Modelo Normal

Um modelo de distribuição de variável aleatória contínua, dentre os mais utilizados, é o Modelo Normal, que estudaremos neste tópico.

Inicialmente, consideremos uma situação que nos permitirá introduzir o modelo.

A temperatura diária em certa região, no mês de janeiro, pode ser considerada uma variável aleatória contínua Z, com função densidade de probabilidades dada pela lei:

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-0{,}5z^2} \;, \qquad \forall z \in \mathbb{R}$$

Qual a probabilidade de que em determinado dia do mês de janeiro a temperatura:

(a) esteja entre 0ºC e 1,5ºC?

(b) esteja entre 1,37ºC e 2,06ºC?

(c) esteja entre -5,29ºC e -3,9ºC?

(d) seja exatamente 3ºC?

(e) seja maior que 2,6ºC?

(f) esteja entre -2ºC e 2ºC?

(g) esteja entre -1,93ºC e -0,4ºC?

(h) esteja entre -1,4ºC e 3,08ºC?

(i) seja maior que -1,03ºC?

(j) seja maior que 6,27ºC?

(k) seja menor que 9ºC?

(l) seja maior que -5ºC?

Aparentemente, essa parece ser uma situação de variável aleatória contínua semelhante às que estudamos no Tópico 3. Entretanto, veremos que se trata de uma situação com algumas diferenças interessantes.

Pelo que estudamos, a área sob a curva da f.d.p. deve ser igual a um. Para comprovarmos tal fato, devemos ter:

$$\int_{-\infty}^{+\infty} f(z)dz = 1$$

Ou seja, devemos demonstrar que:

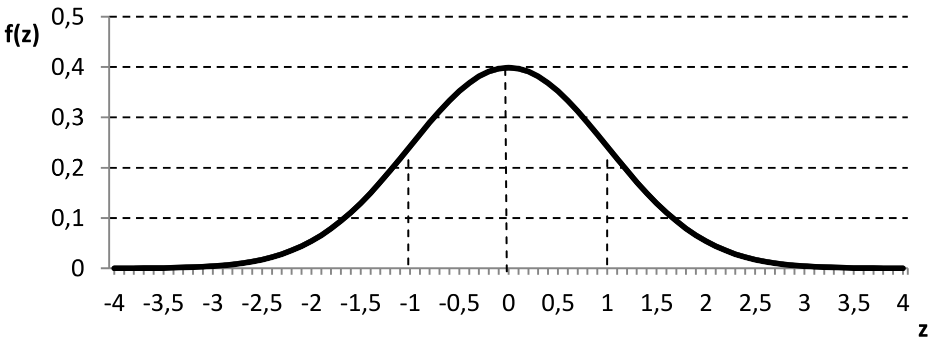
$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-0{,}5z^2} dz = 1$$

Já nessa etapa, temos uma pequena dificuldade: essa função não tem primitiva. Em outras palavras, não é possível calcular essa integral pelo método de Riemann, que foi o método de integração desenvolvido no Cálculo Integral. Precisamos introduzir um artifício para provarmos que essa integral é igual a um. Esse artifício é nomear essa integral de I e calcular I^2, introduzindo coordenadas polares. Como essa demonstração foge aos propósitos de nossa disciplina, nós a omitiremos neste texto.

Entretanto, a representação gráfica dessa função pode ser feita valendo-se do estudo de função que você já desenvolveu no curso de Cálculo Diferencial. Entre outros itens, para construir o gráfico:

- determinamos o campo de definição;
- verificamos se há pontos de descontinuidade, ou não, e se houver, qual é o comportamento da função nos limites, pela direita e pela esquerda, desses pontos;
- determinamos o comportamento da função no limite quando $z \to +\infty$ e quando $z \to -\infty$;
- verificamos se a função tem eixo de simetria, e se tiver, qual é a sua equação;
- determinamos se a função tem assíntota horizontal, vertical, ou inclinada e, em caso afirmativo, calculamos a equação de cada uma delas;
- verificamos se existem e, em caso afirmativo, calculamos os pontos de máximo, ou de mínimo, relativos da função;
- estudamos a concavidade da função, para todos os valores em que a função está definida;
- verificamos se existem e, em caso afirmativo, determinamos os pontos de inflexão.

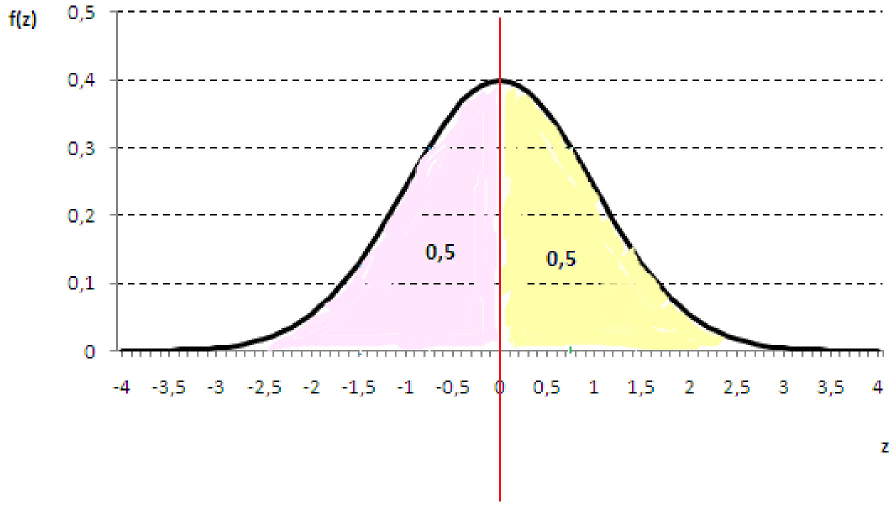
É um bom exercício fazer esse estudo da função densidade de probabilidades definida anteriormente e construir a sua representação gráfica, antes de dar continuidade à leitura. A seguir, essa representação é exibida.



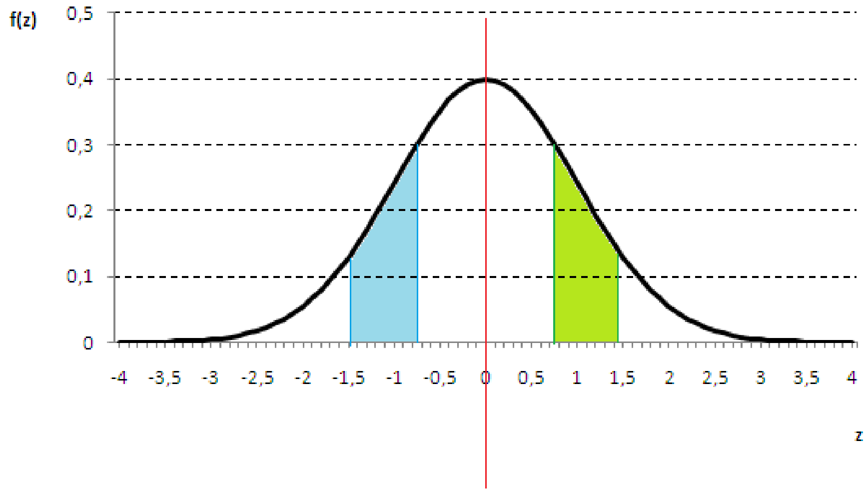
No estudo de função, você deve ter concluído que essa função tem o eixo das abscissas como uma assíntota horizontal, isto é, a curva da função se aproxima do eixo das abscissas, quando $z \to +\infty$ e quando $z \to -\infty$, sem interceptar o eixo, que a reta $x = 0$ é um eixo de simetria da função, que a função tem um ponto de máximo absoluto que é $\left(0\,, \frac{1}{\sqrt{2\pi}}\right)$, que a função tem dois pontos de inflexão que são: $\left(-1, \frac{e^{-0,5}}{\sqrt{2\pi}}\right)$ e

$\left(1, \frac{e^{-0,5}}{\sqrt{2\pi}}\right)$, além de que a concavidade é para baixo no intervalo $[-1,1]$ e para cima fora desse intervalo.

Como conseqüência de que a reta $z = 0$ é um eixo de simetria da função, e que a medida da área total sob a curva é igual a um, temos que cada uma das duas metades da área sob a curva tem medida igual a 0,5.



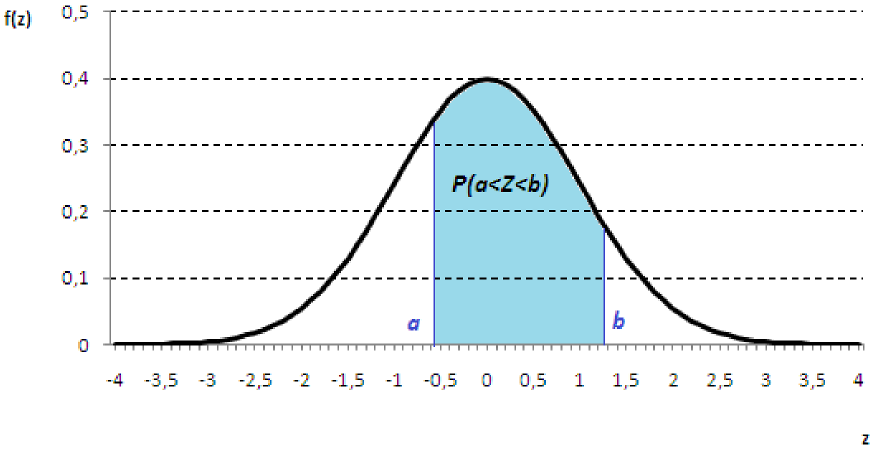
Ainda mais, toda e qualquer área à direita do eixo de simetria, tem outra equivalente à esquerda, como mostra a representação a seguir.



A situação propõe que se calcule a probabilidade da v.a.c. Z (temperatura em ºC naquela localidade no mês de janeiro) pertencer a alguns intervalos. Lembrando que, no caso das variáveis aleatórias contínuas, probabilidade é igual à medida da área sob a curva da função densidade, sendo a v.a.c. Z, com f.d.p. dada por: $f(z) = \frac{1}{\sqrt{2\pi}}\ e^{-0,5z^2}$, $\forall z \in \mathbb{R}$, temos que:

$$P(a < Z < b) = \int_a^b f(z)dz = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-0,5z^2} dz \quad , \forall\, a, b \in \mathbb{R}$$

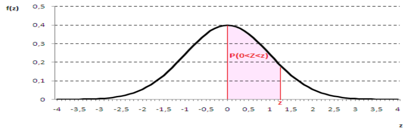
que podemos representar por:



Da mesma forma que não foi possível calcular a área sob a curva da função densidade de probabilidades, diretamente pelo Método de Riemann, porque a f.d.p. não tem primitiva, não é possível calcular essa integral, pelo Método de Riemann, quaisquer que sejam os valores reais de *a* e *b*.

Para resolvermos essa questão, vamos nos valer de métodos numéricos de determinação das medidas de tais áreas, quando quisermos calcular a probabilidade da v.a.c. Z pertencer a algum intervalo real. Como também a determinação desses valores envolve muitos cálculos, vamos utilizar uma tabela, na qual foram determinadas, por métodos numéricos, algumas dessas áreas (essa tabela pode ser encontrada em praticamente todos os livros de Estatística, com mais, ou, menos casas decimais, ou, ainda, tabelas equivalentes). O nosso compromisso é aprender a utilizá-la, já que para tal, dependemos das propriedades sob a curva, que já apresentamos. A seguir a tabela que iremos utilizar (você deve imprimir a página da tabela para acompanhar a sua utilização e para resolver os exercícios).

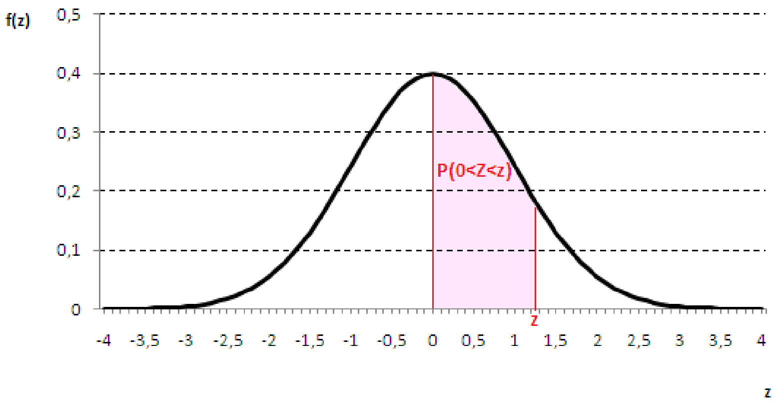
Normal padrão - Z~N(0,1)



z	0	1	2	3	4	5	6	7	8	9	z
0,0	0,000000	0,003989	0,007978	0,011966	0,015953	0,019939	0,023922	0,027903	0,031881	0,035856	**0,0**
0,1	0,039828	0,043795	0,047758	0,051717	0,055670	0,059618	0,063559	0,067495	0,071424	0,075345	**0,1**
0,2	0,079260	0,083166	0,087064	0,090954	0,094835	0,098706	0,102568	0,106420	0,110261	0,114092	**0,2**
0,3	0,117911	0,121720	0,125516	0,129300	0,133072	0,136831	0,140576	0,144309	0,148027	0,151732	**0,3**
0,4	0,155422	0,159097	0,162757	0,166402	0,170031	0,173645	0,177242	0,180822	0,184386	0,187933	**0,4**
0,5	0,191462	0,194974	0,198468	0,201944	0,205401	0,208840	0,212260	0,215661	0,219043	0,222405	**0,5**
0,6	0,225747	0,229069	0,232371	0,235653	0,238914	0,242154	0,245373	0,248571	0,251748	0,254903	**0,6**
0,7	0,258036	0,261148	0,264238	0,267305	0,270350	0,273373	0,276373	0,279350	0,282305	0,285236	**0,7**
0,8	0,288145	0,291030	0,293892	0,296731	0,299546	0,302337	0,305105	0,307850	0,310570	0,313267	**0,8**
0,9	0,315940	0,318589	0,321214	0,323814	0,326391	0,328944	0,331472	0,333977	0,336457	0,338913	**0,9**
1,0	0,341345	0,343752	0,346136	0,348495	0,350830	0,353141	0,355428	0,357690	0,359929	0,362143	**1,0**
1,1	0,364334	0,366500	0,368643	0,370762	0,372857	0,374928	0,376976	0,379000	0,381000	0,382977	**1,1**
1,2	0,384930	0,386861	0,388768	0,390651	0,392512	0,394350	0,396165	0,397958	0,399727	0,401475	**1,2**
1,3	0,403200	0,404902	0,406582	0,408241	0,409877	0,411492	0,413085	0,414657	0,416207	0,417736	**1,3**
1,4	0,419243	0,420730	0,422196	0,423641	0,425066	0,426471	0,427855	0,429219	0,430563	0,431888	**1,4**
1,5	0,433193	0,434478	0,435745	0,436992	0,438220	0,439429	0,440620	0,441792	0,442947	0,444083	**1,5**
1,6	0,445201	0,446301	0,447384	0,448449	0,449497	0,450529	0,451543	0,452540	0,453521	0,454486	**1,6**
1,7	0,455435	0,456367	0,457284	0,458185	0,459070	0,459941	0,460796	0,461636	0,462462	0,463273	**1,7**
1,8	0,464070	0,464852	0,465620	0,466375	0,467116	0,467843	0,468557	0,469258	0,469946	0,470621	**1,8**
1,9	0,471283	0,471933	0,472571	0,473197	0,473810	0,474412	0,475002	0,475581	0,476148	0,476705	**1,9**
2,0	0,477250	0,477784	0,478308	0,478822	0,479325	0,479818	0,480301	0,480774	0,481237	0,481691	**2,0**
2,1	0,482136	0,482571	0,482997	0,483414	0,483823	0,484222	0,484614	0,484997	0,485371	0,485738	**2,1**
2,2	0,486097	0,486447	0,486791	0,487126	0,487455	0,487776	0,488089	0,488396	0,488696	0,488989	**2,2**
2,3	0,489276	0,489556	0,489830	0,490097	0,490358	0,490613	0,490863	0,491106	0,491344	0,491576	**2,3**
2,4	0,491802	0,492024	0,492240	0,492451	0,492656	0,492857	0,493053	0,493244	0,493431	0,493613	**2,4**
2,5	0,493790	0,493963	0,494132	0,494297	0,494457	0,494614	0,494766	0,494915	0,495060	0,495201	**2,5**
2,6	0,495339	0,495473	0,495604	0,495731	0,495855	0,495975	0,496093	0,496207	0,496319	0,496427	**2,6**
2,7	0,496533	0,496636	0,496736	0,496833	0,496928	0,497020	0,497110	0,497197	0,497282	0,497365	**2,7**
2,8	0,497445	0,497523	0,497599	0,497673	0,497744	0,497814	0,497882	0,497948	0,498012	0,498074	**2,8**
2,9	0,498134	0,498193	0,498250	0,498305	0,498359	0,498411	0,498462	0,498511	0,498559	0,498605	**2,9**
3,0	0,498650	0,498694	0,498736	0,498777	0,498817	0,498856	0,498893	0,498930	0,498965	0,498999	**3,0**
3,1	0,499032	0,499065	0,499096	0,499126	0,499155	0,499184	0,499211	0,499238	0,499264	0,499289	**3,1**
3,2	0,499313	0,499336	0,499359	0,499381	0,499402	0,499423	0,499443	0,499462	0,499481	0,499499	**3,2**
3,3	0,499517	0,499534	0,499550	0,499566	0,499581	0,499596	0,499610	0,499624	0,499638	0,499651	**3,3**
3,4	0,499663	0,499675	0,499687	0,499698	0,499709	0,499720	0,499730	0,499740	0,499749	0,499758	**3,4**
3,5	0,499767	0,499776	0,499784	0,499792	0,499800	0,499807	0,499815	0,499822	0,499828	0,499835	**3,5**
3,6	0,499841	0,499847	0,499853	0,499858	0,499864	0,499869	0,499874	0,499879	0,499883	0,499888	**3,6**
3,7	0,499892	0,499896	0,499900	0,499904	0,499908	0,499912	0,499915	0,499918	0,499922	0,499925	**3,7**
3,8	0,499928	0,499931	0,499933	0,499936	0,499938	0,499941	0,499943	0,499946	0,499948	0,499950	**3,8**
3,9	0,499952	0,499954	0,499956	0,499958	0,499959	0,499961	0,499963	0,499964	0,499966	0,499967	**3,9**
4,0	0,499968	0,499970	0,499971	0,499972	0,499973	0,499974	0,499975	0,499976	0,499977	0,499978	**4,0**
4,1	0,499979	0,499980	0,499981	0,499982	0,499983	0,499983	0,499984	0,499985	0,499985	0,499986	**4,1**
4,2	0,499987	0,499987	0,499988	0,499988	0,499989	0,499989	0,499990	0,499990	0,499991	0,499991	**4,2**
4,3	0,499991	0,499992	0,499992	0,499993	0,499993	0,499993	0,499993	0,499994	0,499994	0,499994	**4,3**
4,4	0,499995	0,499995	0,499995	0,499995	0,499996	0,499996	0,499996	0,499996	0,499996	0,499996	**4,4**
4,5	0,499997	0,499997	0,499997	0,499997	0,499997	0,499997	0,499997	0,499998	0,499998	0,499998	**4,5**
4,6	0,499998	0,499998	0,499998	0,499998	0,499998	0,499998	0,499998	0,499998	0,499999	0,499999	**4,6**
4,7	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	**4,7**
4,8	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	0,499999	**4,8**
4,9	0,500000	0,500000	0,500000	0,500000	0,500000	0,500000	0,500000	0,500000	0,500000	0,500000	**4,9**
z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**z**

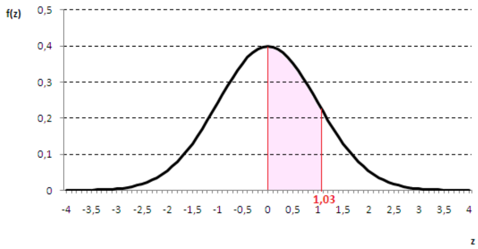
Profa. Dra. Maria Inez Rodrigues Miguel - PUC/SP

Conforme representação no alto da tabela (ampliada logo abaixo), ela fornece apenas a medida da área em um intervalo de zero até algum valor **z**, ou seja, a faixa central, como representada na figura a seguir.



Nessa tabela temos os valores de **z (de 0 a 4,99, com duas casas decimais)** nas bordas da tabela e no miolo temos as medidas das áreas correspondentes ao intervalo de zero a **z**, ou seja, as **probabilidades (com seis casas decimais)** da variável Z pertencer a esse intervalo. A fim de melhor disponibilizar esses valores em uma tabela, os valores inteiros e primeira casa decimal estão nas bordas laterais (esquerda e direita) e os valores da segunda casa decimal estão nas bordas superior e inferior.

Por exemplo, se queremos determinar $P(0 < Z < 1{,}03)$. Note que essa probabilidade corresponde à área representada a seguir:

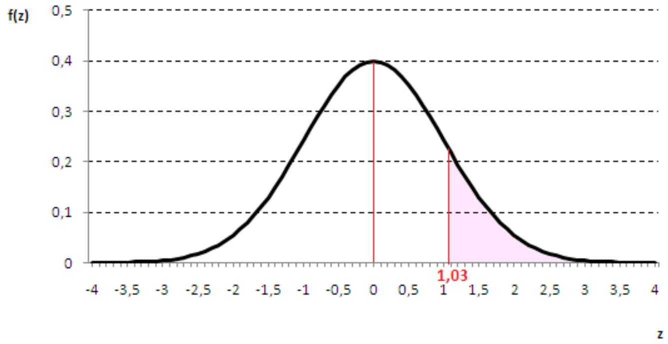


Buscamos na tabela o valor da linha do 1,0 e na coluna do 3, obtendo 0,348495.

z	0	1	2	3	4
0,0	0,000000	0,003989	0,007978	0,011966	0,015953
0,1	0,039828	0,043795	0,047758	0,051717	0,055670
0,2	0,079260	0,083166	0,087064	0,090954	0,094835
0,3	0,117911	0,121720	0,125516	0,129300	0,133072
0,4	0,155422	0,159097	0,162757	0,166402	0,170031
0,5	0,191462	0,194974	0,198468	0,201944	0,205401
0,6	0,225747	0,229069	0,232371	0,235653	0,238914
0,7	0,258036	0,261148	0,264238	0,267305	0,270350
0,8	0,288145	0,291030	0,293892	0,296731	0,299546
0,9	0,315940	0,318589	0,321214	0,323814	0,326391
1,0	0,341345	0,343752	0,346136	0,348495	0,350830
1,1	0,364334	0,366500	0,368643	0,370762	0,372857

Assim, $P(0 < Z < 1{,}03) = 0{,}348495.$

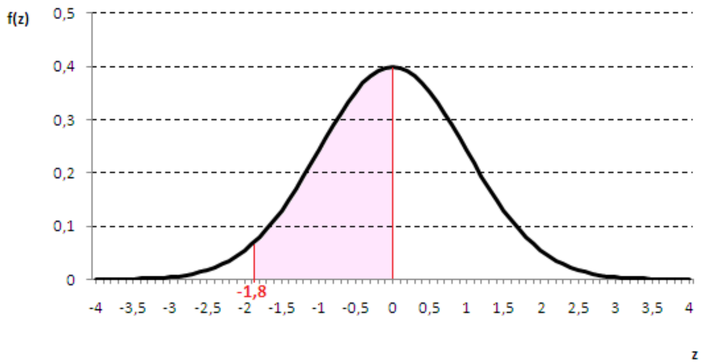
Outro exemplo: queremos determinar $P(Z > 1{,}03)$, que corresponde à área representada a seguir:



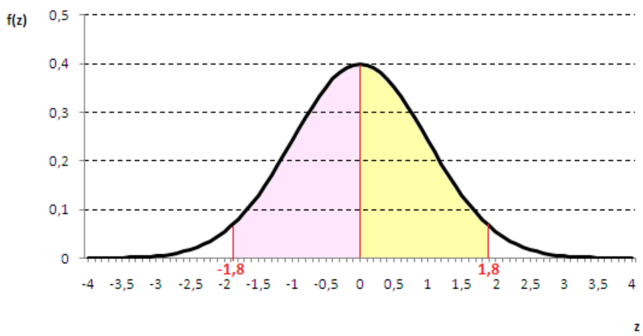
O valor a ser procurado na tabela é o mesmo: 1,03, do qual se obtém 0,348495. Só que esse valor é da faixa central e queremos a medida da área da cauda da direita. Como de zero a infinito a área tem valor 0,5 (metade da área total, que é um), a área da cauda, que é a que queremos, vale $0{,}5 - 0{,}348495 = 0{,}1515{,}5$. Assim,

$P(Z > 1{,}03) = 0{,}151505$.

Outro exemplo: queremos determinar $P(-1{,}8 < Z < 0)$, que corresponde à área representada a seguir:



Considerando o eixo de simetria $(z = 0)$, essa área é exatamente igual à área correspondente a $P(0 < Z < 1{,}8)$, como mostrado na representação a seguir.

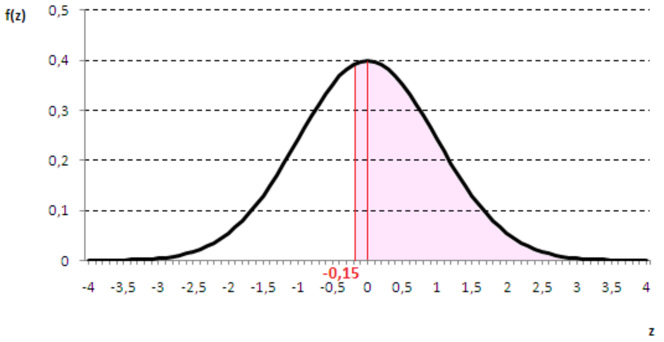


Assim, $P(-1{,}8 < Z < 0) = P(0 < Z < 1{,}8)$. O valor a ser procurado na tabela é 1,8. Buscando o valor correspondente à linha do 1,8 e à coluna do zero (a segunda casa decimal é zero, porque $1{,}8 = 1{,}80$), encontramos 0,464070. Portanto, temos que:

$$P(-1{,}8 < Z < 0) = 0{,}464070$$

Observe que, fazer um rascunho da representação da área que se quer determinar a probabilidade, facilita a visualização das propriedades da curva para que possamos determinar corretamente essa probabilidade.

Outro exemplo: queremos determinar $P(Z > -0{,}15)$, que corresponde à área representada a seguir:

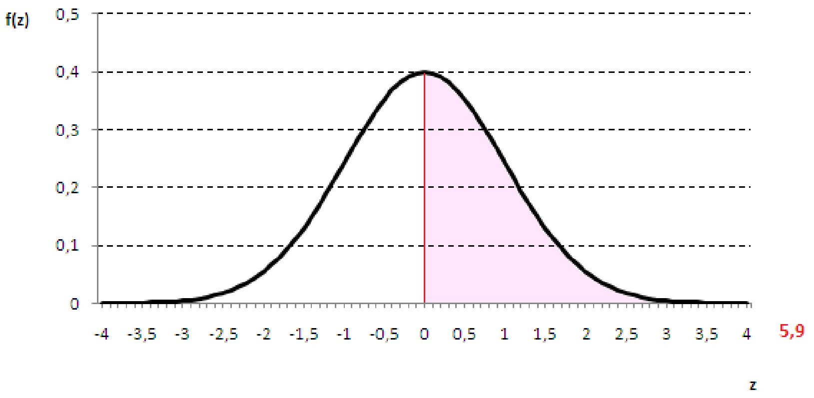


$$P(Z > -0{,}15) = P(-0{,}15 < Z < 0) + P(0 < Z < +\infty)$$

Lembrando que $P(-0{,}15 < Z < 0) = P(0 < Z < 0{,}15)$, basta procurarmos 0,15 na tabela (linha do 0,1 e coluna do 5), obtendo: 0,059618. $P(0 < Z < +\infty) = 0{,}5$, já que representa metade da área total. Portanto, temos que:

$$P(Z > -0{,}15) = 0{,}059618 + 0{,}5 = 0{,}559618$$

Outro exemplo: queremos determinar $P(0 < Z < 5{,}9)$, que corresponde à área representada a seguir:

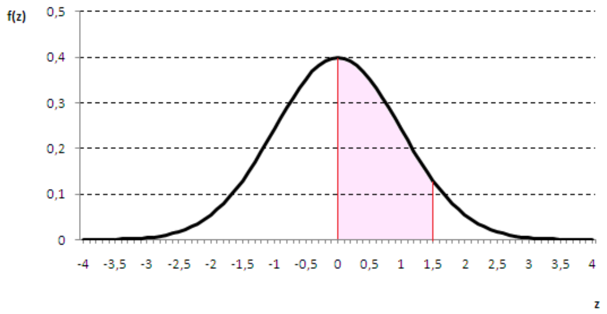


Note que na tabela não consta o valor 5,9, além de que a última linha da tabela, que contém arredondamento até 6 casas decimais, tem todos os valores iguais a 0,500000. Isso significa que, com aproximação de 6 casas decimais, a área da cauda para valores maiores, ou iguais, a 4,9 são desprezíveis (é como se a curva fosse limitada, de modo que a área de $-4{,}9$ a $+4{,}9$ fosse igual a um). Dessa forma, qualquer valor maior ou igual a 4,9 tem valor da tabela igual a 0,5, ou seja, a área de zero até qualquer valor maior ou igual a 4,9 já é 0,5, com aproximação de 6 casas decimais.

Assim, $P(0 < Z < 5{,}9) = 0{,}5$, como também $P(Z > 5{,}9) = 0$, isto é, a cauda à direita de 5,9 é zero com aproximação de seis casas decimais.

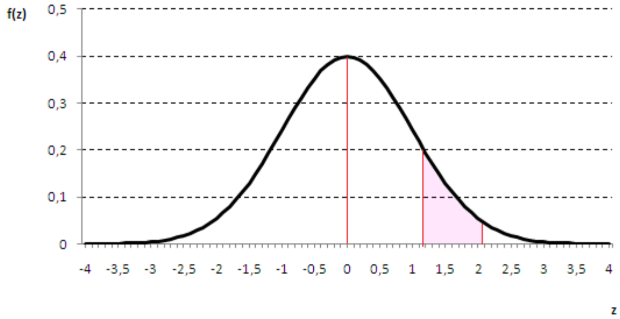
Voltemos à situação proposta no início e calculemos as probabilidades pedidas.

a) esteja entre 0ºC e 1,5ºC



$$P(0 < Z < 1{,}5) = 0{,}433193$$

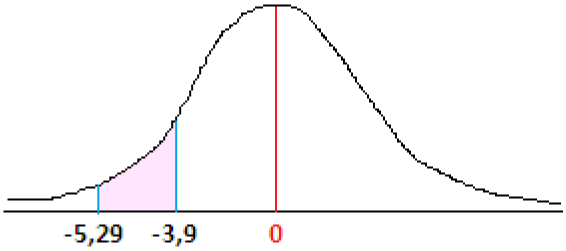
b) esteja entre 1,37ºC e 2,06ºC



$$P(1{,}37 < Z < 2{,}06) = P(0 < Z < 2{,}06) - P(0 < Z < 1{,}37) =$$
$$= 0{,}480301 - 0{,}414657 = 0{,}065644$$

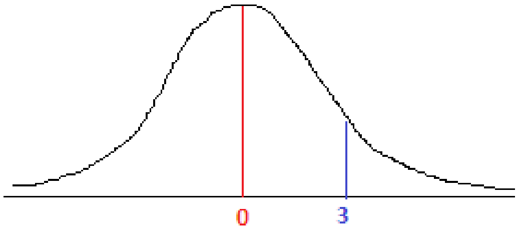
Observe que neste último item, fizemos a diferença das duas faixas centrais que a tabela fornece, para obtermos a faixa de área desejada. Outro aspecto a salientar, mais uma vez, é a importância de uma representação gráfica para favorecer a determinação correta das probabilidades pedidas. Não precisa ser uma representação detalhada e precisa, mas, apenas representativa, como a que faremos nos itens que se seguem.

c) Esteja entre -5,29ºC e -3,9ºC



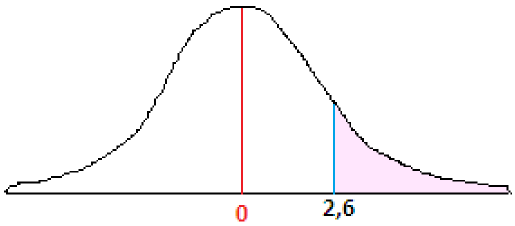
$P(-5{,}29 < Z < -3{,}9) = 0{,}5 - 0{,}499952 = 0{,}000048$. Note que, neste caso, o valor na tabela para 5,29 já extrapola o valor 4,9 e, portanto, vale 0,5, com aproximação de seis casas decimais.

d) Seja exatamente 3ºC



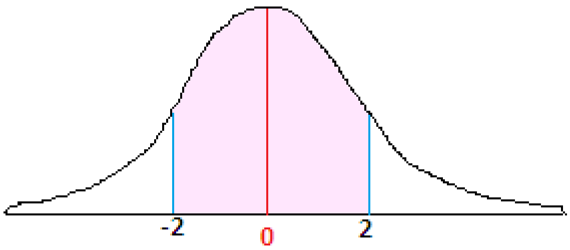
$P(Z = 3) = 0$. Em um ponto, a medida da área é zero, assim como a probabilidade. Esse resultado vem ao encontro de que o valor exato de uma medida não é possível determinar.

e) Seja maior que 2,6ºC



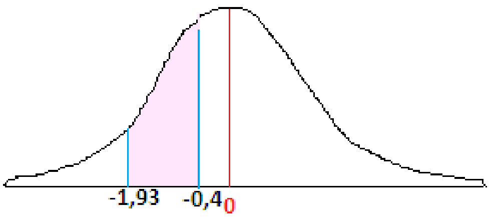
$$P(Z > 2{,}6) = 0{,}5 - 0{,}495339 = 0{,}004661$$

f) Esteja entre -2ºC e +2ºC



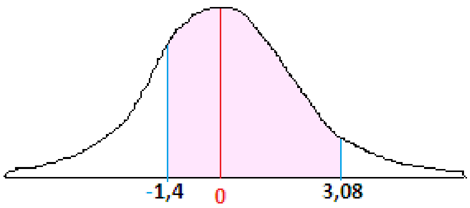
$$P(-2 < Z < 2) = 0{,}477250 + 0{,}477250 = 0{,}9545$$

g) Esteja entre -1,93ºC e -0,4ºC



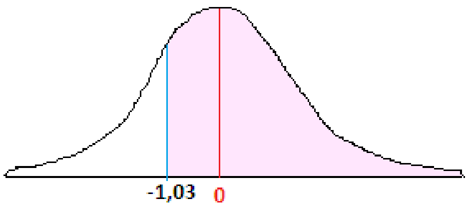
$$P(-1{,}93 < Z < -0{,}4) = 0{,}473197 - 0{,}155422 = 0{,}317775$$

h) Esteja entre -1,4ºC e +3,08ºC



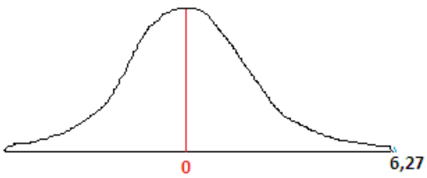
$$P(-1{,}4 < Z < 3{,}08) = 0{,}419243 + 0{,}498965 = 0{,}918208$$

i) Seja maior que -1,03ºC



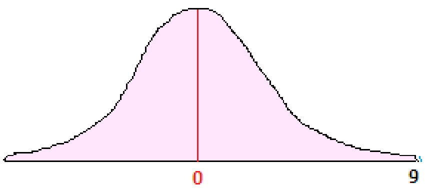
$$P(Z > -1{,}03) = 0{,}348495 + 0{,}5 = 0{,}848495$$

j) Seja maior que 6,27ºC



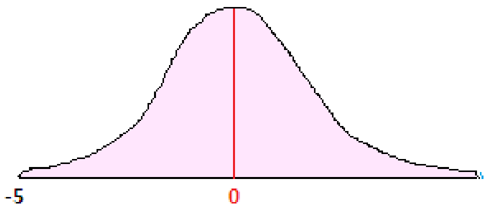
$$P(Z > 6{,}27) = 0$$

k) Seja menor que 9ºC



$$P(Z < 9) = 1$$

l) Seja maior que -5ºC



$$P(Z > -5) = 1$$

A seguir propomos um exercício para você trinar o uso da Tabela da v.a.c. Z com f.d.p. dada pela lei: $f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-0,5z^2}$, $\forall z \in \mathbb{R}$.

Exercício 1. Sendo Z uma v.a.c. com f.d.p. dada pela Lei :$\boldsymbol{f(z)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}\boldsymbol{\pi}}}\, \boldsymbol{e}^{\mathbf{-0,5}\boldsymbol{z}^{\mathbf{2}}}$ *,* $\forall \boldsymbol{z} \in \mathbb{R}$, *determine as seguintes probabilidades: a)* $\boldsymbol{P}(\mathbf{0} \leq \boldsymbol{Z} < 2,37)$; *b)* $\boldsymbol{P}(\boldsymbol{Z} < -2)$; *c)* $\boldsymbol{P}(\boldsymbol{Z} \geq \mathbf{-2,4})$; *d)* $\boldsymbol{P}(\mathbf{-0,5} \leq \boldsymbol{Z} \leq \mathbf{0,08})$; *e)* $\boldsymbol{P}(\mathbf{1,05} < Z < 3)$; *f)* $\boldsymbol{P}(\boldsymbol{Z} < -5,8)$; *g)* $\boldsymbol{P}(\boldsymbol{Z} \leq \mathbf{7,08})$; *h)* $\boldsymbol{P}(\mathbf{-0,16} < Z < -3,9)$; *i)* $\boldsymbol{P}(\mathbf{-1} < Z < 1)$; *j)* $\boldsymbol{P}(\mathbf{-2} < Z < 2)$; *k)* $\boldsymbol{P}(\mathbf{-3} < Z < 3)$.