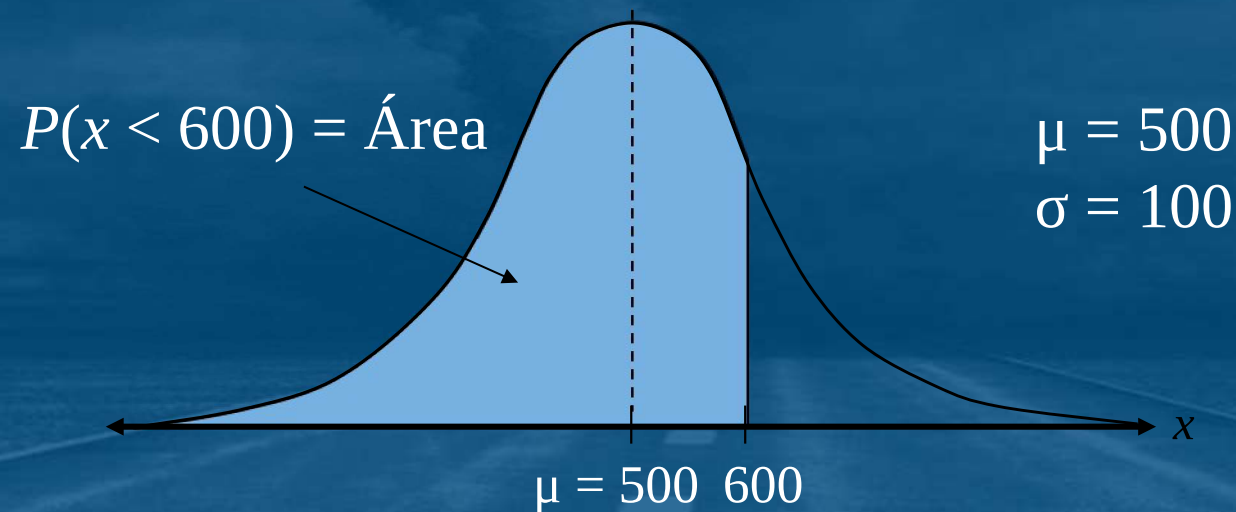
Pela tabela normal padrão, a área é igual a 0,8276.

Seção 5.2

Distribuições normais: encontrando probabilidades

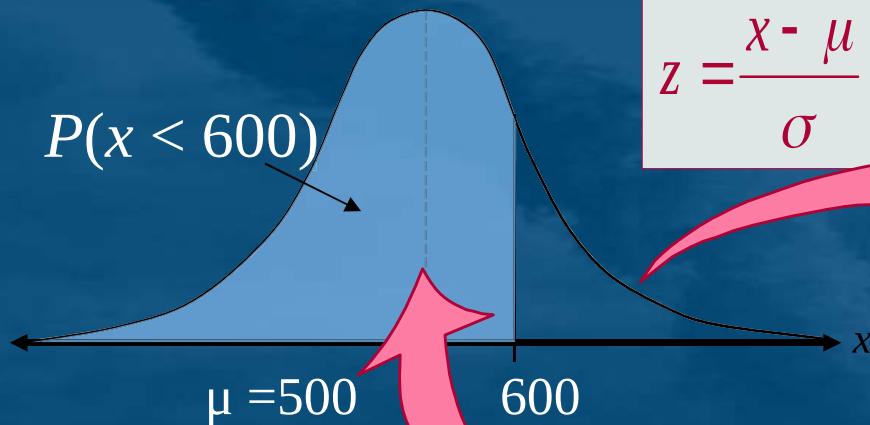
Probabilidade e distribuições normais

- *Se uma variável aleatória x é normalmente distribuída, você pode encontrar a probabilidade de que x cairá em um dado intervalo, calculando a área sob a curva normal daquele intervalo.*



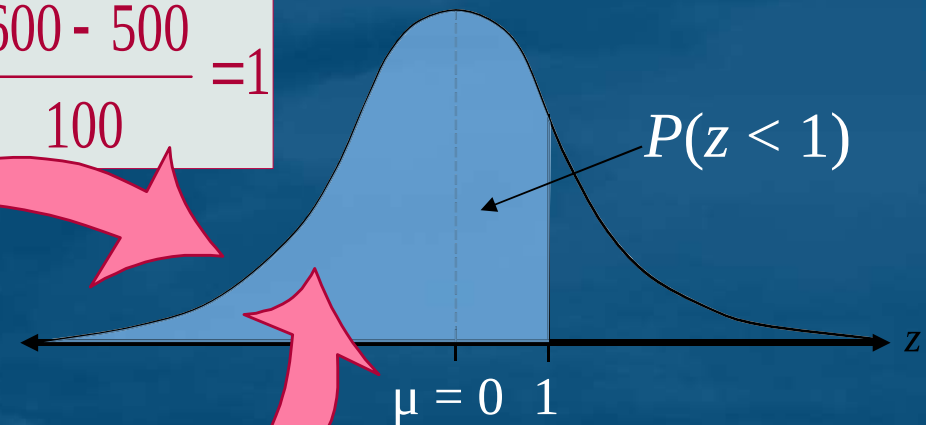
Distribuição normal

$$\mu = 500 \quad \sigma = 100$$



Distribuição normal padrão

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



$$z = \frac{x - \mu}{\sigma} = \frac{600 - 500}{100} = 1$$

Mesma área

$$P(x < 500) = P(z < 1)$$

Exemplo: encontrando probabilidades para distribuições normais

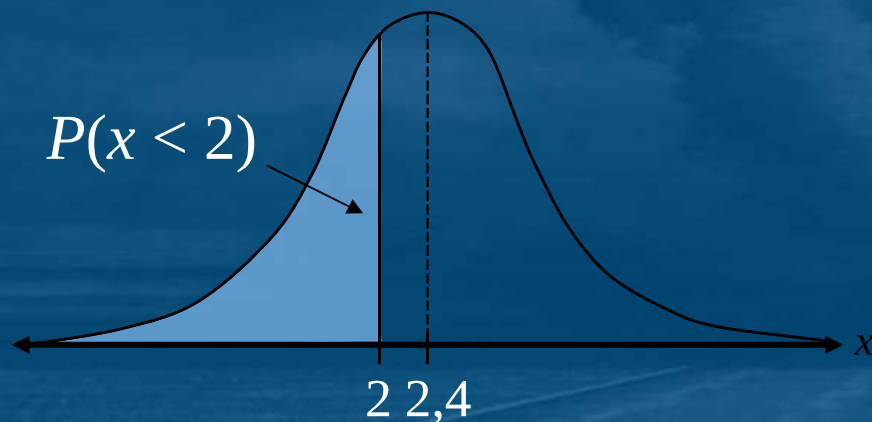
Uma pesquisa indica que pessoas usam seus computadores uma média de 2,4 anos antes de trocá-los por uma máquina nova. O desvio padrão é de 0,5 ano. Um proprietário de computador é selecionado aleatoriamente. Encontre a probabilidade de que ele use sua máquina por menos de dois anos antes de comprar uma nova. Assuma que a variável x é normalmente distribuída.



Solução: encontrando probabilidades para distribuições normais

Distribuição normal

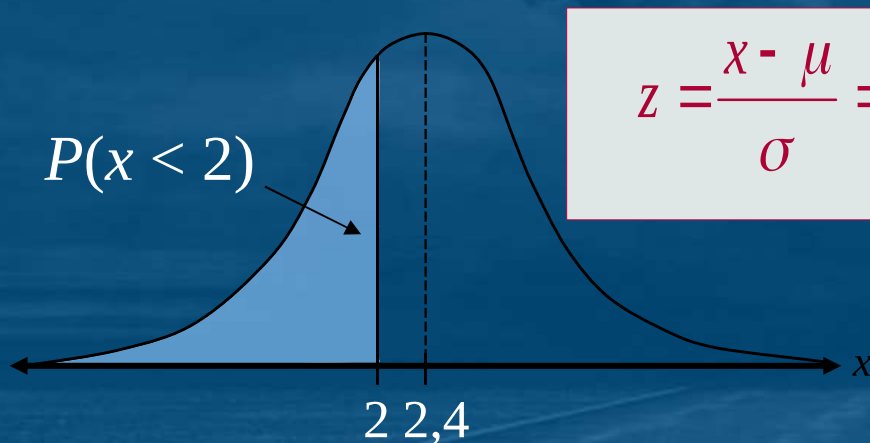
$$\mu = 2,4 \quad \sigma = 0,5$$



Solução: encontrando probabilidades para distribuições normais

Distribuição normal

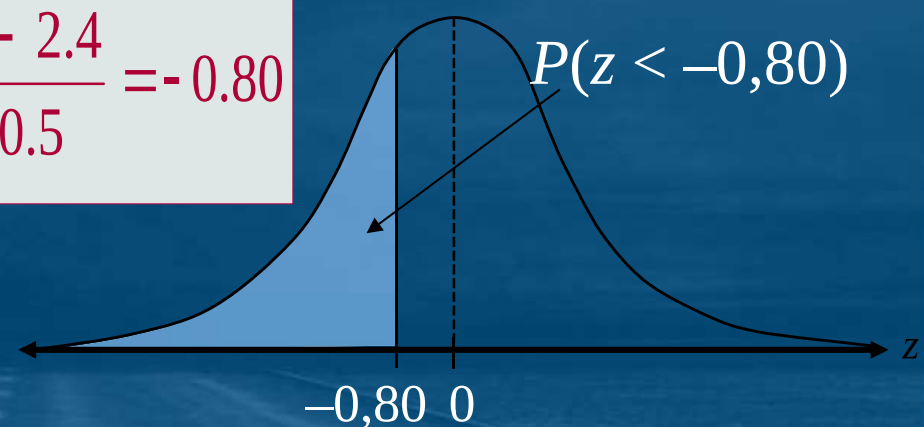
$$\mu = 2,4 \quad \sigma = 0,5$$



$$z = \frac{x - \mu}{\sigma} = \frac{2 - 2,4}{0,5} = -0,80$$

Distribuição normal padrão

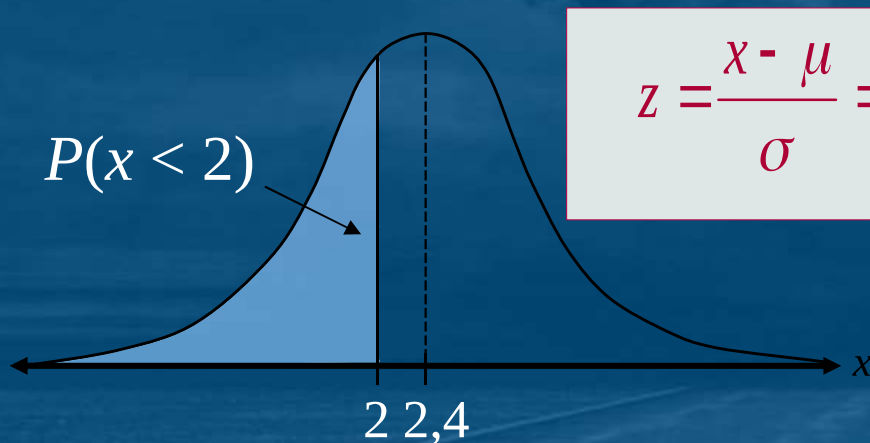
$$\mu = 0 \quad \sigma = 1$$



Solução: encontrando probabilidades para distribuições normais

Distribuição normal

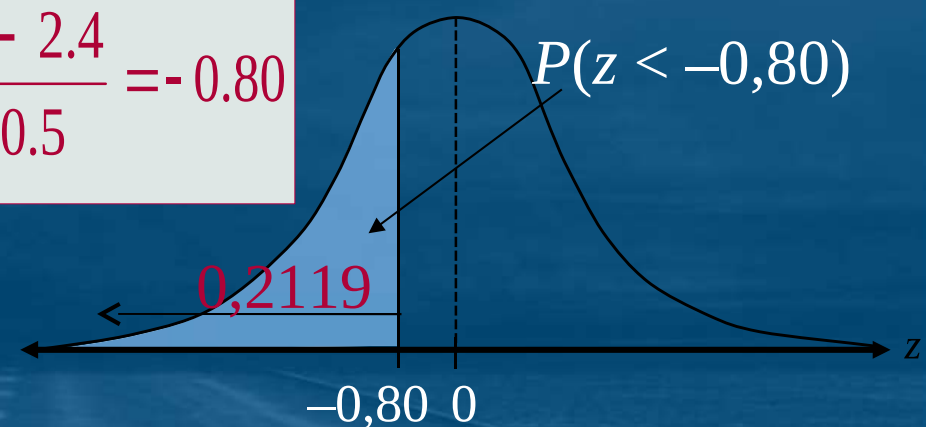
$$\mu = 2,4 \quad \sigma = 0,5$$



$$z = \frac{x - \mu}{\sigma} = \frac{2 - 2.4}{0.5} = -0.80$$

Distribuição normal padrão

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



$$P(x < 2) = P(z < -0,80) = 0,2119$$

Exemplo: encontrando probabilidades para distribuições normais

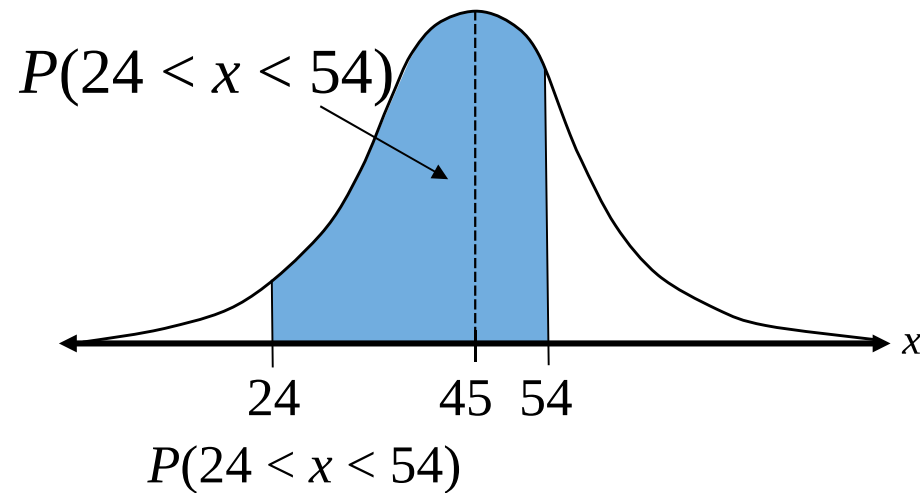
Uma pesquisa indica que para cada ida ao supermercado, um comprador gasta uma média de 45 minutos com um desvio padrão de 12 minutos no mercado. O período de tempo gasto no mercado é normalmente distribuído e representado pela variável x . Um cliente entra no mercado. Encontre a probabilidade de que ele passe entre 24 e 54 minutos dentro do mercado.



Solução: encontrando probabilidades para distribuições normais

Distribuição normal

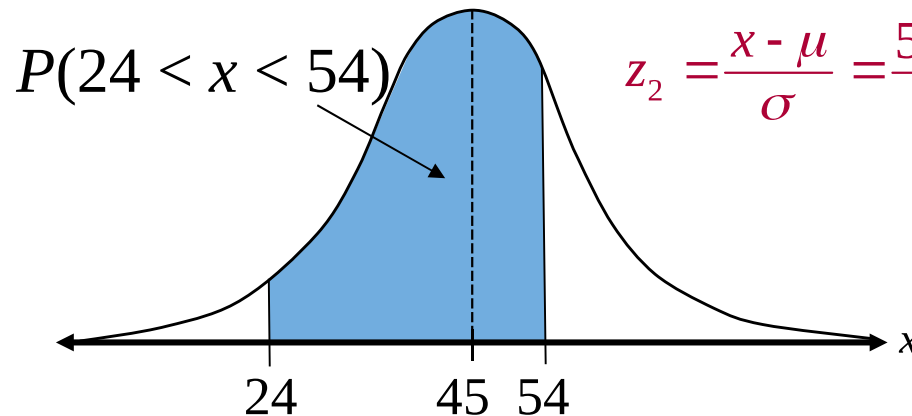
$$\mu = 45 \quad \sigma = 12$$



Solução: encontrando probabilidades para distribuições normais

Distribuição normal

$$\mu = 45 \quad \sigma = 12$$



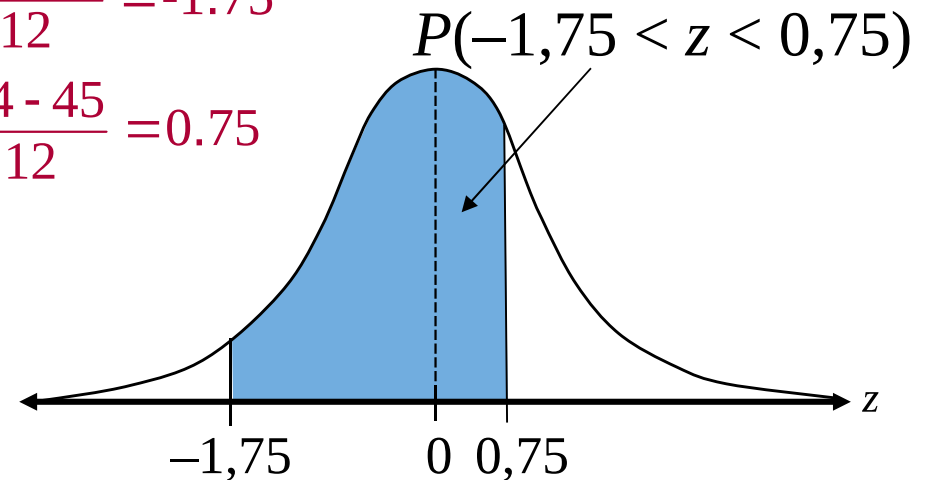
$$P(24 < x < 54)$$

$$z_1 = \frac{x - \mu}{\sigma} = \frac{24 - 45}{12} = -1.75$$

$$z_2 = \frac{x - \mu}{\sigma} = \frac{54 - 45}{12} = 0.75$$

Distribuição normal padrão

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



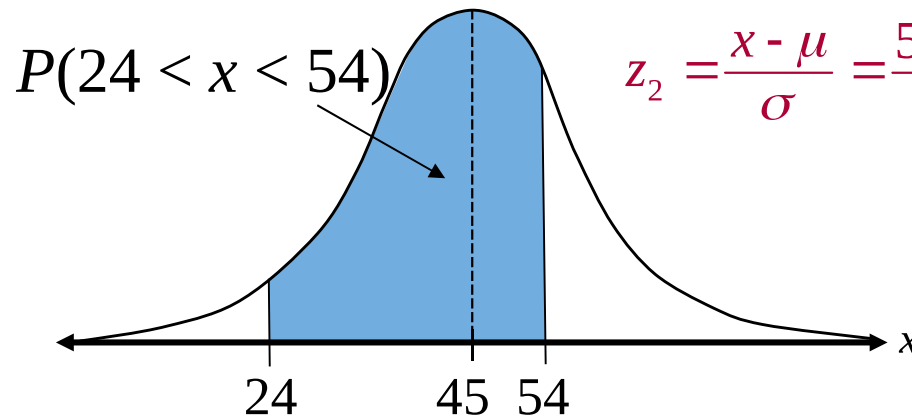
$$P(-1.75 < z < 0.75)$$

$$P(24 < x < 54) = P(-1.75 < z < 0.75)$$

Solução: encontrando probabilidades para distribuições normais

Distribuição normal

$$\mu = 45 \quad \sigma = 12$$



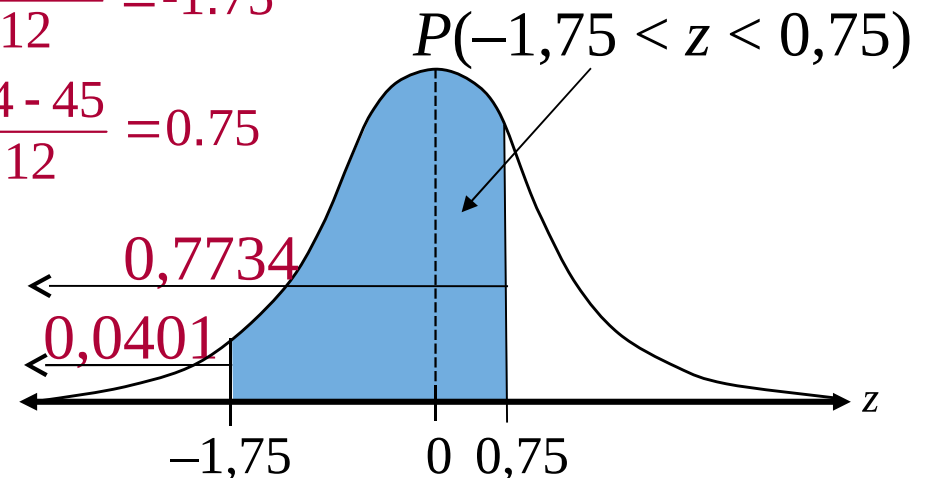
$$P(24 < x < 54) = P(-1,75 < z < 0,75) \\ = 0,7734 - 0,0401 = \mathbf{0,7333}$$

Distribuição normal padrão

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

$$z_1 = \frac{x - \mu}{\sigma} = \frac{24 - 45}{12} = -1,75$$

$$z_2 = \frac{x - \mu}{\sigma} = \frac{54 - 45}{12} = 0,75$$



Exemplo: encontrando probabilidades para distribuições normais

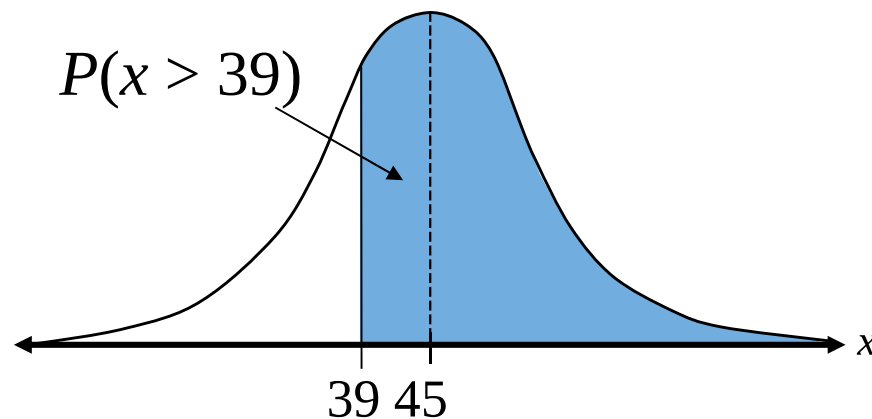
Encontre a probabilidade de que o cliente fique no mercado mais de 39 minutos. (Lembre-se: $\mu = 45$ minutos e $\sigma = 12$ minutos.)



Solução: encontrando probabilidades para distribuições normais

Distribuição normal

$$\mu = 45 \quad \sigma = 12$$

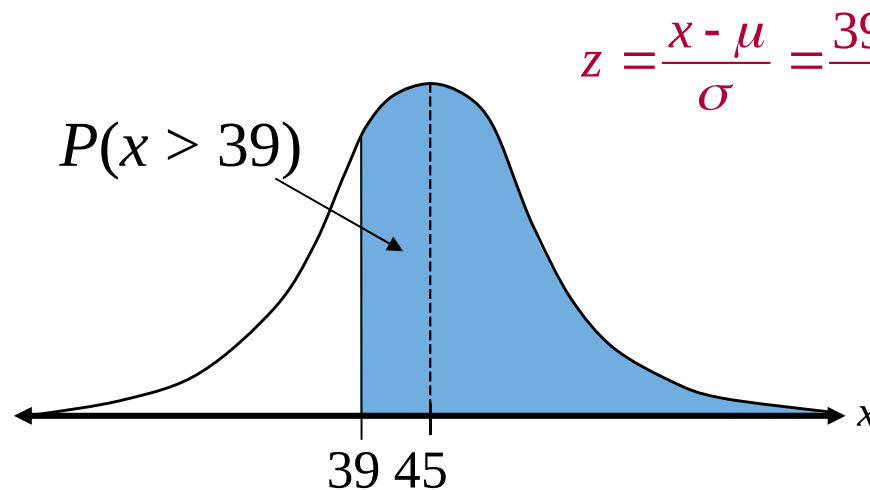


$$P(x > 39)$$

Solução: encontrando probabilidades para distribuições normais

Distribuição normal

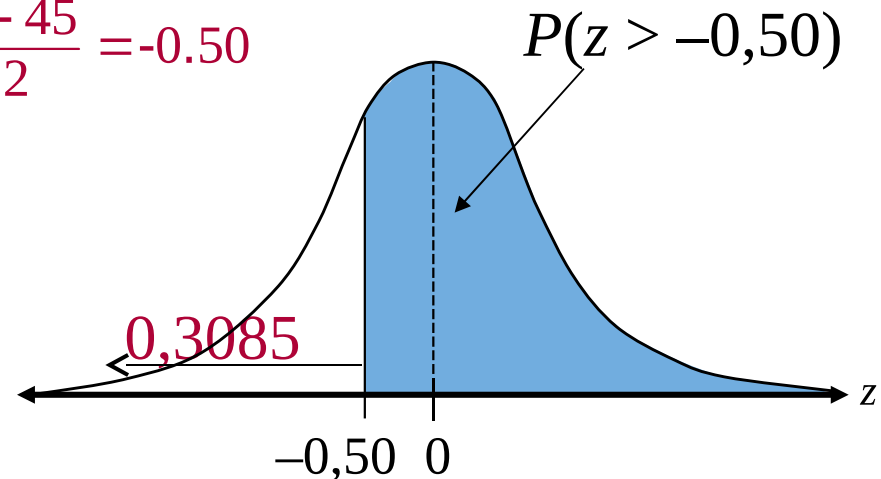
$$\mu = 45 \quad \sigma = 12$$



$$z = \frac{x - \mu}{\sigma} = \frac{39 - 45}{12} = -0.50$$

Distribuição normal padrão

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



$$P(x > 39) = P(z > -0,50) = 1 - 0,3085 = \mathbf{0,6915}$$

Exemplo: encontrando probabilidades para distribuições normais

Se 200 clientes entram no mercado, quantos deles você esperaria que permanecessem por mais de 39 minutos?

Solução:

Lembre-se: $P(x > 39) = 0,6915$

$200(0,6915) = 138,3$ (ou cerca de 138) clientes



Seção 5.4

Distribuições amostrais e o Teorema Central do Limite

Distribuições de amostras de médias amostrais

População com μ, σ

Exemplo 1
 \bar{X}_1

Exemplo 2
 \bar{X}_2

Exemplo 3
 \bar{X}_3

Exemplo 4
 \bar{X}_4

Exemplo 5
 \bar{X}_5

A distribuição da amostragem consiste dos valores das médias amostrais, $\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4, \bar{X}_5, \dots$

Propriedades de distribuições de amostras de médias amostrais

1. *A média das médias amostrais, $\mu_{\bar{x}}$, é igual à média populacional μ .*

$$\mu_{\bar{x}} = \mu$$

2. O desvio padrão das médias amostrais, $\sigma_{\bar{x}}$, é igual ao desvio padrão da população, σ dividido pela raiz quadrada do tamanho da amostragem, n .

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- Chamado de **erro padrão da média**

Exemplo: distribuições de amostras de médias amostrais

Os valores populacionais {1, 3, 5, 7} são escritos em pedaços de papel e postos em uma caixa. Dois pedaços de papel são aleatoriamente selecionados, sendo recolocados na caixa após cada seleção.

- a. Encontre a média, a variação e o desvio padrão da população.

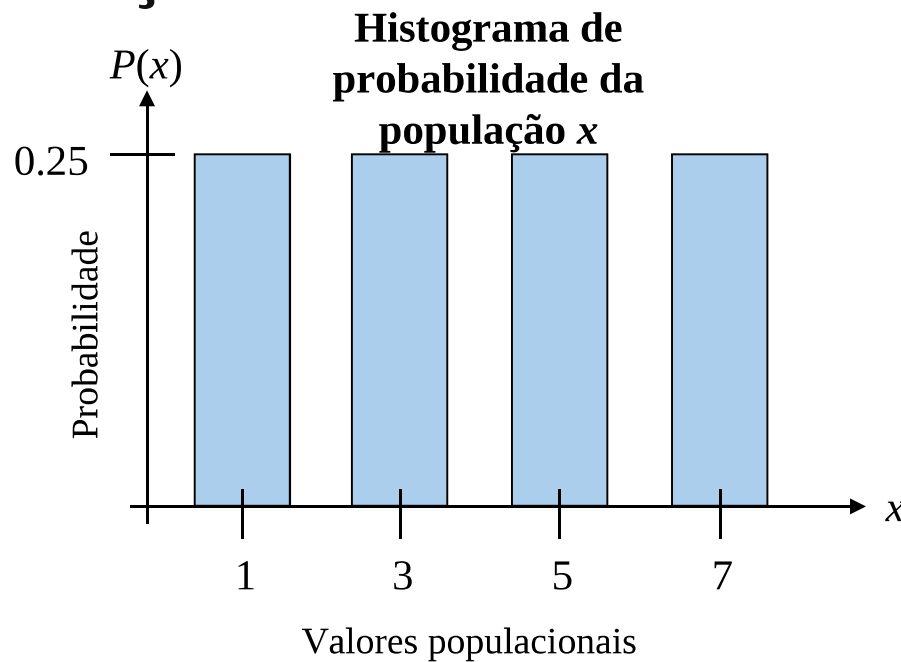
Solução: Média $\mu = \frac{\sum x}{N} = 4$

Variância $\sigma^2 = \frac{\sum (x - \mu)^2}{N} = 5$

Desvio padrão $\sigma = \sqrt{5} \approx 2.236$

- b. Faça o gráfico do histograma de probabilidade dos valores populacionais.

Solução:



Todos os valores têm a mesma probabilidade de serem selecionados (distribuição uniforme)

- c. Liste todas as amostragens possíveis de tamanho $n = 2$ e calcule a média de cada amostragem.

Solução:

Amostragem \bar{x}		Amostragem \bar{x}		Essas médias formam a distribuição amostral das médias amostrais
1, 1	1	5, 1	3	
1, 3	2	5, 3	4	
1, 5	3	5, 5	5	
1, 7	4	5, 7	6	
3, 1	2	7, 1	4	
3, 3	3	7, 3	5	
3, 5	4	7, 5	6	
3, 7	5	7, 7	7	

d. Construa a distribuição de probabilidade das médias amostrais.

Solução:

\bar{x}	f	Probabilidade
1	1	0,0625
2	2	0,1250
3	3	0,1875
4	4	0,2500
5	3	0,1875
6	2	0,1250
7	1	0,0625

- e. Encontre a média, a variância e o desvio padrão da distribuição amostral das médias amostrais.

Solução:

A média, a variância, e o desvio padrão de 16 amostras são:

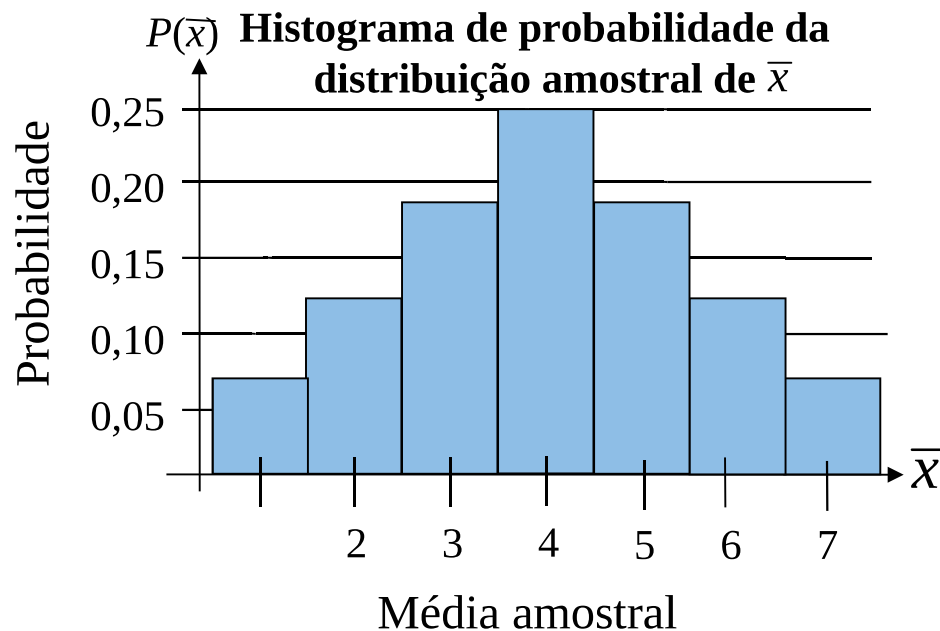
$$\mu_{\bar{x}} = 4 \quad \sigma_{\bar{x}}^2 = \frac{5}{2} = 2.5 \quad \sigma_{\bar{x}} = \sqrt{2.5} \approx 1.581$$

Esses resultados satisfazem as propriedades de distribuições de amostras de médias amostrais.

$$\mu_{\bar{x}} = \mu = 4 \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{5}}{\sqrt{2}} \approx \frac{2.236}{\sqrt{2}} \approx 1.581$$

- f. Faça o gráfico do histograma de probabilidade das médias amostrais.

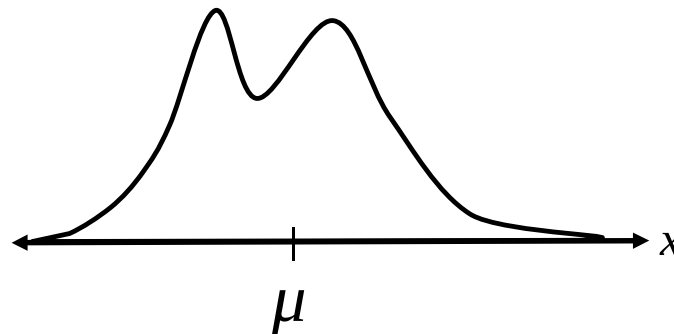
Solução:



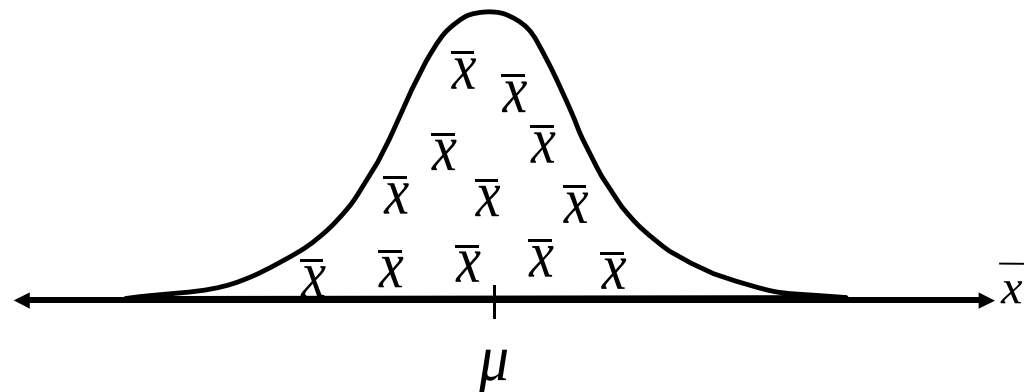
O gráfico é simétrico e em formato de sino. Aproxima-se de uma distribuição normal

O Teorema Central do Limite

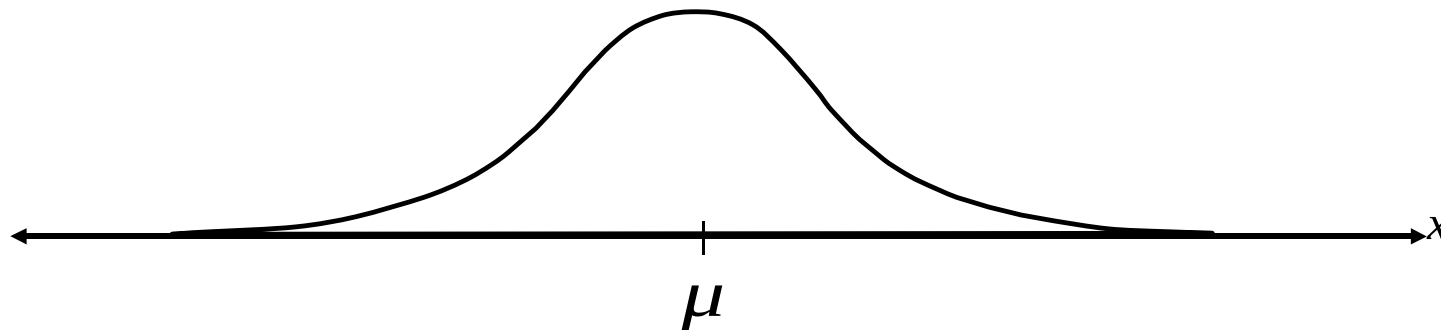
1. Se amostragens de tamanho $n \geq 30$ são tiradas de qualquer população de média $= \mu$ e desvio padrão $= \sigma$,



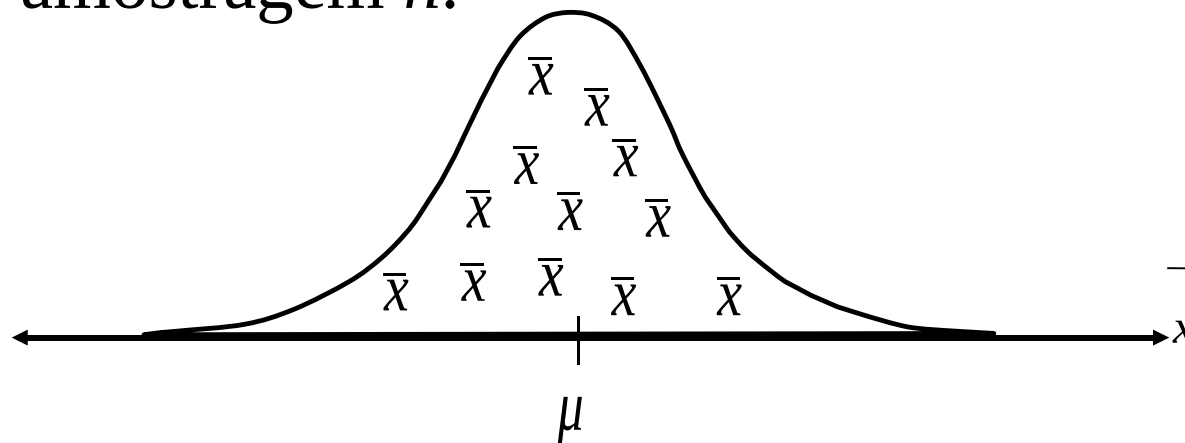
então a distribuição de amostras da média amostral aproxima-se de uma distribuição normal. Quanto maior o tamanho da amostragem, melhor a aproximação.



2. Se a própria população é normalmente distribuída,



a distribuição de amostras das médias amostrais é normalmente distribuída para **qualquer** tamanho de amostragem n .



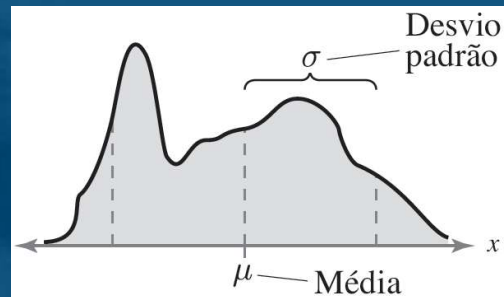
- *Em ambos os casos, a distribuição de amostras de médias amostrais tem uma média igual à média da população.*

- *A distribuição da amostra de médias amostrais tem uma variância igual a $1/n$ vez a variância da população e um desvio padrão igual ao desvio padrão da população dividido pela raiz quadrada de n .*

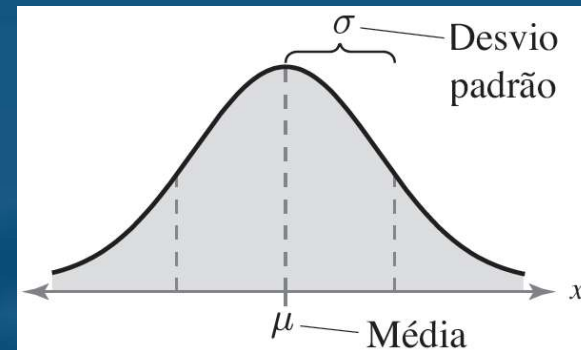
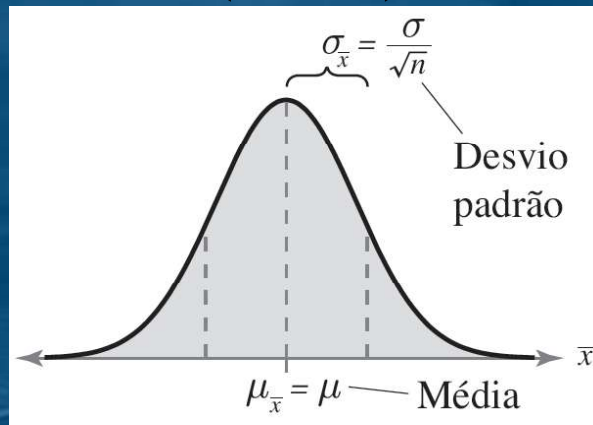
$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \quad \text{Variância}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{Desvio padrão (erro padrão da média)}$$

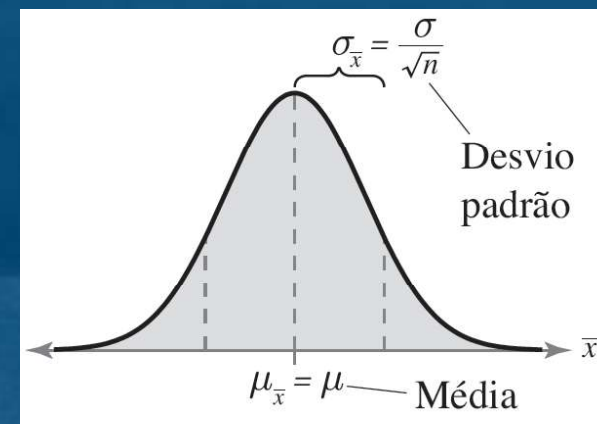
1. Distribuição populacional qualquer
2. Distribuição populacional normal



Distribuição das médias amostrais ($n \geq 30$)



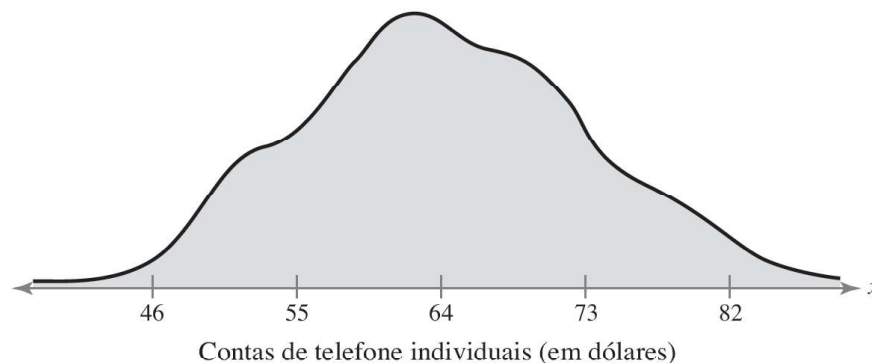
Distribuição das médias amostrais (n qualquer)



Exemplo: interpretando o Teorema Central do Limite

As contas dos telefones dos habitantes de uma cidade têm uma média de \$ 64 e um desvio padrão de \$ 9. Amostragens aleatórias de 36 contas de telefone são tiradas dessa população e a média de cada amostragem é determinada. Encontre a média e o erro padrão da média da distribuição amostral. Então esboce um gráfico da distribuição amostral das médias amostrais

Distribuição para todas as contas de telefone



Solução: interpretando o Teorema Central do Limite

- *A média da distribuição de amostras é igual à média da população*

$$\mu_{\bar{x}} = \mu = 64$$

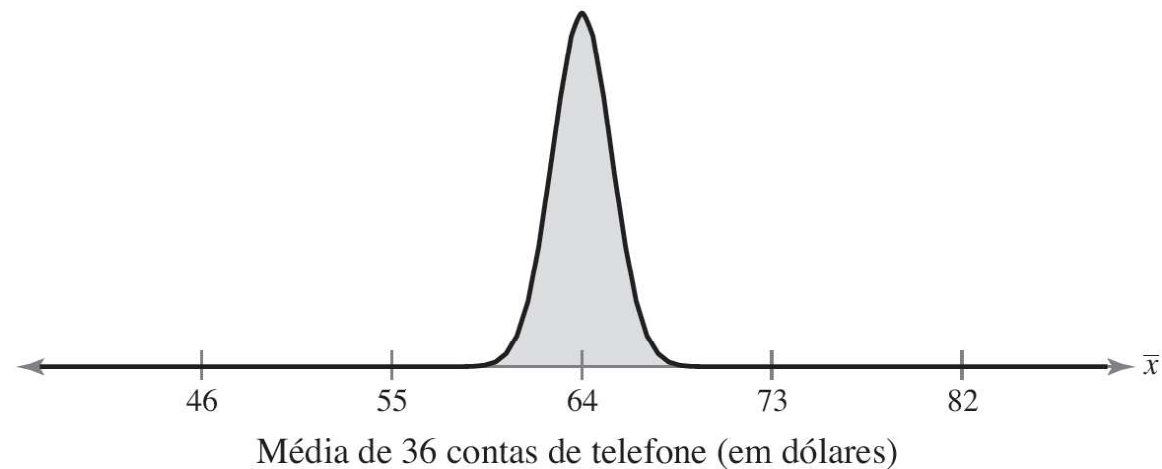
- *O erro padrão da média é igual ao desvio padrão populacional dividido pela raiz quadrada de n*

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{36}} = 1.5$$

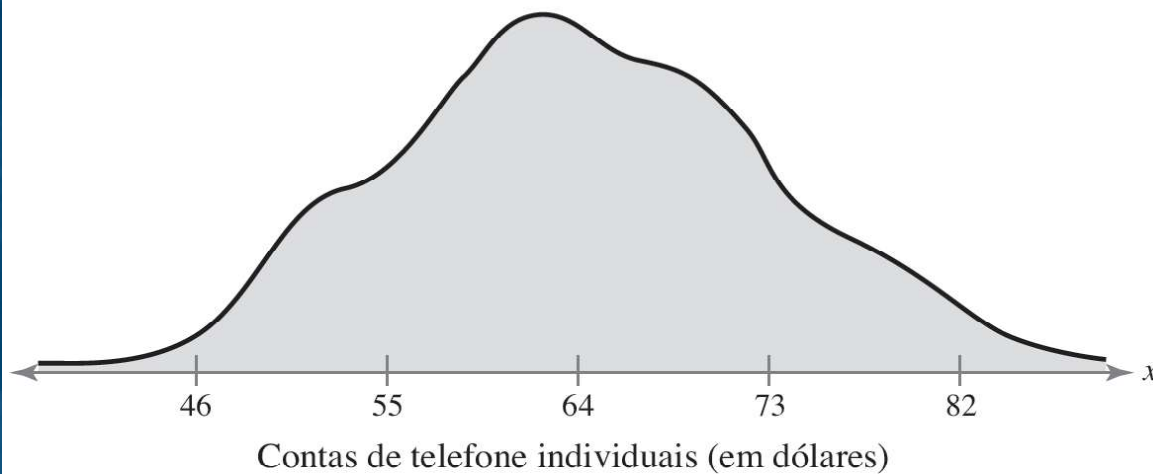
- Já que o tamanho da amostragem é maior que 30, a distribuição das amostras pode ser aproximada por uma distribuição normal com*

$$\mu_{\bar{x}} = 64 \qquad \sigma_{\bar{x}} = 1.5$$

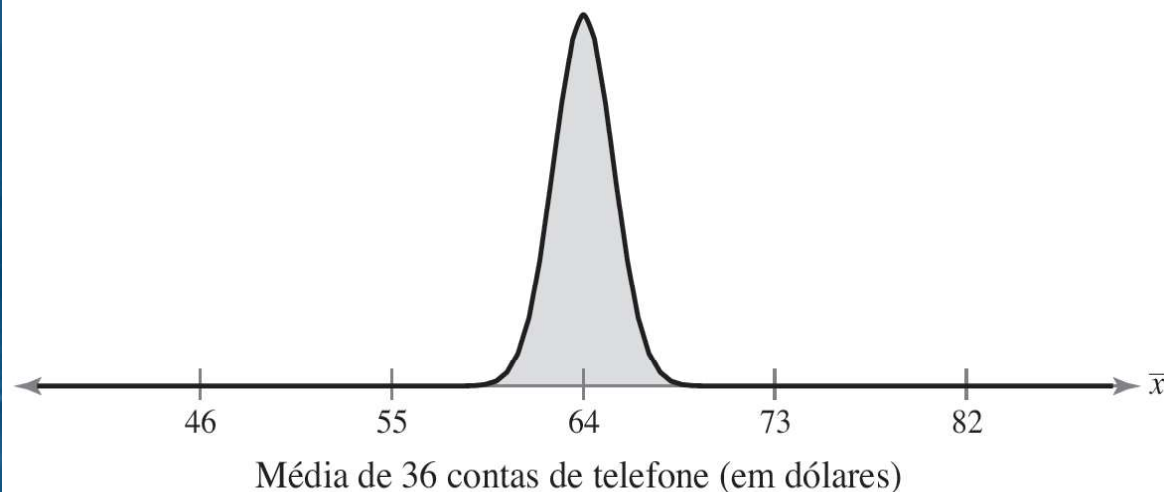
Distribuição das médias das amostras com $n = 36$



Distribuição para todas as contas de telefone



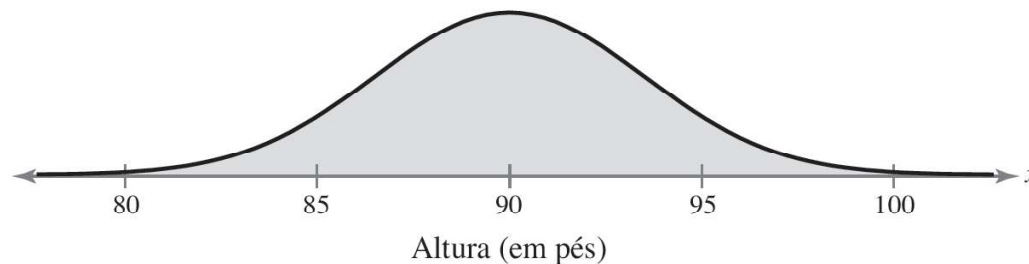
Distribuição das médias das amostras com $n = 36$



Exemplo: interpretando o Teorema Central do Limite

As alturas das árvores de carvalho branco adultas são normalmente distribuídas, com uma média de 90 pés e um desvio padrão de 3,5 pés. Amostras aleatórias de tamanho 4 são tiradas dessa população, e a média de cada amostra é determinada. Encontre a média e o erro padrão da média da distribuição amostral. Então esboce um gráfico da distribuição amostral das médias amostrais.

Distribuição das alturas da população



Solução: interpretando o Teorema Central do Limite

- *A média da distribuição amostral é igual à média populacional*

$$\mu_{\bar{x}} = \mu = 90$$

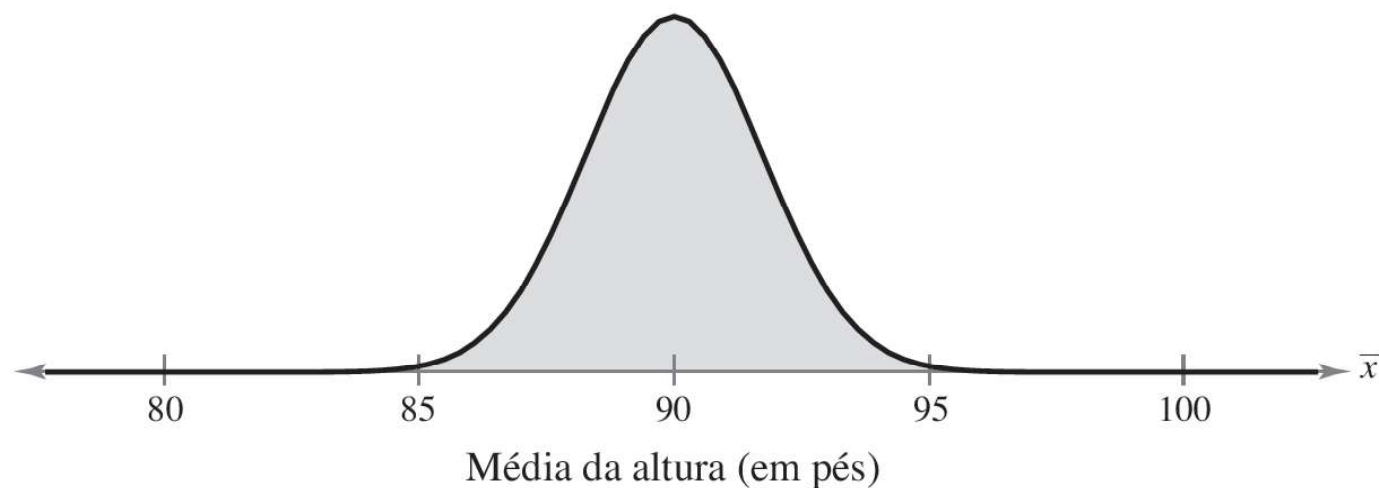
- *O erro padrão da média é igual ao desvio padrão da população dividido pela raiz quadrada de n.*

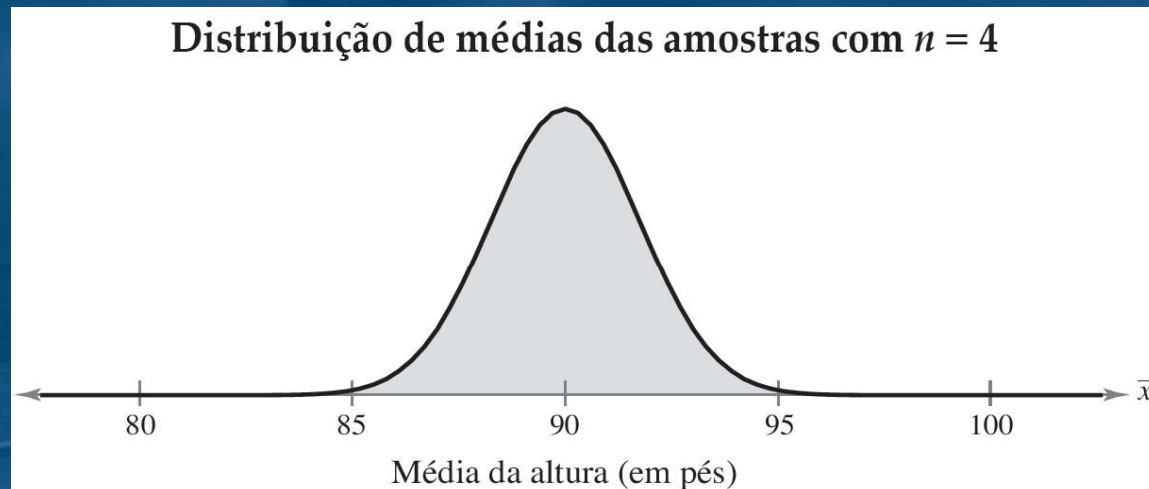
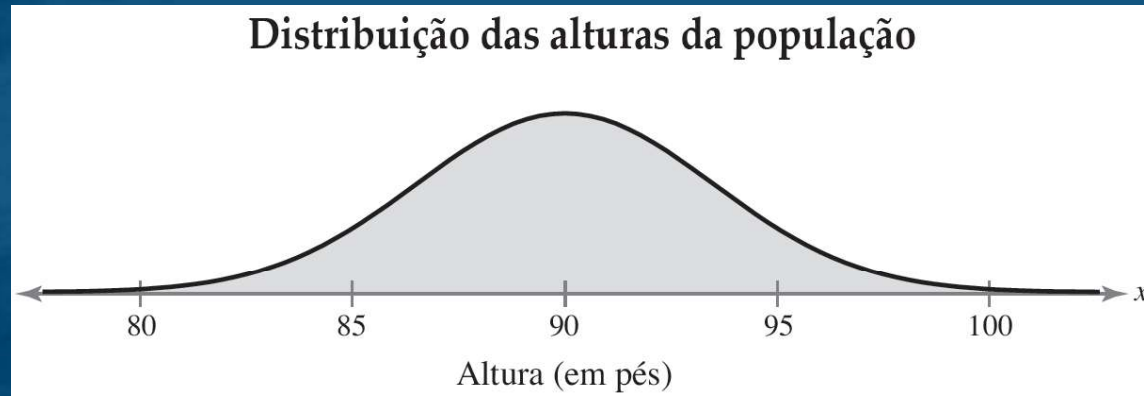
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{4}} = 1.75$$

Já que a população é normalmente distribuída, a distribuição amostral da média amostral também é normalmente distribuída.

$$\mu_{\bar{x}} = 90 \qquad \sigma_{\bar{x}} = 1.75$$

Distribuição de médias das amostras com $n = 4$





Exemplo: probabilidades para x e \bar{x}

Um auditor de um certo banco afirma que os balanços dos cartões de crédito são normalmente distribuídos com uma média de R\$ 2.870 e um desvio padrão de R\$ 900.

1. Qual é a probabilidade de que um portador de cartão de crédito aleatoriamente selecionado tenha um balanço menor que R\$ 2.500?



Solução:

Foi pedido que encontrássemos a probabilidade associada com um certo valor da variável aleatória x .

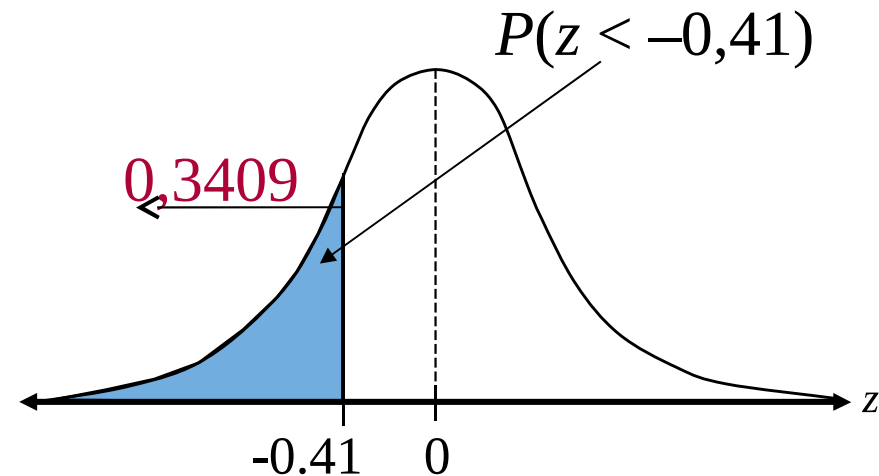
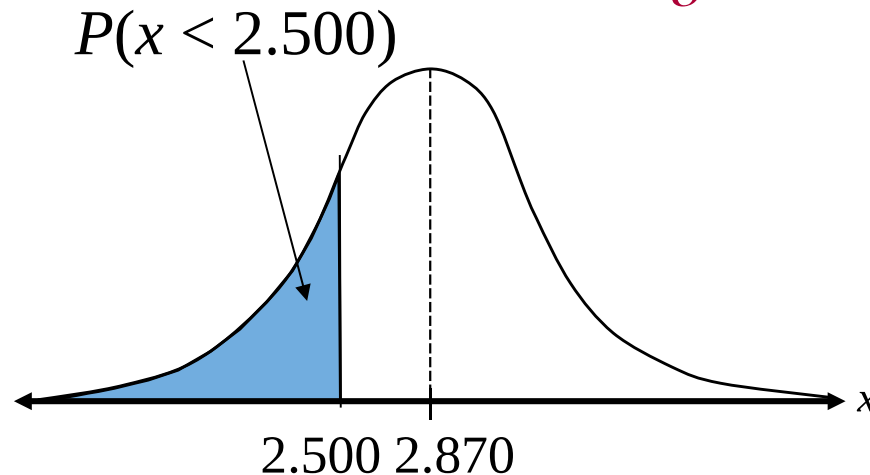
Distribuição normal

$$\mu = 2.870 \quad \sigma = 900$$

Distribuição normal padrão

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

$$z = \frac{x - \mu}{\sigma} = \frac{2500 - 2870}{900} \approx -0.41$$



$$P(x < 2.500) = P(z < -0,41) = \mathbf{0,3409}$$

Exemplo: probabilidades para x e \bar{x}

2. *Você seleciona aleatoriamente 25 portadores de cartão de crédito.*

Qual é a probabilidade de que a média dos balanços dos seus cartões de crédito seja menor que R\$ 2.500?



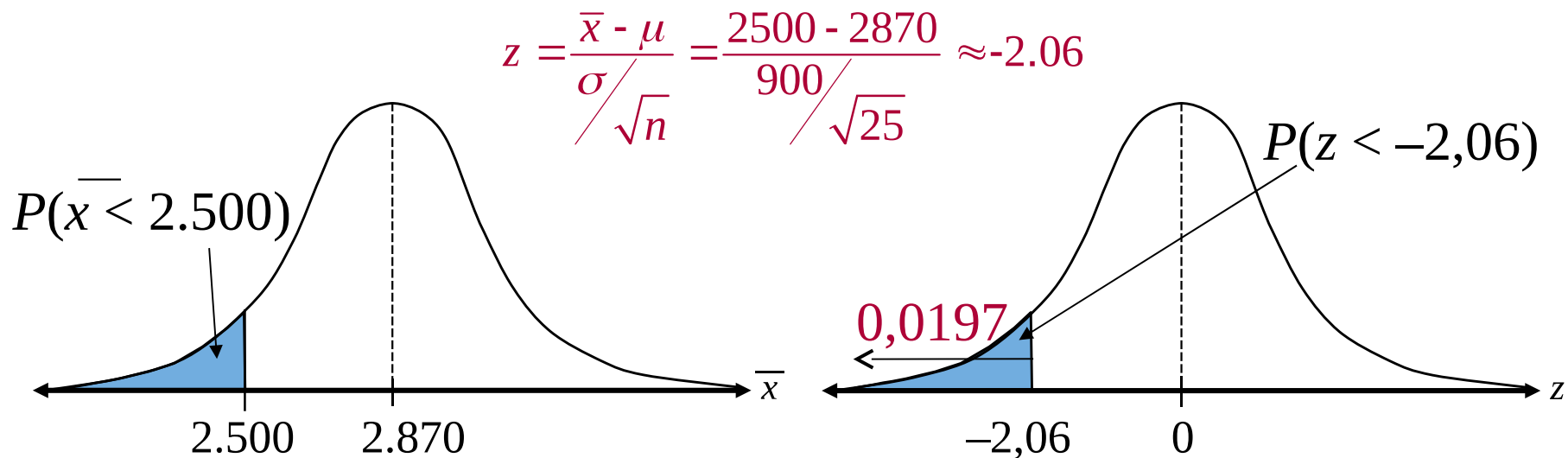
Solução:

Foi pedido que encontrássemos a probabilidade associada com uma média amostral \bar{x} .

$$\mu_{\bar{x}} = \mu = 2870 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{900}{\sqrt{25}} = 180$$

Distribuição normal
 $\mu = 2.870$ $\sigma = 180$

Distribuição normal padrão
 $\mu = 0$ $\sigma = 1$



$$P(\bar{x} < 2.500) = P(z < -2.06) = \mathbf{0,0197}$$

Solução: probabilidades para x e \bar{x}

- *Existe uma chance de 34% para o indivíduo ter um balanço menor que \$ 2.500*
- *Existe apenas 2% de chance para que a média de uma amostragem de 25 tenha um balanço menor que \$ 2.500 (evento incomum)*
- *É possível que a amostragem seja incomum ou é possível que a afirmação do auditor, de que a média é \$ 2.870, seja incorreta*