Capítulo 5

Distribuições de Probabilidades Contínuas

Texto baseado no livro:

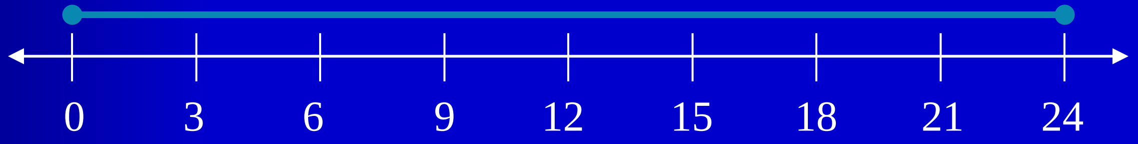
Estatística Aplicada - Larson / Farber – Editora Pearson – 2010

Variável aleatória contínua

Variável aleatória contínua

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

Horas gastas estudando durante um dia



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

Considere a variável aleatória contínua X em que definimos a função $f(x)$ denominada *função densidade de probabilidade* e que tem as seguintes propriedades:

A probabilidade da variável aleatória X é sempre definida num intervalo de valores dessa variável X, por exemplo, (x_1, x_2).

A probabilidade da variável aleatória X é medida pela área sob a curva da função densidade $f(x)$ num determinado intervalo.

A área total sob a curva $f(x)$ é igual a 1 ou 100%.

Variável Aleatória Contínua

Para a variável aleatória contínua *X* que assume valores do conjunto dos números reais há uma função matemática *f*(*x*) com as seguintes premissas:

A função densidade de probabilidade *f*(*x*) é sempre positiva, para todo *x* pertencente a *X*.

A área sob a função *f*(*x*) entre os limites menos infinito e mais infinito da variável aleatória contínua *X* é igual a 1 ou 100%:

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

Variável Aleatória Contínua

A probabilidade da *VA* contínua *X* dentro do intervalo (*a*, *b*) com ambos limites incluídos é medida pela área definida pela função *f*(*x*) entre os limites *a* e *b*:

$$P(a \leq x \leq b) = \int_a^b f(x)\,dx$$

Um ponto *f*(*x*) da função densidade não é a probabilidade do valor *x* da variável aleatória *X*, pois, por exemplo, o ponto *f*(*x*=*a*) da função densidade é zero.

Como deve-se representar os limites do cálculo da probabilidade de uma variável aleatória contínua dentro do intervalo (a, b), $P(a \leq X \leq b)$ ou $P(a < X < b)$?

As duas representações podem ser utilizadas, incluindo a representação com limites mistos, por exemplo: $P(a \leq X < b)$ ou $P(a < X \leq b)$!

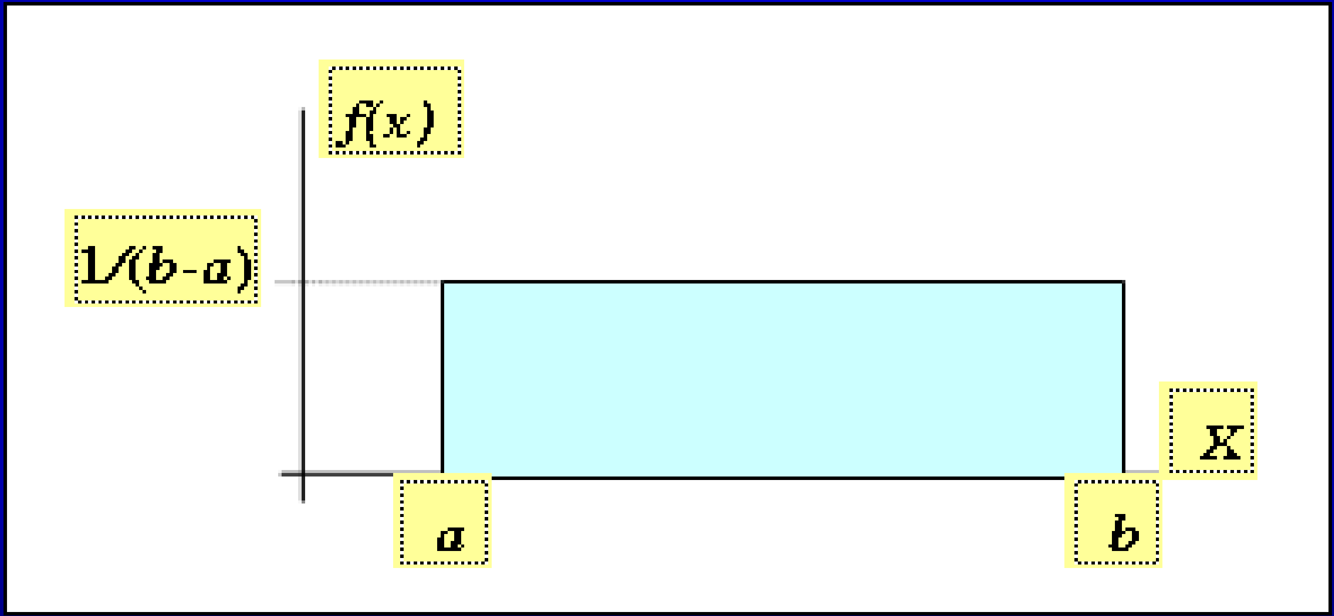
Seja a *VA* contínua *X* com função densidade de probabilidade $f(x)$:

- *Valor esperado* de *X* : $\mu_X = \int_{-\infty}^{+\infty} x\, f(x)\, dx$
- *Variância* de *X*: $\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x)\, dx$ e
- *Desvio padrão* de *X*: $\sigma_X = +\sqrt{\sigma_X^2}$

Uma distribuição de variável aleatória contínua é uma *distribuição uniforme se sua* função densidade de probabilidade é constante dentro de um intervalo de valores da variável aleatória X.

Cada um dos possíveis valores que X com distribuição uniforme pode assumir tem a mesma probabilidade de ocorrer.

A variável aleatória X tem *distribuição uniforme* de probabilidades no intervalo (a,b) se a função densidade $f(x)$ for $f(x)=\dfrac{1}{b-a}$, com as seguintes condições $b \ge a$ e $a \le x \le b$.



A *média* μ_X e a *variância* σ_X^2 da variável aleatória X com *distribuição uniforme* de probabilidades no intervalo (a,b) são:

- *Média*: $\mu_X = \dfrac{a+b}{2}$
- *Variância*: $\sigma_X^2 = \dfrac{(b-a)^2}{12}$

Exemplo

Calcular a média e o desvio padrão da variável aleatória X com distribuição uniforme no intervalo (100, 200).

Solução. A média da variável aleatória contínua X é 150 obtida com a fórmula: $\mu_X = \frac{a+b}{2} = \frac{100+200}{2} = 150$. Da mesma maneira, a variância 833,33 foi obtida com a fórmula: $\sigma_X^2 = \frac{(b-a)^2}{12} = \frac{(200-100)^2}{12} = 833{,}33$.

O desvio padrão é igual a: $\sigma_X = \sqrt{833{,}33} = 28{,}87$

Objetivos da Seção

Distribuição Normal

Interpretar gráficos de distribuição de probabilidade normal

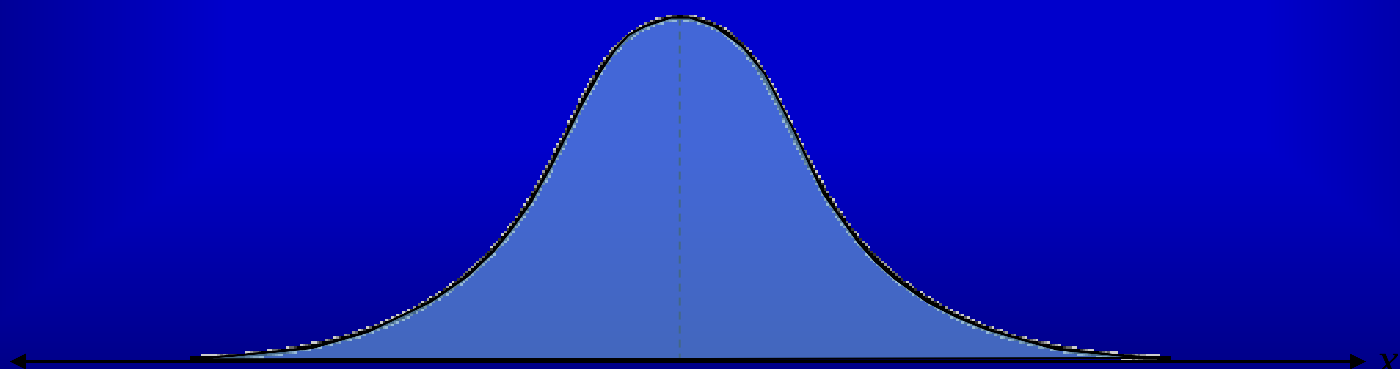
Encontrar áreas sob a curva normal padrão

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**

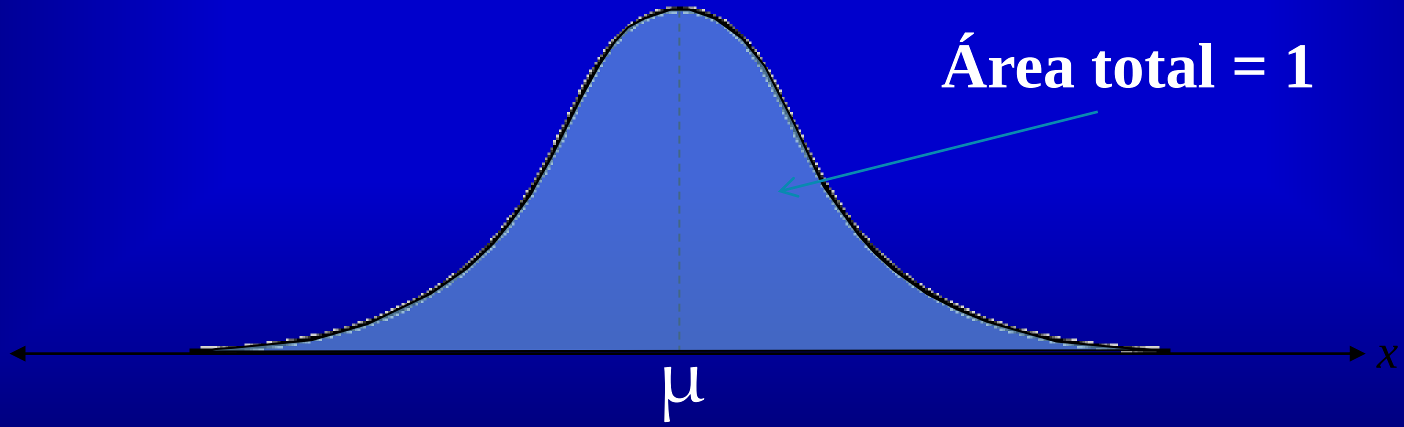


1. A média, a mediana e a moda são iguais.

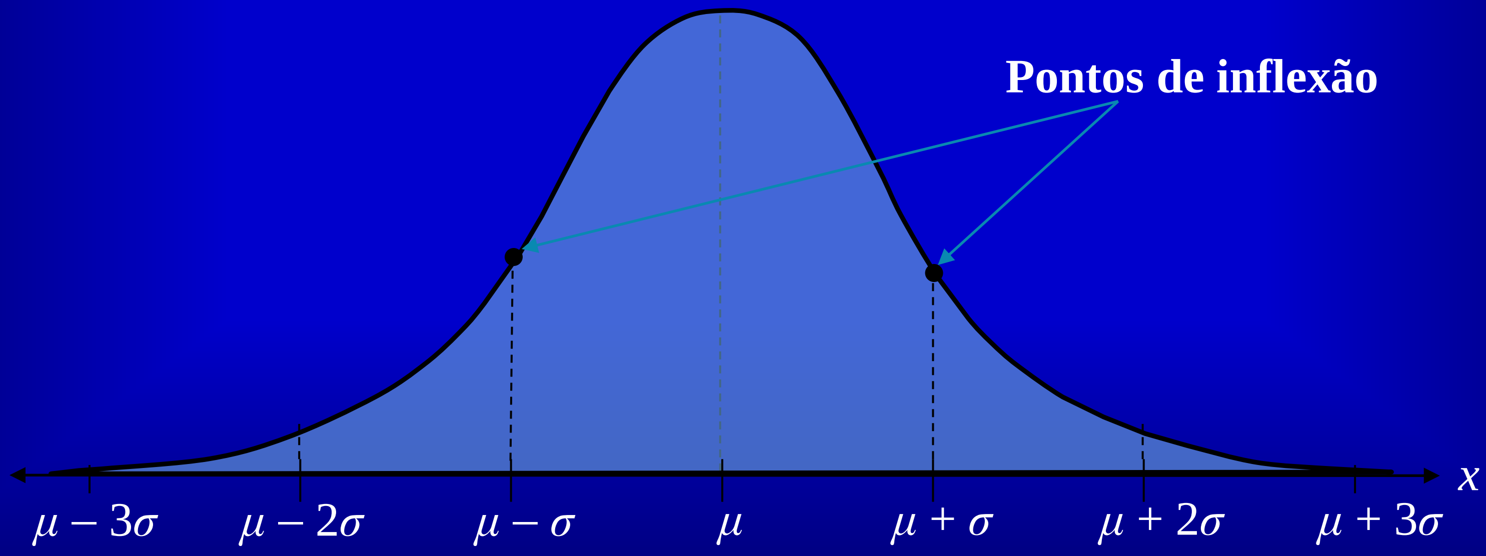
2. A curva normal tem formato de sino e é simétrica em relação à média.

3. A área total abaixo da curva é igual a 1.

4. A curva normal se aproxima do eixo *x*, mas nunca o toca, conforme se afasta da média.



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

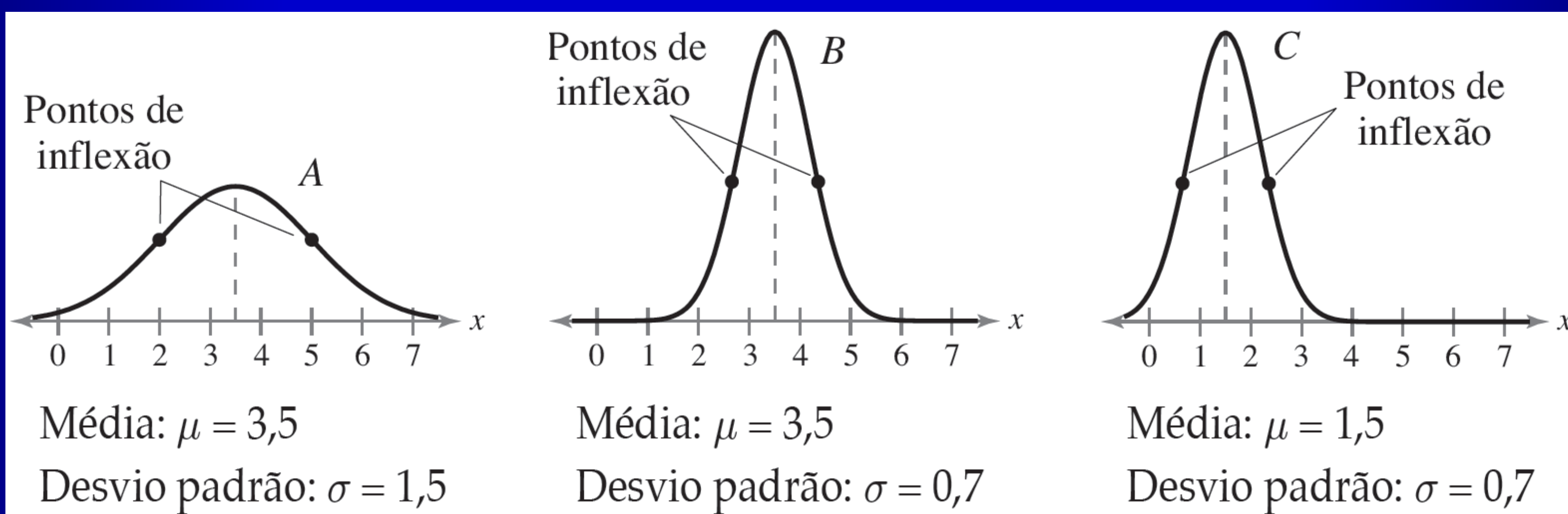


Médias e desvios padrão

Uma distribuição normal pode ter qualquer média e qualquer desvio padrão positivo

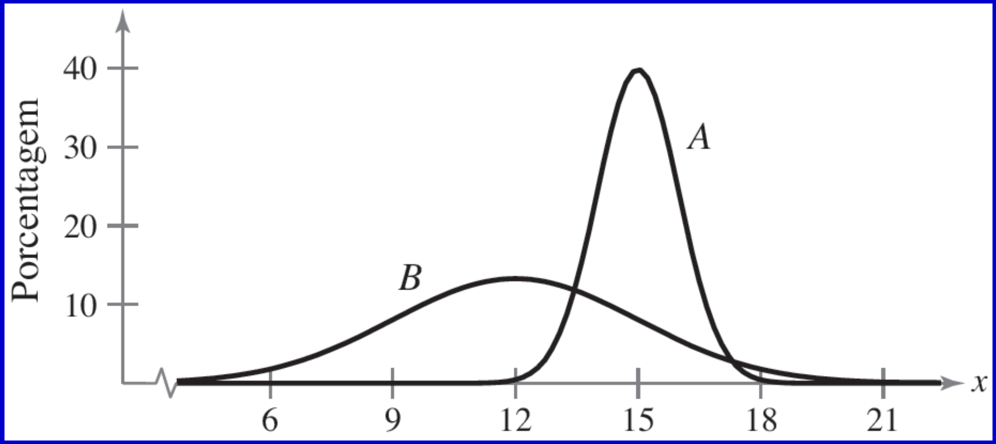
A média dá a localização da linha de simetria

O desvio padrão descreve a dispersão dos dados



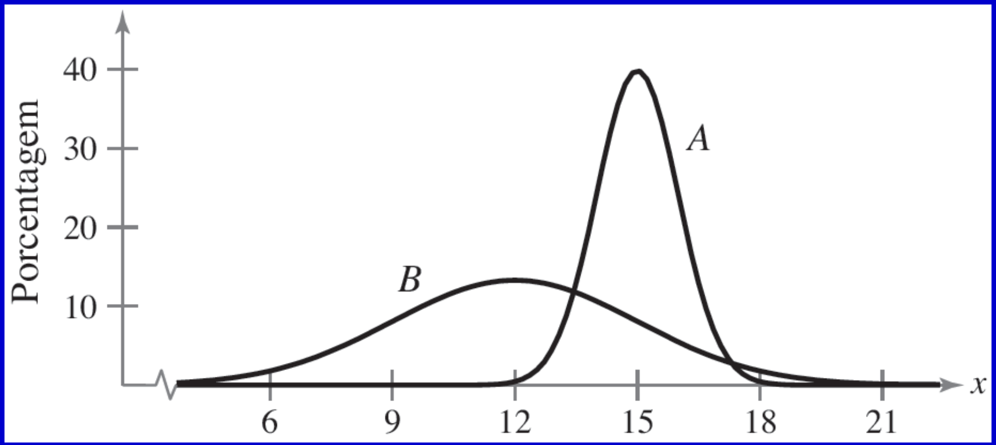
Exemplo: entendendo média e desvio padrão

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?



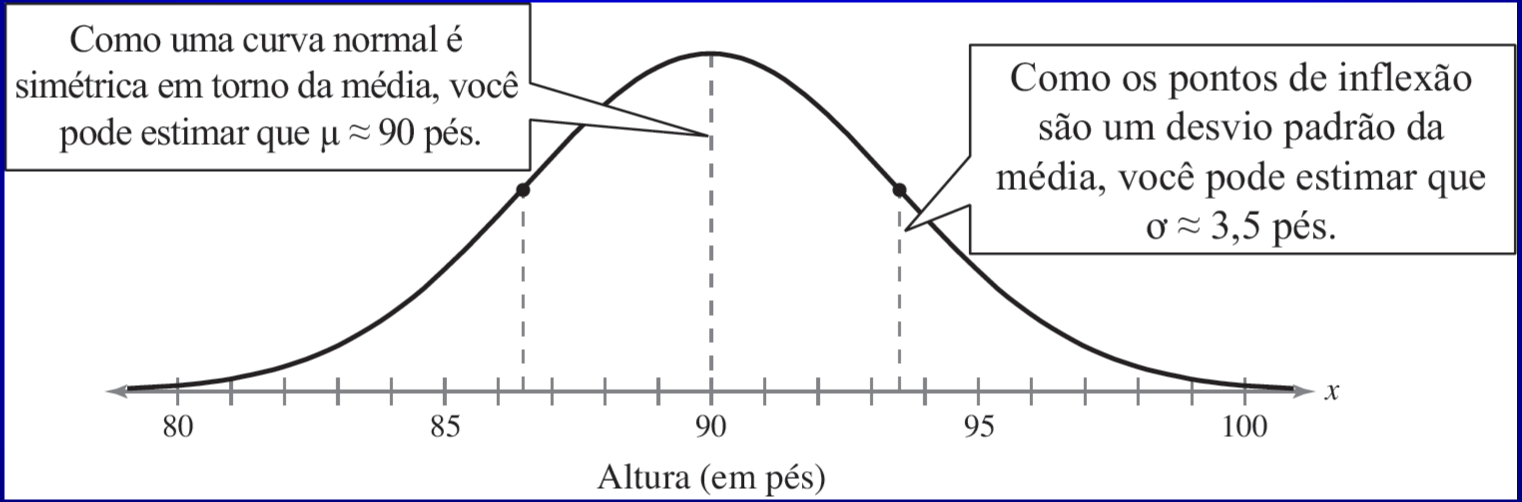
Solução:

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

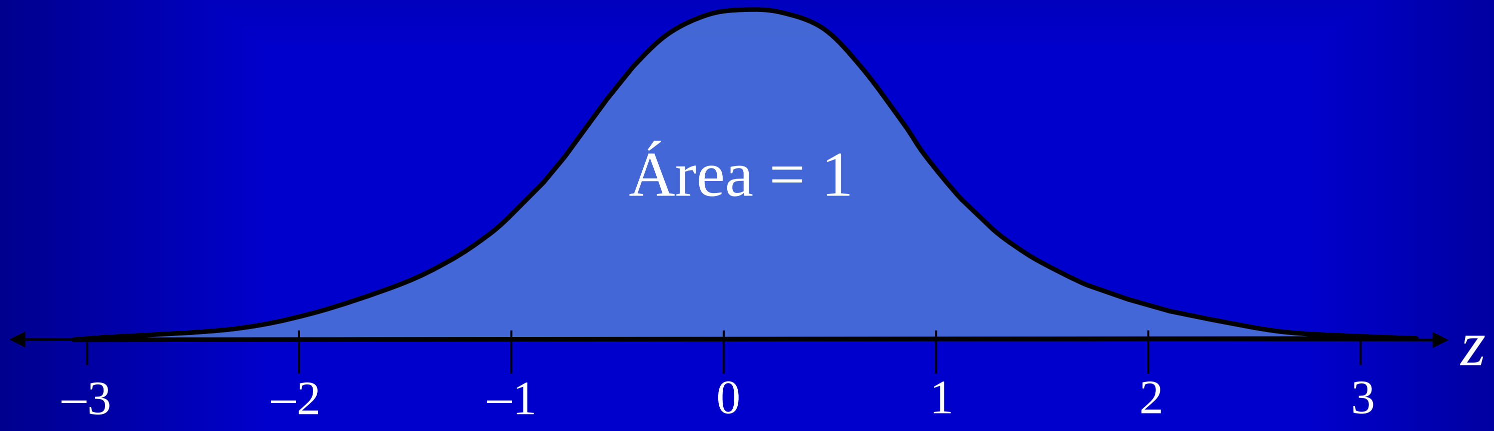
Exemplo: interpretando gráficos

As alturas 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 distribuição normal padrão



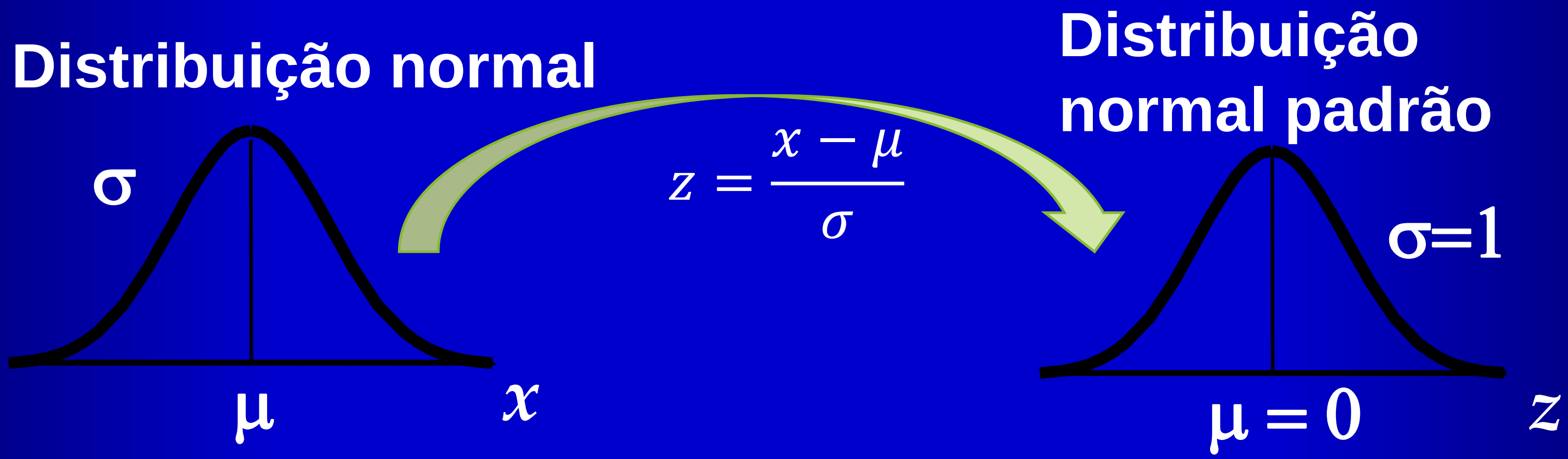
Qualque valor de *x* pode ser transformado em um escore *z* usando a formula:

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

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

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

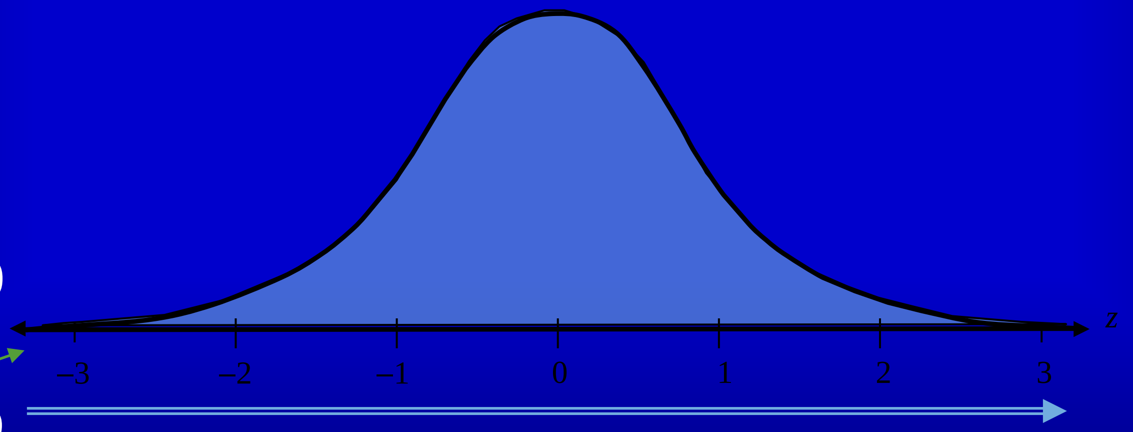


Usar a tabela normal padrão para encontrar a área cumulativa abaixo da curva normal padrão

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.

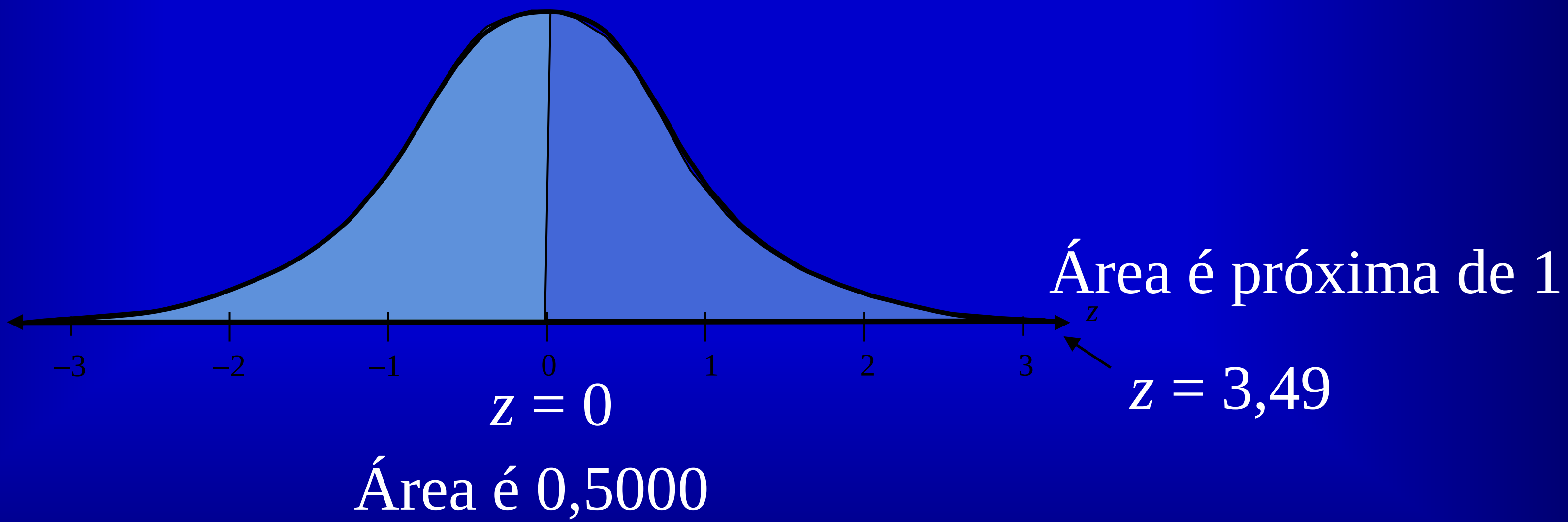
Área é próxima de 0



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
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

Solução:

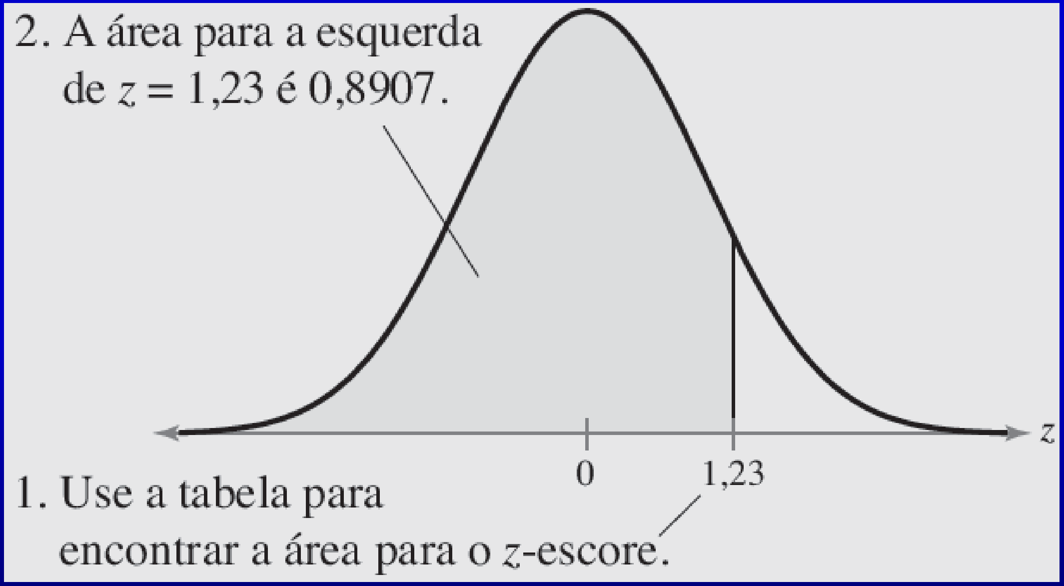
Encontre –0,2 na coluna à esquerda.

Cruze a fileira para a coluna sob 0.04.

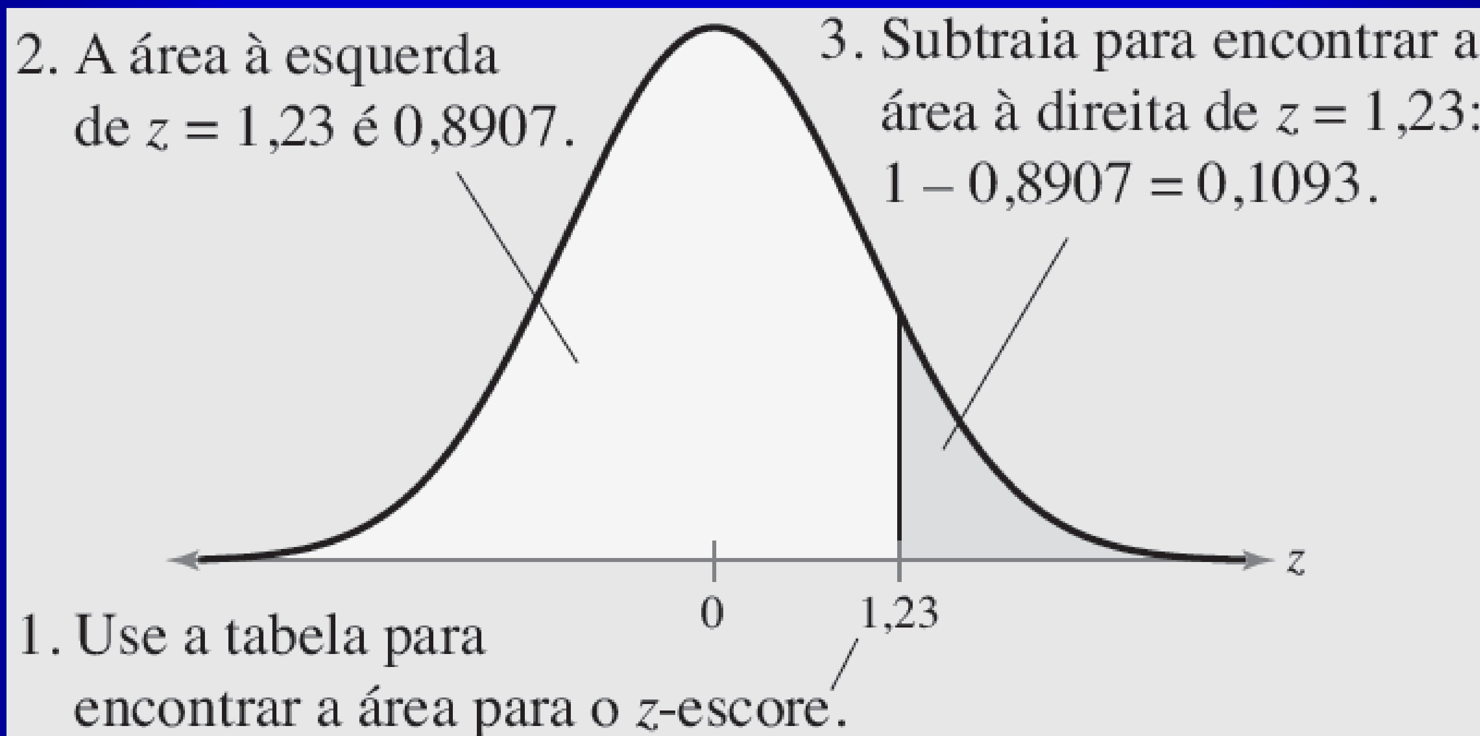
A área à esquerda de *z* = –0,24 é 0,4052.

Encontrando áreas sob a curva normal padrão

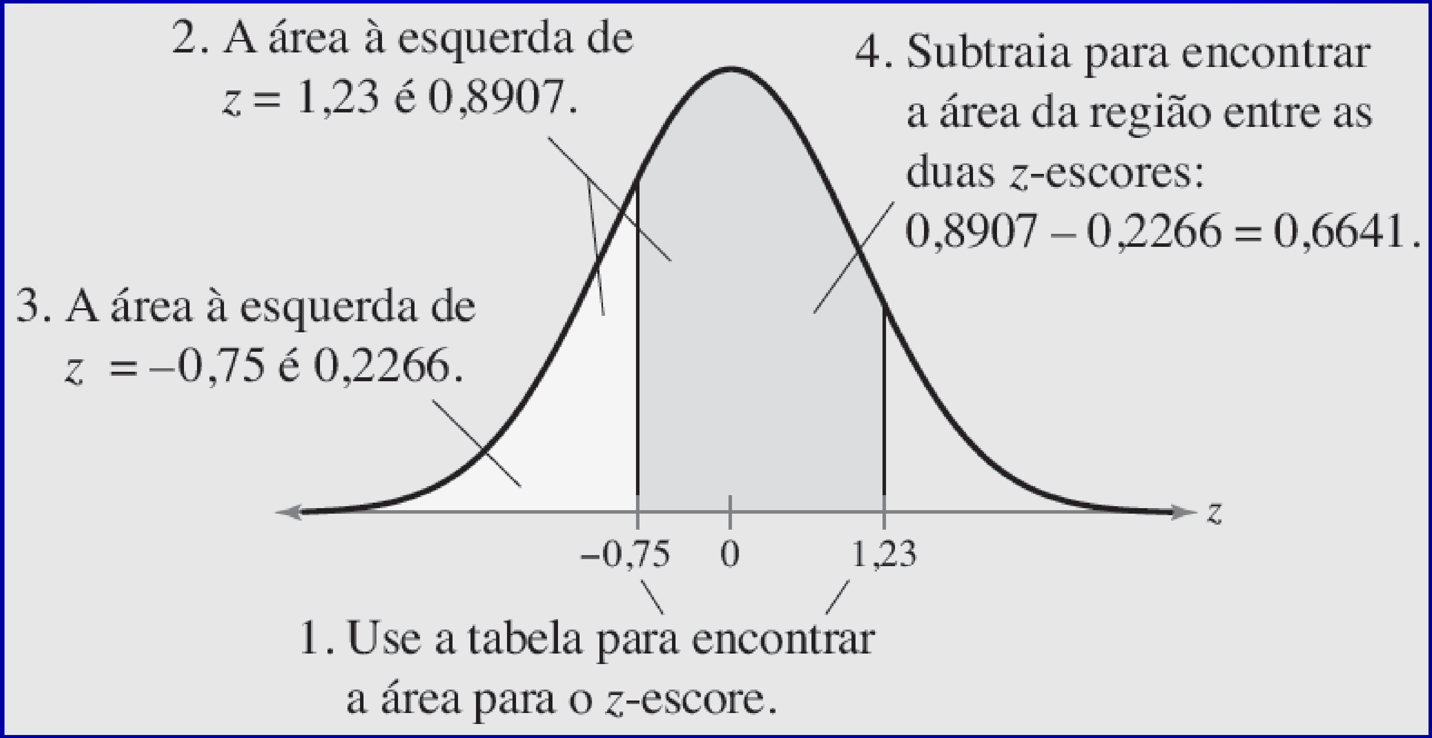
1. Esboce a curva normal padrão e preencha a área apropriada abaixo da curva.
2. Encontre a área seguindo as direções para cada caso.
    a) Para encontrar a área à *esquerda* de z, encontre a área que corresponda a z na tabela normal padrão.



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.



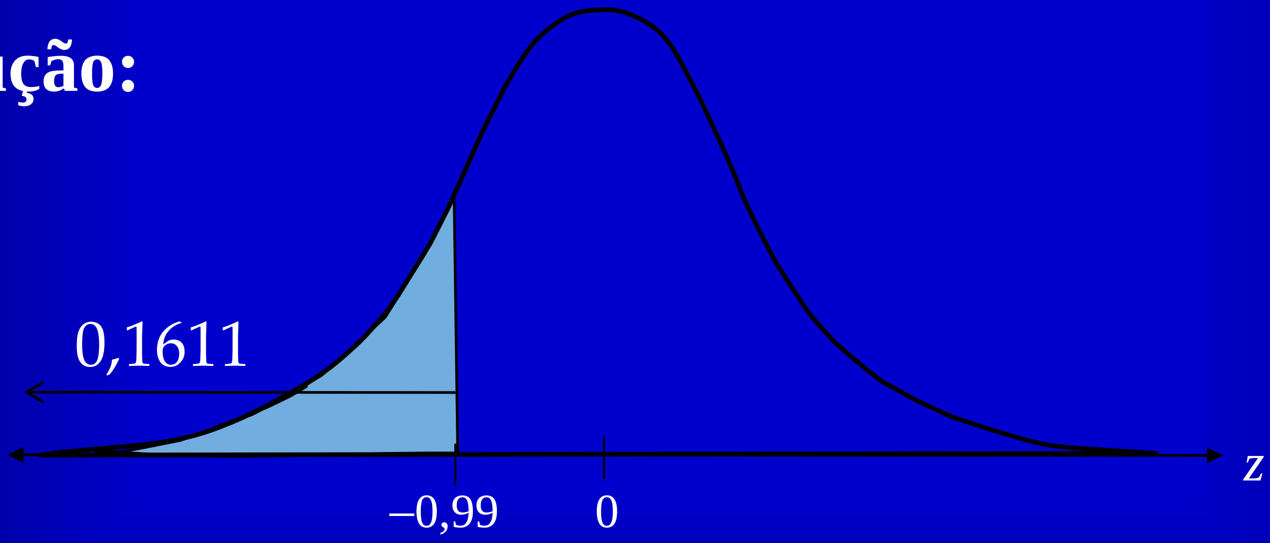
c) Para encontrar a área *entre* dois escores *z*, encontre a área correspondente a cada escore *z* na tabela normal padrão. Então subtraia a área menor da área maior.



Exemplo: encontrando a área sob a curva normal padrão

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

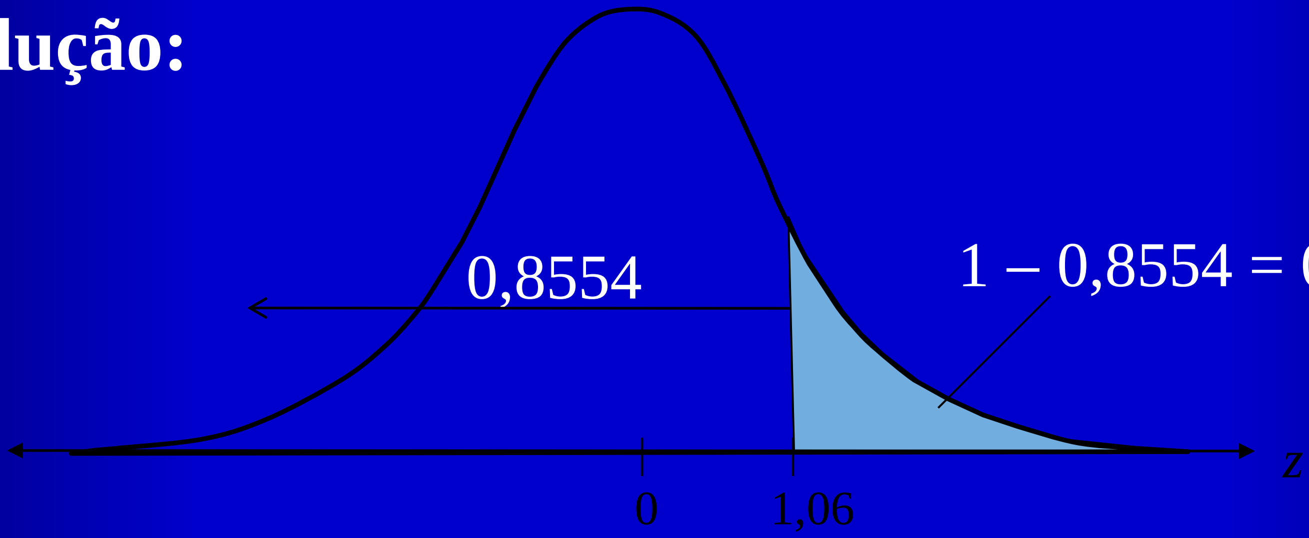
Solução:



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

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

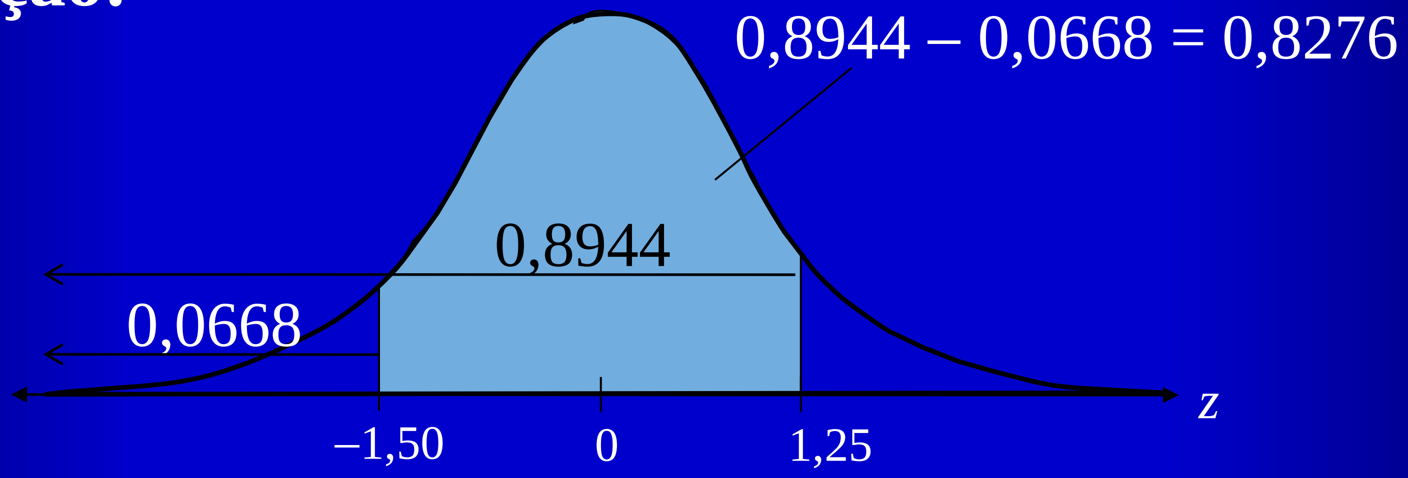
Solução:



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:



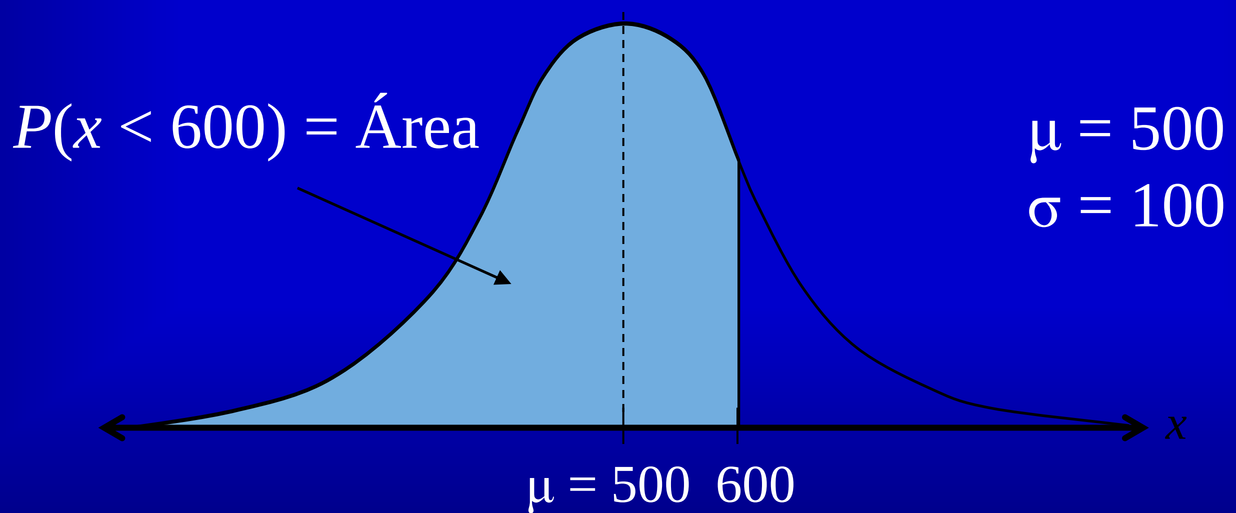
Pela tabela normal padrão, a área é igual a 0,8276.

Seção 5.2

Encontrando probabilidades para valores normalmente distribuídos

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.



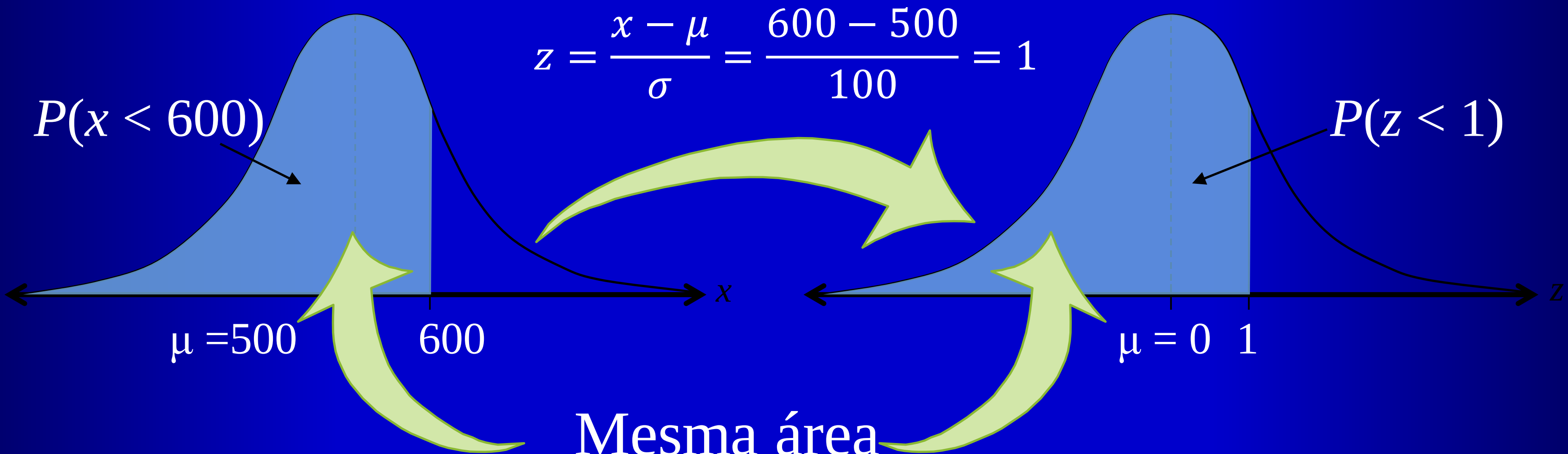
Distribuição normal

$\mu = 500 \;;\; \sigma = 100$

Distribuição normal padrão

$\mu = 0 \;;\; \sigma = 1$

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



$$P(x < 600) = P(z < 1) = 0{,}8413$$

Exemplo: encontrando probabilidades para distribuições normais

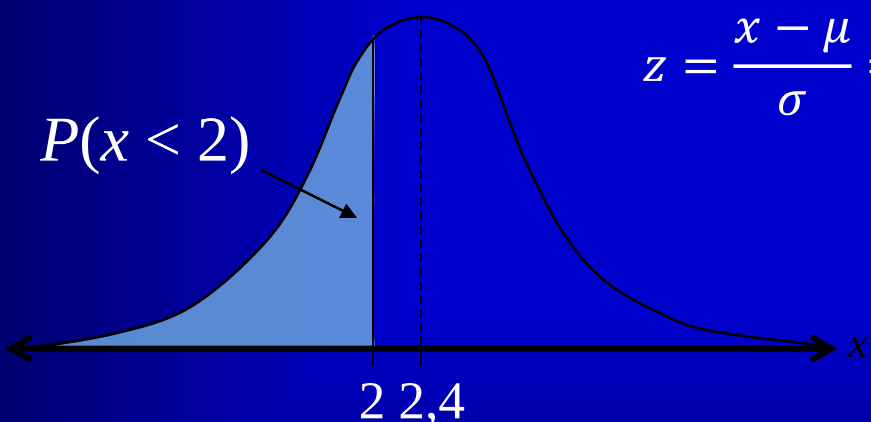
Uma pesquisa indica que pessoas usam seus computadores uma média de 2,4 anos antes de trocá-los por uma máquina nova. O desvio padrão é de 0,5 ano. Um proprietário de computador é selecionado aleatoriamente. Encontre a probabilidade de que ele use sua máquina por menos de dois anos antes de comprar uma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

$$P(X < 2)$$

Solução: encontrando probabilidades para distribuições normais

Distribuição normal

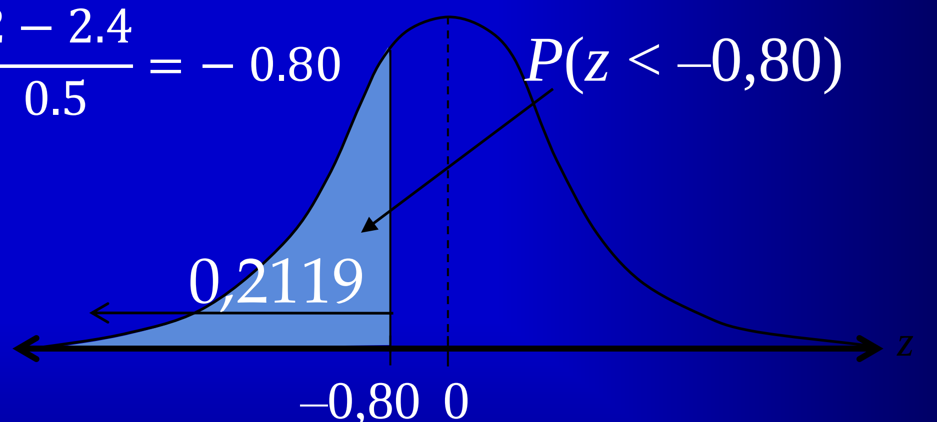
$\mu = 2{,}4 \quad \sigma = 0{,}5$



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

Distribuição normal padrão

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



$P(x < 2) = P(z < -0{,}80) = \mathbf{0{,}2119}$

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

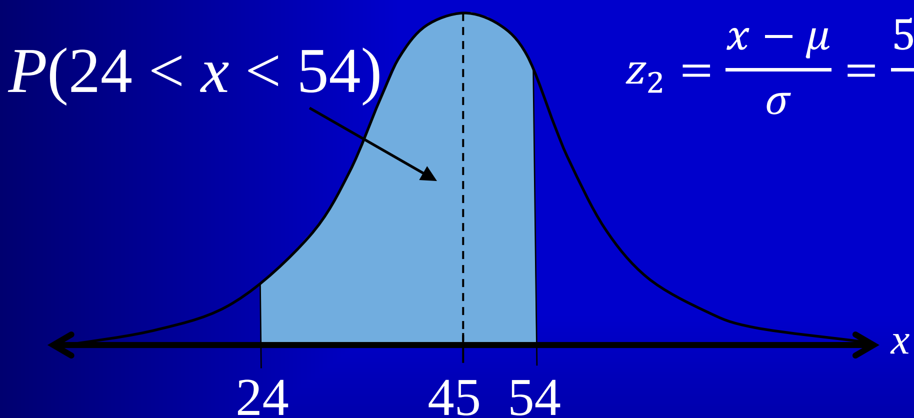
$\mu = 45 \quad \sigma = 12$

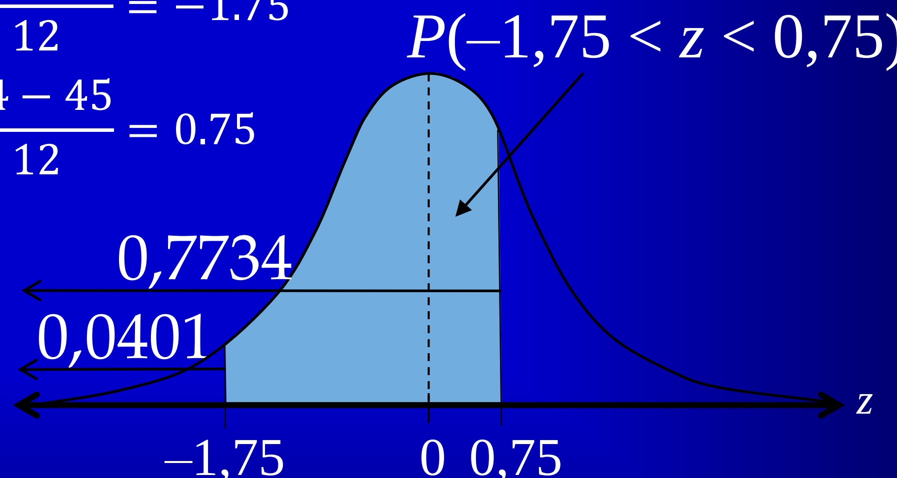
Distribuição normal padrão

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

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

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





$P(24 < x < 54) = P(-1{,}75 < z < 0{,}75)$

$= 0{,}7734 - 0{,}0401 =$ **$0{,}7333$**

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

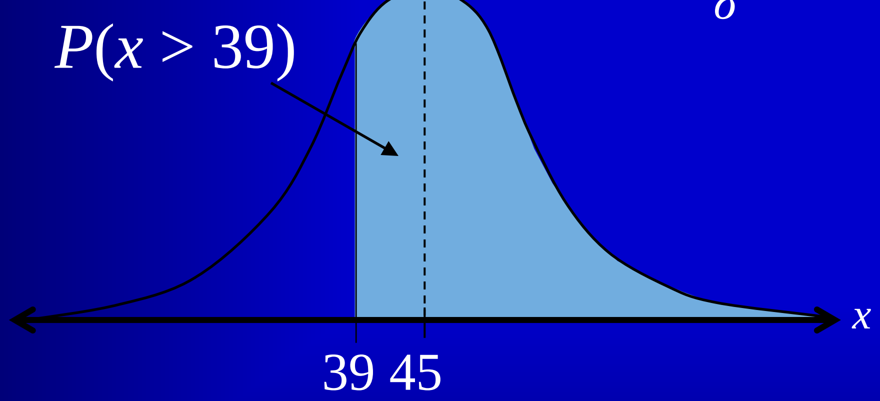
Distribuição normal

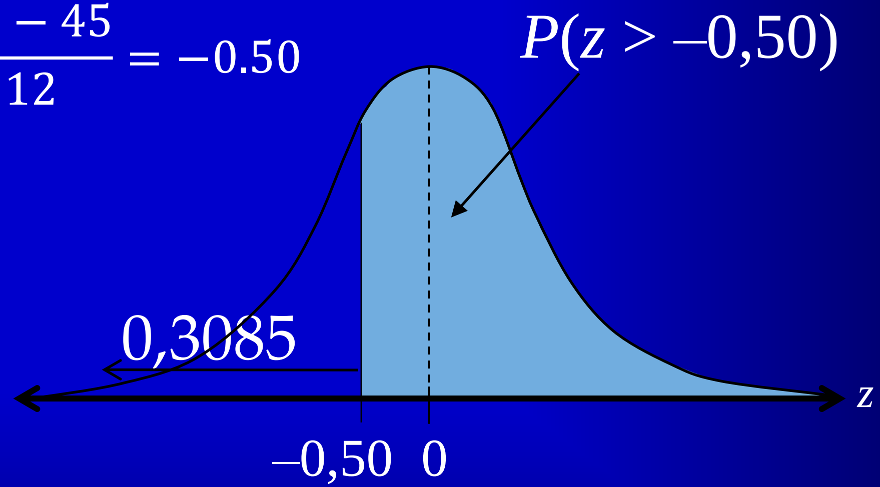
$\mu = 45 \quad \sigma = 12$

Distribuição normal padrão

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

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





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

Exemplo: encontrando probabilidades para distribuições normais

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

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

Exemplo: usando tecnologia para encontrar probabilidades normais

Assuma que os níveis de colesterol em homens do Brasil são normalmente distribuídos, com uma média de 215 miligramas por decilitro e um desvio padrão de 25 miligramas por decilitro. Você seleciona aleatoriamente um homem. Qual é a probabilidade de que seu nível de colesterol seja menor que 175? Use uma ferramenta tecnológica para encontrar a probabilidade.

$$P(X < 175) = P\left(Z < \frac{175 - 215}{25} = -1{,}6\right) = 0{,}0548$$

Solução: usando tecnologia para encontrar probabilidades normais

É preciso especificar a média, o desvio padrão, e o(s) valor(es) –*x*(s) que determinam o intervalo.

=DISTNORM(175;215;25;VERDADEIRO)

Objetivos da Seção 5.3

Encontrar um escore *z* dada a área sob a curva normal

Transformar um escore *z* em um valor *x*

Encontrar o valor de um dado específico de uma distribuição normal dada a probabilidade

Encontrando valores dada uma probabilidade

Na seção 5.2 foi dada uma variável aleatória *x* normalmente distribuída e foi pedido que encontrassem a probabilidade

Nesta seção, será dada uma probabilidade e será pedido o valor da variável aleatória *x*

5,2

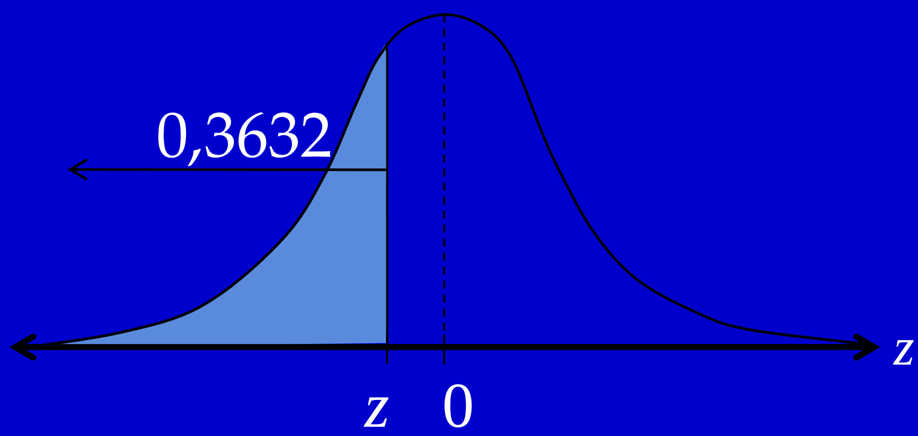
x *z* Probabilidade

5,3

Exemplo: encontrando um escore z dada uma área

Encontre o escore z que corresponda à área cumulativa de 0,3632.

Solução:



$$P(Z < z) = 0,3632 \Rightarrow z = -0,35$$

Solução: encontrando um escore *z* dada uma área

Localize 0,3632 no corpo da tabela normal padrão

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
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

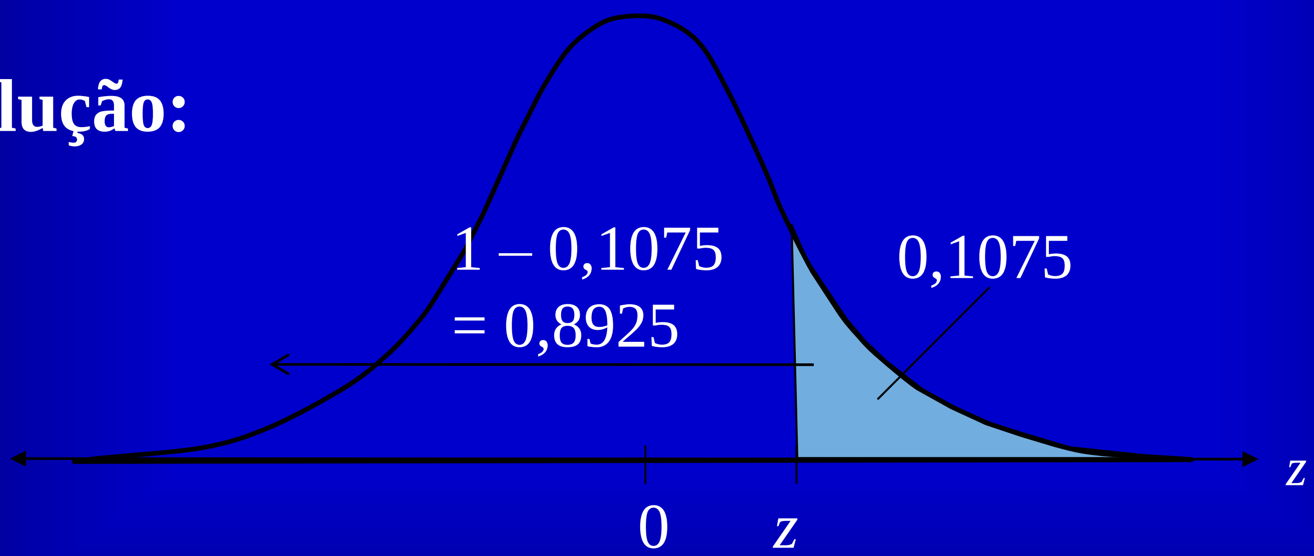
O escore *z* é – 0,35.

Os valores no começo da fileira correspondente e no topo da coluna fornecem o escore *z*

Exemplo: encontrando um escore *z* dada uma área

Encontre o escore *z* que tenha 10,75% da área da distribuição à sua direita.

Solução:



Porque a área à direita é 0,1075 e a área cumulativa é 0,8925.

Solução: encontrando um escore *z* dada uma área

Localize 0,8925 no corpo da tabela normal padrão

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

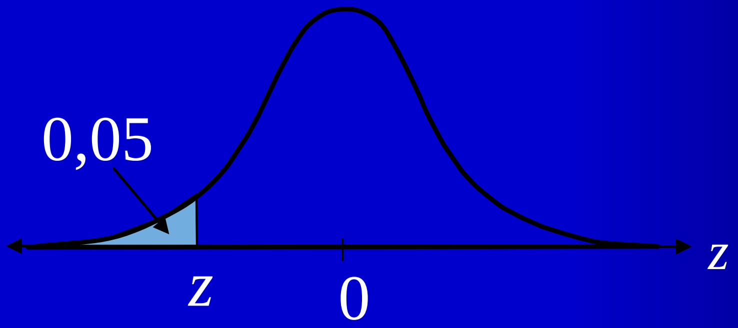
O escore *z* é 1,24.

Os valores no começo da fileira correspondente e no topo da coluna fornecem o escore *z*

Exemplo: encontrando um escore *z* dado um percentil

Encontre o escore *z* que corresponda a 0,05.

Solução:



As áreas mais próximas de 0,05 na tabela são 0,0495 (*z* = – 1,65) e 0,0505 (*z* = –1,64). Porque 0,05 está entre as duas áreas na tabela, use o escore *z* que está entre –1,64 e –1,65.

O escore *z* é –1,645.

Transformando um escore *z* em um escore *x*

Para transformar um escore *z* para um valor *x* em uma dada população, use a fórmula:

$$x = \mu + z\sigma$$

Exemplo: encontrando um valor x

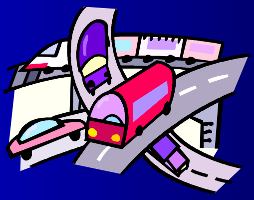
As velocidades dos veículos em um trecho de uma rodovia são normalmente distribuídas, com uma média de 67 km por hora e um desvio padrão de 4 km por horas. Encontre as velocidades *x* correspondentes aos escores *z* de 1,96 , –2,33 e 0.

Solução: Use a fórmula $x = \mu + z\sigma$

$z = 1{,}96$: $x = 67 + 1{,}96(4) = 74{,}84$ km por hora

$z = -2{,}33$: $x = 67 + (-2{,}33)(4) = 57{,}68$ km por hora

$z = 0$: $x = 67 + 0(4) = 67$ km por hora

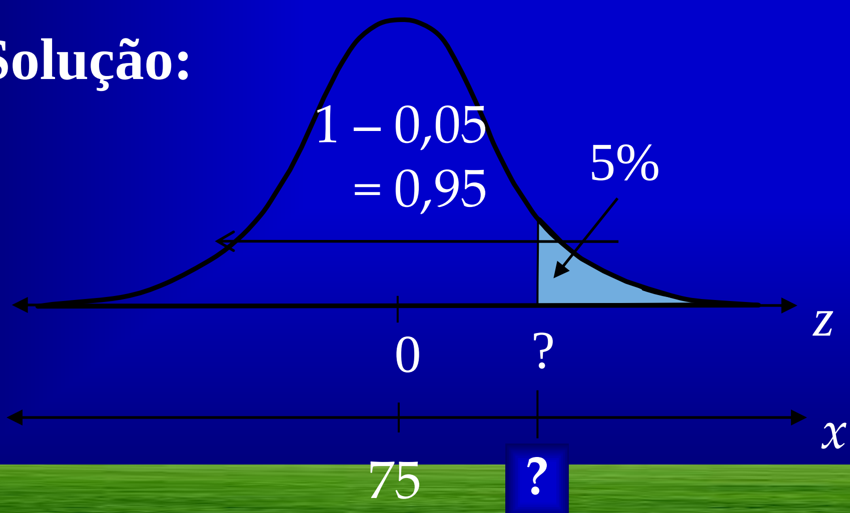


Note que 74,84 km/h está acima da média, e 67 km/h é igual à média.

Exemplo: encontrando um dado de valor específico

As pontuações para um teste de serviço civil são normalmente distribuídos, com uma média de 75 e um desvio padrão de 6,5. Para ser adequado ao emprego de serviço civil, você precisa ter uma pontuação dentro dos primeiros 5%. Qual é a menor pontuação que você pode conseguir e ainda assim ser adequado ao emprego?

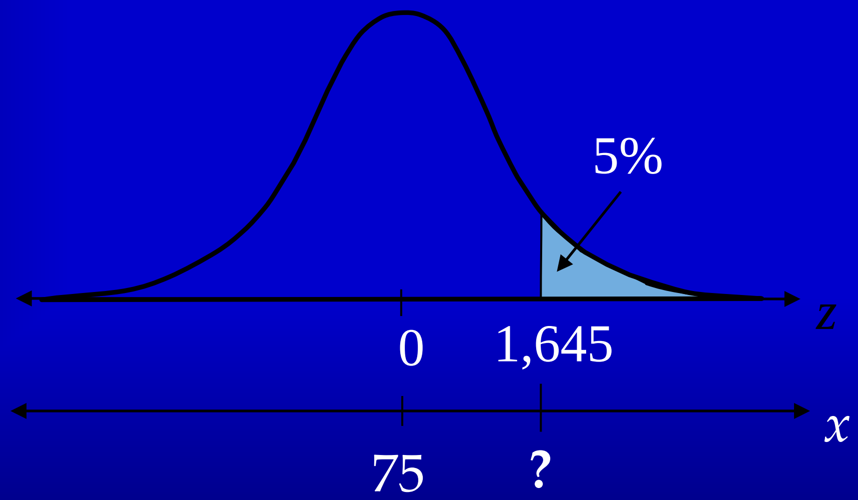
Solução:



Uma pontuação no teste acima dos primeiros 5% é qualquer pontuação acima do 95º percentil. Encontre o escore *z* que corresponda à área cumulativa de 0,95.

Solução: encontrando um dado de valor específico

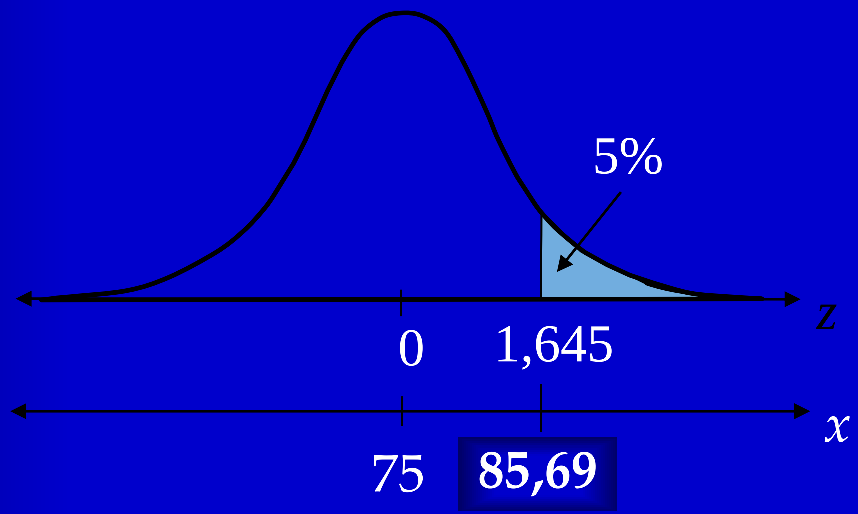
Pela tabela normal padrão, as áreas mais próximas de 0,95 são 0,9495 (z = 1,64) e 0,9505 (z = 1,65). Como 0,95 está entre as duas áreas na tabela, use o escore z que está entre 1,64 e 1,65, isto é, $\boldsymbol{z = 1{,}645}$.



Usando a equação: $\frac{x-\mu}{\sigma} = z \;\Rightarrow\; x = \mu + z\sigma$

Que é o equivalente a: $x = \mu + z\sigma$

$$x = 75 + 1{,}645(6{,}5) \approx 85{,}69$$



A pontuação mais baixa que você pode obter e ainda assim estar qualificado para o emprego é 86.