

Modelos probabilísticos para variáveis aleatórias contínuas

Uma variável é considerada contínua quando pode assumir qualquer valor dentro de um intervalo, ou seja, se houver um número incontável de resultados possíveis, representados por um intervalo sobre o eixo real.

Os principais modelos de probabilidades que são designados para descreverem o comportamento de uma variável aleatória contínua são:

- a) Uniforme
- b) Normal
- c) Exponencial
- d) Gama
- e) Weibull
- f) Beta
- g) Lognormal

Neste material abordaremos as distribuições Exponencial e Normal.

Distribuição Exponencial

A distribuição exponencial é o modelo probabilístico usual para situações tais como tempo de espera em uma fila, tempo de sobrevivência de um paciente após o início de um tratamento, tempo de vida de material eletrônico.

Definição: Uma v. a. X tem distribuição exponencial se sua função de densidade de probabilidade (f.d.p.) é da forma:

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & , x > 0 \text{ e } \alpha > 0 \\ 0 & , \text{ para quaisquer outros valores} \end{cases}$$

A notação que utilizaremos para representar que uma v. a. contínua X tem distribuição exponencial é: $X \sim \text{Exp}(\alpha)$.

A média e a variância da distribuição exponencial são:

$$E(X) = \frac{1}{\alpha}$$

$$\text{Var}(X) = \frac{1}{\alpha^2}$$

Apresentamos na Figura 1 o gráfico da distribuição exponencial para três valores diferentes para o parâmetro (α).

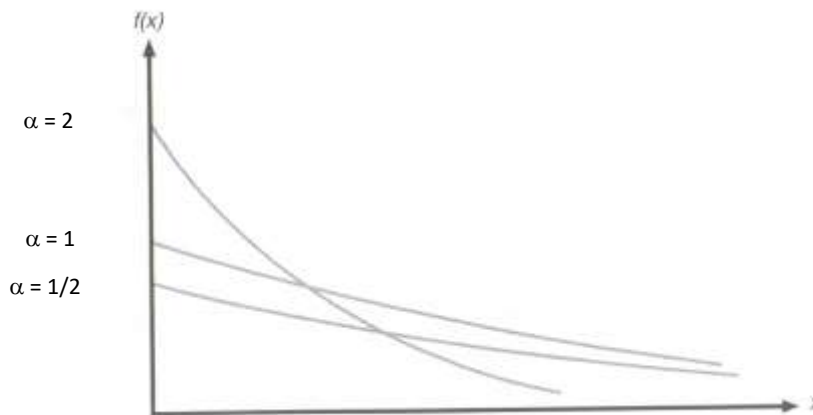


Figura 1 – Distribuição exponencial para $\alpha = 2$, $\alpha = 1$ e $\alpha = 0,5$.

Para uma variável aleatória contínua X distribuída exponencialmente, para qualquer intervalo de a até b , a probabilidade $P(a \leq X \leq b)$ é calculada por:

$$P(a \leq X \leq b) = \int_a^b \alpha e^{-\alpha x} dx = -e^{-\alpha x} \Big|_a^b = e^{-\alpha a} - e^{-\alpha b}$$

Lembre-se de que a inclusão ou não dos extremos a e b não altera o cálculo efetuado acima.

Propriedade de falta de memória: Considere uma variável aleatória contínua X com distribuição exponencial e parâmetro α . Considere também s e $t > 0$, então temos que:

$$\begin{aligned} P(X \geq t + s / X \geq s) &= \frac{P(X \geq t + s \cap X \geq s)}{P(X \geq s)} = \frac{P(X \geq t + s)}{P(X \geq s)} = \\ &= \frac{\int_{t+s}^{\infty} \alpha e^{-\alpha x} dx}{\int_s^{\infty} \alpha e^{-\alpha x} dx} = \frac{-e^{-\alpha x} \Big|_{t+s}^{\infty}}{-e^{-\alpha x} \Big|_s^{\infty}} = \frac{e^{-\alpha(t+s)}}{e^{-\alpha(s)}} = e^{-\alpha t} = P(X \geq t) \end{aligned}$$

Supondo que X representa o tempo de vida de um equipamento, podemos fazer a seguinte interpretação para a propriedade da falta de memória: a probabilidade do equipamento durar pelo menos $t+s$ anos, sabendo-se que já está funcionando a pelo menos s , é igual a probabilidade de um equipamento novo durar pelo menos t anos. Em outras palavras, a informação da “idade” do equipamento pode ser esquecida e o que importa, para o cálculo da probabilidade, é quantos anos a mais queremos que dure.

Exemplo 1

A vida útil de uma lâmpada é modelada através da distribuição exponencial com parâmetro $1/8000$.

a) Calcule o tempo médio de duração dessas lâmpadas.

X : vida útil de uma lâmpada

$X \sim \text{Exp}(1/8000)$.

$$E(X) = \frac{1}{\alpha} = \frac{1}{1/8000} = 8000 \text{ horas}$$

b) Calcule a probabilidade de que uma lâmpada dure pelo menos 4000 horas.

Como a distribuição exponencial não tem um limite superior mas, tem o zero como limite inferior, para realizar o cálculo de que a lâmpada dure pelo menos 4000 horas iremos utilizar o complementar, ou seja,

$$P(X \geq 4000) = 1 - P(0 < X < 4000) = 1 - \left(e^{-\frac{1}{8000} \cdot 0} - e^{-\frac{1}{8000} \cdot 4000} \right) = 0,6065$$

- c) Sabe-se que o fabricante garante a reposição de uma lâmpada caso ela dure menos de 50 horas. Determine a probabilidade de haver troca por defeito na fabricação.

$$P(X < 50) = P(0 < X < 50) = e^{-\frac{1}{8000} \cdot 0} - e^{-\frac{1}{8000} \cdot 50} = 0,0062$$

- d) Uma lâmpada é colocada em teste. Calcule a probabilidade de que ela dure pelo menos 10000 horas, sabendo-se que ela já está em funcionamento a pelo menos 6000 horas.

Para resolver essa probabilidade condicional podemos utilizar a propriedade de falta de memória da distribuição exponencial:

$$P(X \geq 10000 | X \geq 6000) = P(X \geq 4000) = 0,6065$$

Exercício 1

A vida de certa marca de lâmpada tem uma distribuição aproximadamente exponencial com média de 1000 horas.

- Determinar a porcentagem das lâmpadas que queimarão antes de 1000 horas.
- Após quantas horas terão queimado 50% das lâmpadas?

Respostas: a) 0,6312 b) 693,1472 horas

Exercício 2

Uma fábrica utiliza dois métodos para a produção de lâmpadas. 70% das lâmpadas são produzidas pelo método A e as demais pelo método B. A duração da lâmpada depende do método pelo qual ela foi produzida, sendo que as produzidas pelo método A seguem uma distribuição exponencial com parâmetro $1/80$ e as do método B seguem uma exponencial de parâmetro $1/100$. Qual a probabilidade de que, se escolhermos uma lâmpada ao acaso, ela dure mais de 100 horas?

Resposta: 0,31

Distribuição Normal

A distribuição normal conhecida também como distribuição gaussiana é sem dúvida a mais importante distribuição contínua.

Sua importância se deve a vários fatores, entre eles podemos citar o Teorema Central do Limite, o qual é um resultado fundamental em aplicações práticas e teóricas, pois ele garante que mesmo que os dados não sejam distribuídos segundo uma normal a média dos dados converge para uma distribuição normal conforme o número de dados aumenta.

O requisito mínimo para adotarmos o modelo normal para um conjunto de dados contínuos é que seu histograma seja aproximadamente simétrico em torno do ponto central, que também deve ser o ponto de máximo.

Definição: A variável aleatória X , que tome todos os valores reais $-\infty < x < +\infty$, tem uma distribuição normal (ou gaussiana) se sua função densidade de probabilidade (f.d.p.) for da forma:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$$

Os parâmetros da distribuição normal são a média e a variância, onde $-\infty < \mu < +\infty$ e $\sigma > 0$. Utilizaremos a notação $X \sim N(\mu, \sigma^2)$ para representar que a v. a. X tem distribuição Normal com parâmetros μ e σ^2 .

Quando μ e σ^2 são desconhecidos (caso mais comum), estes valores serão estimados por \bar{x} e s^2 , respectivamente, a partir da amostra.

A Figura 2 mostra a função densidade de probabilidade de uma v. a. com distribuição normal:

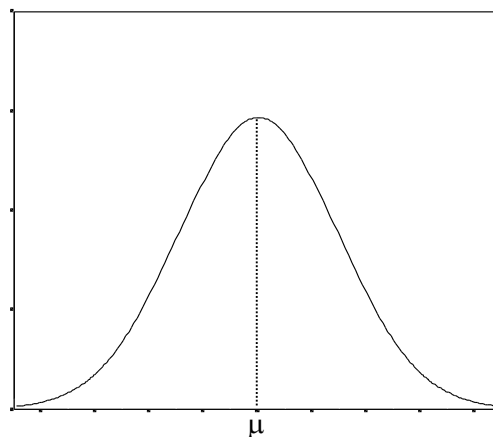


Figura 2 – Função densidade de probabilidade da distribuição normal.

Propriedades:

- ✓ A curva que representa a distribuição de probabilidade é frequentemente descrita como curva em forma de sino ou curva de Gauss;

- ✓ A distribuição é simétrica em torno da média, assim, as medidas de tendência central (média, mediana e moda) apresentam o mesmo valor.
- ✓ A distribuição normal fica delimitada pelo seu desvio padrão e sua média. Para cada combinação de valores de média e desvio padrão gera uma distribuição Normal diferente;
- ✓ A área sob a curva corresponde à 1.
- ✓ O ponto de máximo da curva da distribuição normal ocorre quando $x = \mu$, ou seja, em torno da média registra-se uma probabilidade maior de ocorrência. À medida que nos afastamos da média, as probabilidades de ocorrência vão diminuindo.
- ✓ Se X tiver a distribuição normal com média = 0 e variância = 1, ou seja, $X \sim N(0, 1)$, diremos que X possui a distribuição normal reduzida ou **distribuição normal padrão**. Isto é, sua função densidade de probabilidade pode ser escrita como

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

A integral da função densidade de probabilidade da distribuição normal, não pode ser avaliada pelo método tradicional (teorema fundamental do cálculo). Ela só pode ser calculada por métodos numéricos. E por isso ela é encontrada tabelada em qualquer livro texto de Probabilidade ou Estatística.

A distribuição Normal depende dos parâmetros: média (μ) e desvio padrão (σ). A depender dos valores da média e do desvio, diferentes serão os formatos das curvas. A Figura 3 apresenta três curvas distintas, com mesmo desvio padrão, porém com médias diferentes:

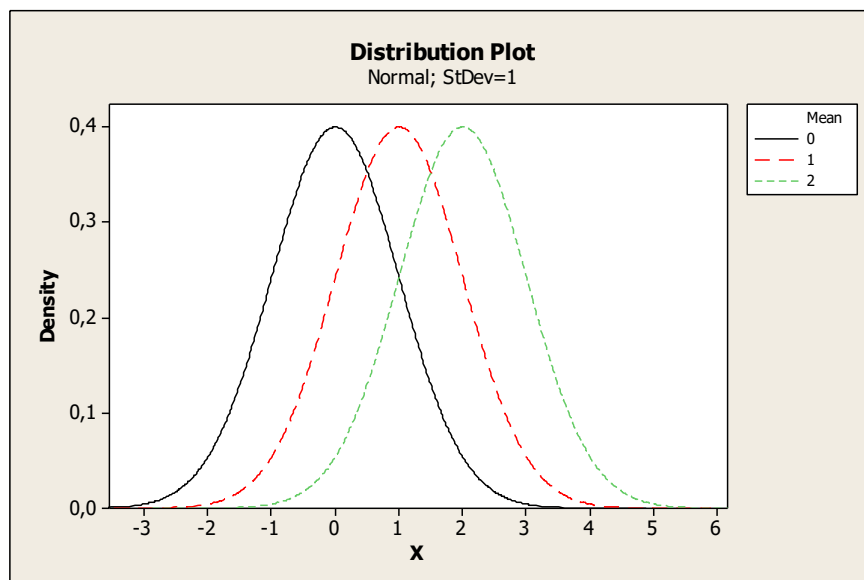


Figura 3 – Distribuição Normal com mesmo desvio padrão ($\sigma=1$) e médias diferentes ($\mu=0$, $\mu=1$ e $\mu=2$).

A Figura 4 apresenta três curvas distintas, com a mesma média, porém com desvios diferentes.

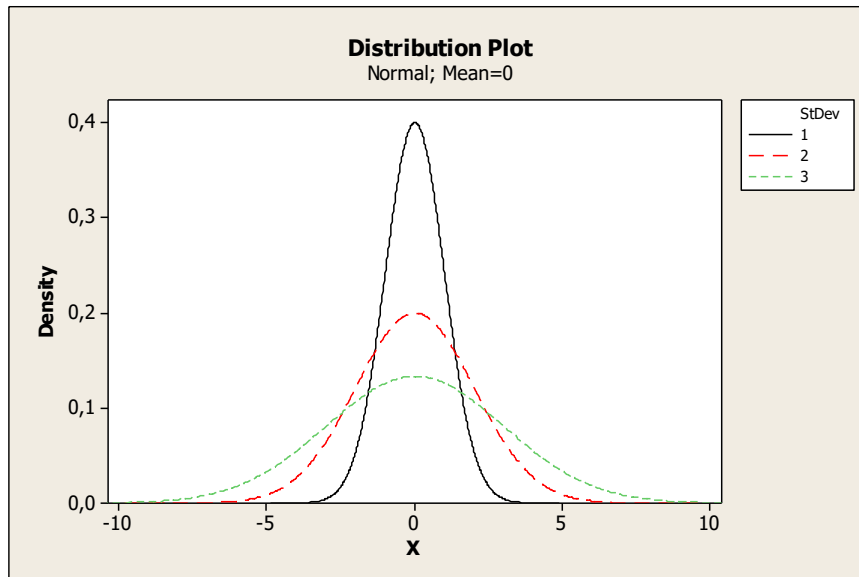


Figura 4 – Distribuições com mesma média ($\mu=0$) e desvios diferentes ($\sigma=1$, $\sigma=2$ e $\sigma=3$).

Definição: Regra de 2σ . No caso de uma distribuição normal $N(\mu; \sigma^2)$, existe uma regra simples e muito útil. Qualquer que sejam os valores dos parâmetros, a área (e portanto, a probabilidade) que fica entre $\mu - \sigma$ e $\mu + \sigma$ é aproximadamente igual a 0,68. A área que fica localizada a dois σ de μ (isto é, a área entre $\mu - 2\sigma$ e $\mu + 2\sigma$) é aproximadamente igual a 0,95. A uma distância de 3σ de μ fica uma área (e probabilidade) de 0,997, aproximadamente. Assim, é de 5% a chance de um valor selecionado de uma $N(\mu; \sigma^2)$ se afastar por mais de 2σ de μ . A chance de se afastar mais de 3σ é bastante pequena.

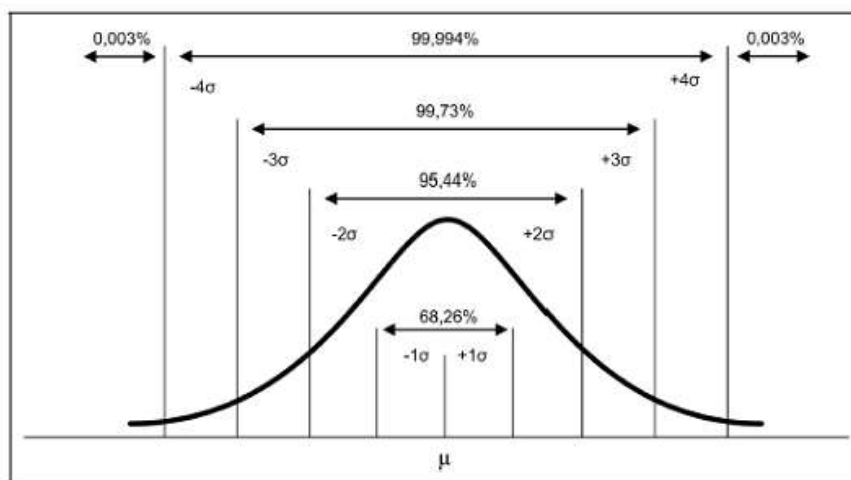


Figura 5 – Ilustração da regra de 2σ para a distribuição normal.

Fonte: <http://www.portalaction.com.br/probabilidades/62-distribuicao-normal>

Distribuição Normal Padrão

O cálculo de probabilidades, para variáveis adequadamente descritas pela distribuição normal, é realizado por meio da distribuição normal padrão, também chamada de

padronizada ou reduzida. A variável aleatória Z tem distribuição normal padrão, pois sua média é igual a zero e sua variância é igual a 1, ou seja, $Z \sim N(0, 1)$.

Para obter tal distribuição, isto é, quando se tem uma variável X com distribuição normal com média μ diferente de 0 (zero) e/ou desvio padrão σ diferente de 1 (um), devemos reduzi-la a uma variável Z , efetuando o seguinte cálculo:

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

Onde, z é o valor da variável normal padronizada Z , x é o valor da variável normal X ; μ é a média da variável aleatória X e σ é o desvio padrão da variável aleatória X .

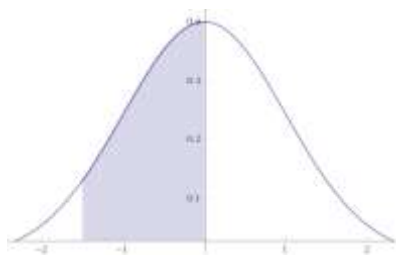
Existem vários tipos de tabelas que nos fornecem as probabilidades sob a curva normal padrão. A tabela que vamos utilizar é aquela que fornece a probabilidade da variável Z assumir um valor entre zero e um particular valor z .

Exemplo 2

Dado que Z é uma variável aleatória normal padrão, ou seja, $\mu=0$ e $\sigma=1$, calcule as probabilidades à seguir:

- a. $P(-1,5 < Z < 0)$

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



Buscando a probabilidade na tabela da distribuição normal padrão, temos:

↓

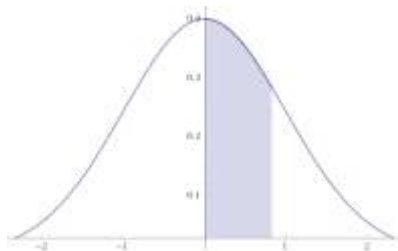
z	0	1	2	3	4
0,0	0,0000	0,0040	0,0080	0,0120	0,0160
0,1	0,0398	0,0438	0,0478	0,0517	0,0557
0,2	0,0793	0,0832	0,0871	0,0910	0,0948
.
.
.
1,4	0,4192	0,4207	0,4222	0,4236	0,4251
1,5	0,4332	0,4345	0,4357	0,4370	0,4382
1,6	0,4452	0,4463	0,4474	0,4484	0,4495

→

$P(-1,5 < Z < 0) = P(0 < Z < 1,5) = 0,4332$ (Já que a curva da distribuição normal é simétrica em torno da média).

- b. $P(0 < Z < 0,83)$

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



Buscando a probabilidade na tabela da distribuição normal padrão, temos:

↓

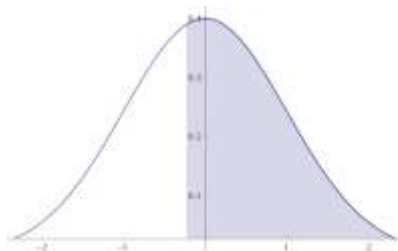
z	0	1	2	3
0,0	0,0000	0,0040	0,0080	0,0120
0,1	0,0398	0,0438	0,0478	0,0517
0,2	0,0793	0,0832	0,0871	0,0910
0,3	0,1179	0,1217	0,1255	0,1293
0,4	0,1554	0,1591	0,1628	0,1664
0,5	0,1915	0,1950	0,1985	0,2019
0,6	0,2257	0,2291	0,2324	0,2357
0,7	0,2580	0,2611	0,2642	0,2673
0,8	0,2881	0,2910	0,2939	0,2967

→

$$P(0 < Z < 0,83) = 0,2967$$

c. $P(Z > -0,23) = 0,59095$

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



Buscando a probabilidade na tabela da distribuição normal padrão, temos:

↓

z	0	1	2	3
0,0	0,0000	0,0040	0,0080	0,0120
0,1	0,0398	0,0438	0,0478	0,0517
0,2	0,0793	0,0832	0,0871	0,0910

→

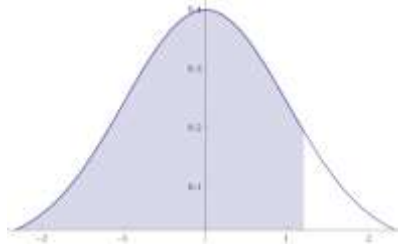
$$P(-0,23 < Z < 0) = 0,0910$$

$$\text{Para encontrar } P(Z > -0,23) = P(-0,23 < Z < 0) + 0,5 = 0,5910$$

Foi necessário somar 0,5 à probabilidade porque $P(Z > 0) = 0,5$.

d. $P(Z < 1,20)$

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



Buscando a probabilidade na tabela da distribuição normal padrão, temos:

↓

z	0	1	2
0,0	0,0000	0,0040	0,0080
0,1	0,0398	0,0438	0,0478
0,2	0,0793	0,0832	0,0871
0,3	0,1179	0,1217	0,1255
.	.	.	.
.	.	.	.
.	.	.	.
1,0	0,3413	0,3438	0,3461
1,1	0,3643	0,3665	0,3686
1,2	0,3849	0,3869	0,3888
1,3	0,4032	0,4049	0,4066

→

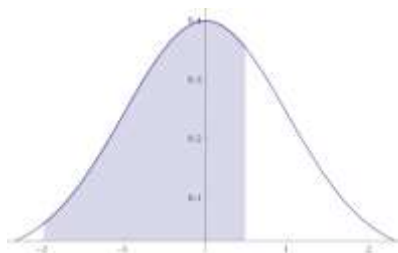
$$P(0 < Z < 1,20) = 0,3849$$

Para encontrar $P(Z < 1,20) = P(0 < Z < 1,20) + 0,5 = 0,8849$.

Foi necessário somar 0,5 à probabilidade porque $P(Z < 0) = 0,5$.

e. $P(-1,98 < Z < 0,49)$

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



Buscando as probabilidades na tabela da distribuição normal padrão, temos:



z	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
.
.
.
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767

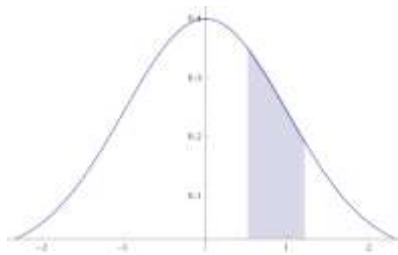
$$P(0 < Z < 0,49) = 0,1879$$

$$P(-1,98 < Z < 0) = 0,4761$$

$$\text{Para encontrar } P(-1,98 < Z < 0,49) = P(-1,98 < Z < 0) + P(0 < Z < 0,49) = 0,6640.$$

f. $P(0,52 < Z < 1,22)$

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



Buscando as probabilidades na tabela da distribuição normal padrão, temos:

z	0	1	2	3
0,0	0,0000	0,0040	0,0080	0,0120
0,1	0,0398	0,0438	0,0478	0,0517
0,2	0,0793	0,0832	0,0871	0,0910
0,3	0,1179	0,1217	0,1255	0,1293
0,4	0,1554	0,1591	0,1628	0,1664
0,5	0,1915	0,1950	0,1985	0,2019
0,6	0,2257	0,2291	0,2324	0,2357
.
.
.
1,2	0,3849	0,3869	0,3888	0,3907
1,3	0,4032	0,4049	0,4066	0,4082

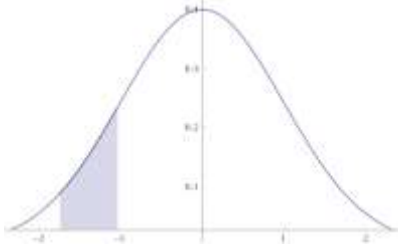
$$P(0 < Z < 0,52) = 0,1985$$

$$P(0 < Z < 1,22) = 0,3888$$

$$\text{Para encontrar } P(0,52 < Z < 1,22) = P(0 < Z < 1,22) - P(0 < Z < 0,52) = 0,1903.$$

g. $P(-1,75 < Z < -1,04)$

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



Buscando as probabilidades na tabela da distribuição normal padrão, temos:

z	0	1	2	3	4	5	6
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026
.
.
.
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770
.
.
.
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608

$$P(-1,75 < Z < 0) = 0,4599$$

$$P(-1,04 < Z < 0) = 0,3508$$

$$\text{Para encontrar } P(-1,75 < Z < -1,04) = P(-1,75 < Z < 0) - P(-1,04 < Z < 0) = 0,1091.$$

Exemplo 3

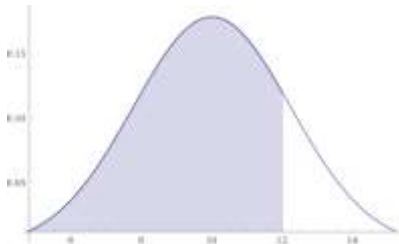
Suponha que as medidas da corrente em um pedaço de fio sigam a distribuição normal, com um média de 10 miliamperes e uma variância de 5 miliamperes. Qual a probabilidade:

Como temos uma variável (X : medida da corrente em um pedaço de fio) com distribuição normal com $\mu=10$ e $\sigma^2=5$, é necessário padronizá-la para poder consultar as probabilidades disponíveis na tabela da distribuição normal padrão. A padronização de uma variável $X \sim N(\mu, \sigma^2)$ em uma variável $Z \sim N(0, 1)$ é realizada efetuando o seguinte cálculo:

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

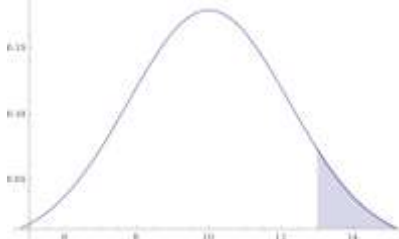
- a) Da medida da corrente ser de no máximo 12 miliamperes.

Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



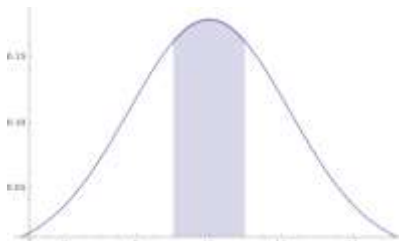
$$P(X \leq 12) = P\left(Z \leq \frac{12 - 10}{\sqrt{5}}\right) = P(Z \leq 0,89) = 0,5 + 0,3133 = 0,8133$$

- b) Da medida da corrente ser de pelo menos 13 miliamperes.
Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



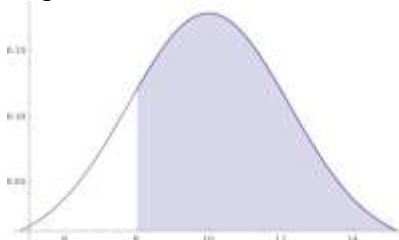
$$P(X \geq 13) = P\left(Z \geq \frac{13 - 10}{\sqrt{5}}\right) = P(Z \geq 1,34) = 0,5 - 0,4099 = 0,0901$$

- c) Um valor entre 9 e 11 miliamperes.
Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



$$\begin{aligned} P(9 < X < 11) &= P\left(\frac{9 - 10}{\sqrt{5}} < Z < \frac{11 - 10}{\sqrt{5}}\right) = P(-0,45 < Z < +0,45) \\ &= 0,1736 + 0,1736 = 0,3472 \end{aligned}$$

- d) Maior do que 8 miliamperes.
Graficamente, a probabilidade desejada pode ser representada da seguinte maneira:



$$P(X > 8) = P\left(Z > \frac{8 - 10}{\sqrt{5}}\right) = P(Z > -0,89) = 0,5 + 0,3133 = 0,8133$$

Exercício 3

Considere que a pontuação obtida por diferentes candidatos em um concurso público segue uma distribuição aproximadamente normal, com média igual a 140 pontos e desvio padrão igual a 20 pontos. Suponha que um candidato é escolhido ao acaso. Calcule as probabilidades a seguir:

- a) Apresentar uma pontuação entre 140 e 165,6.
- b) Apresentar uma pontuação entre 127,4 e 140.
- c) Apresentar uma pontuação entre 117,2 e 157.
- d) Apresentar uma pontuação inferior a 127.
- e) Apresentar uma pontuação superior a 174,2.
- f) Apresentar uma pontuação inferior a 167,4.
- g) Apresentar uma pontuação entre 155,4 e 168,4.

Respostas: a) 0,3997 b) 0,2357 c) 0,6752 d) 0,2578 e) 0,0436 f) 0,9147 g) 0,1428

Exercício 4

As vendas diárias de um mercado de bairro seguem, aproximadamente, uma distribuição normal, com média igual a R\$5.000,00 e desvio padrão igual a R\$2.000,00. Calcule a probabilidade de que, em um determinado dia, as vendas:

- a) Sejam superiores a R\$3.500,00?
- b) Sejam inferiores a R\$3.000,00?
- c) Estejam entre R\$3.800,00 e R\$5.300,00?
- d) Estejam entre R\$2.100,00 e 7.800,00?

Respostas: a) 0,7734 b) 0,1587 c) 0,2854 d) 0,8457

