Variáveis aleatórias contínuas

- Conceito
- Função densidade de probabilidade e função de distribuição
- Valor esperado e variância

Distribuições de probabilidade

- Uniforme
- Exponencial
- Normal
- χ^2
- t
- F

Variáveis aleatórias contínuas

Definição: São contínuas todas as variáveis cujo espaço amostral S_X é **não enumerável**.

⇨ Se X é uma variável aleatória contínua, X pode assumir qualquer valor num intervalo **[a; b]** ou no intervalo **$(-\infty;+\infty)$**.

⇨ O espaço S_X será sempre definido como um intervalo do conjunto dos **reais**, sendo, portanto, um conjunto infinito.

Exemplos:

- tempo de reação de uma mistura
- vida útil de um componente eletrônico
- peso de uma pessoa
- produção de leite de uma vaca
- quantidade de chuva que ocorre numa região

1. Função densidade de probabilidade

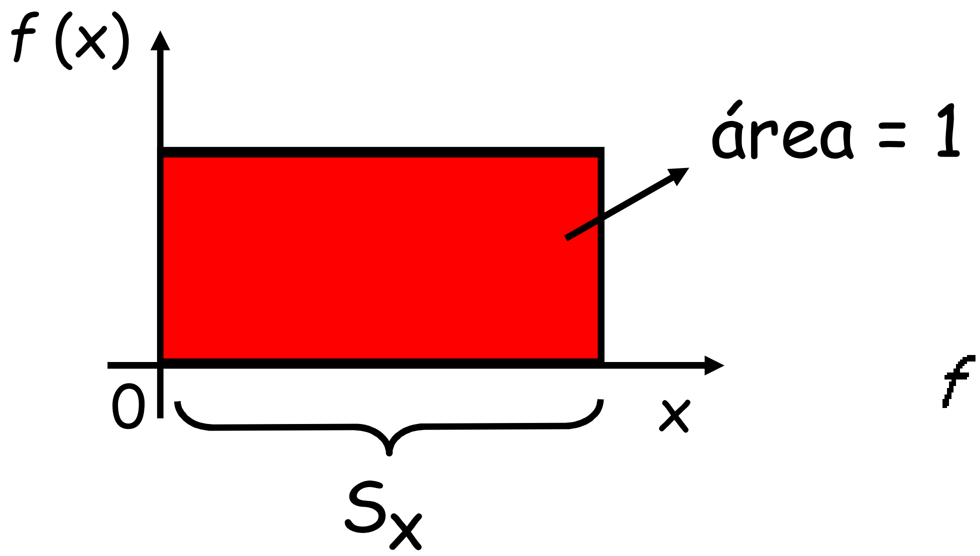
Definição: Seja X uma variável aleatória contínua e S_X o seu espaço amostral. Uma função **f** associada a variável X é denominada **função densidade de probabilidade (fdp)** se satisfizer duas condições:

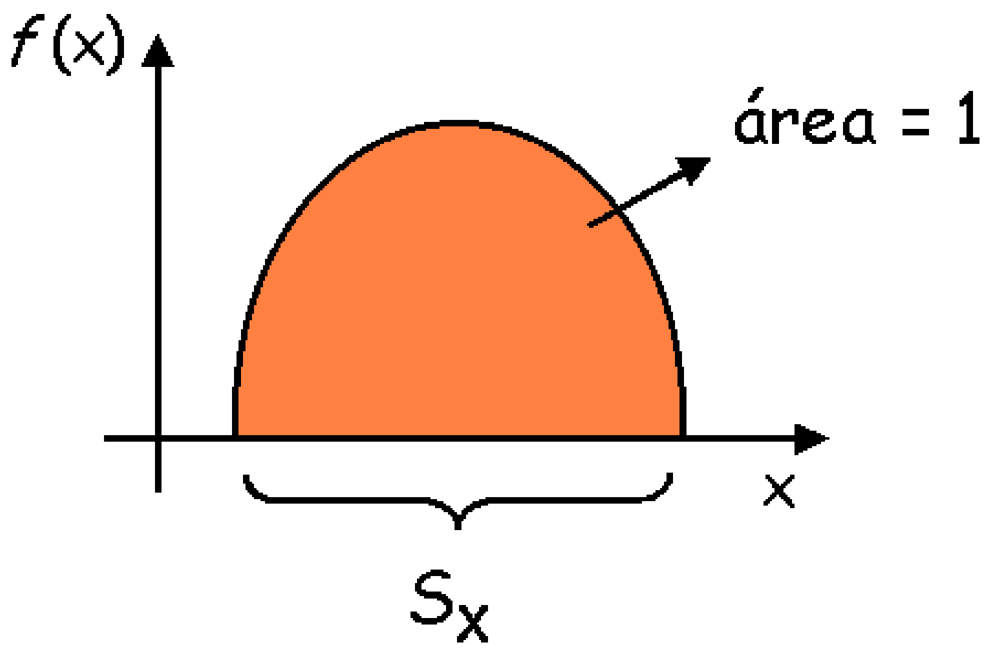
1. $f(x) \geq 0, \ \forall \ x \in S_X$

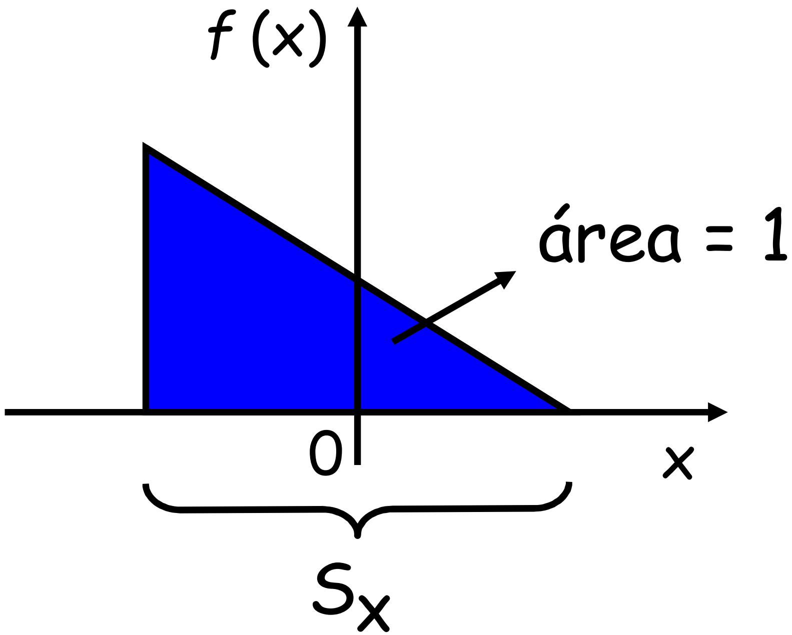
2. $\int_{S_X} f(x)dx = 1 \quad = P(X \in S_X)$

Esta área corresponde à probabilidade de a variável X pertencer ao espaço amostral S_X

É toda a função que não assuma valores negativos, ou cujo gráfico esteja acima do eixo das abscissas, e cuja área compreendida entre a função e o eixo das abscissas seja igual a 1 (um).

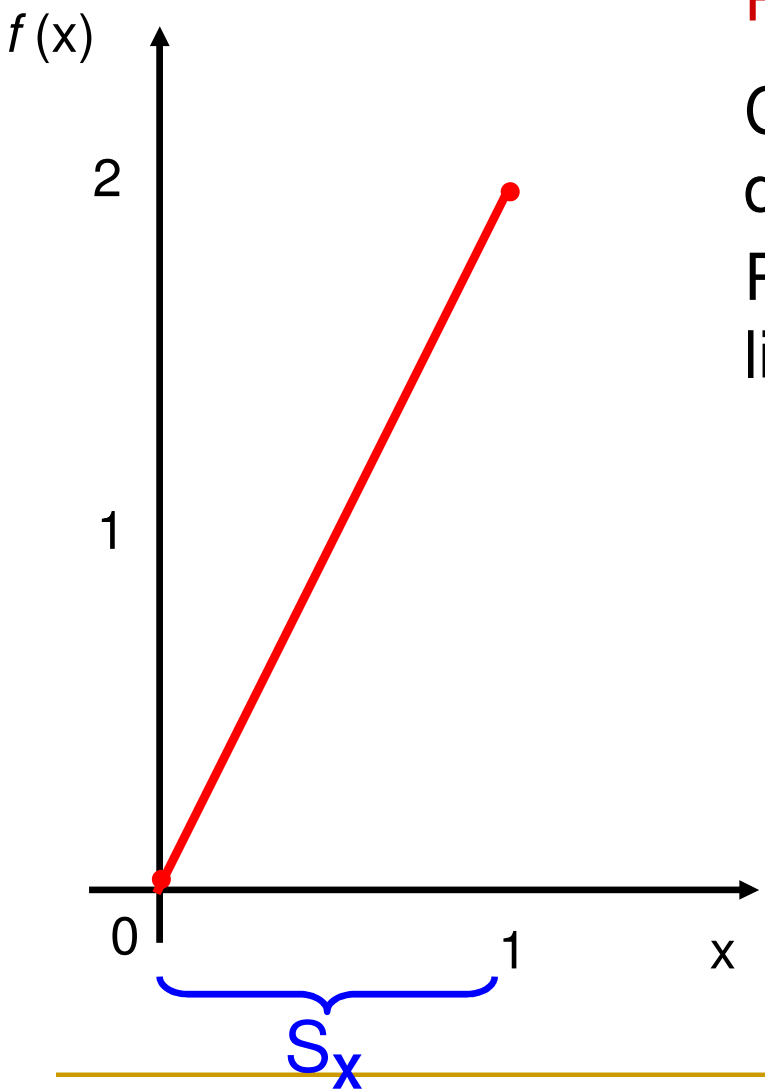

f (x)
área = 1
0
x
Sx


f (x)
área = 1
x
Sx


f (x)
área = 1
0
x
Sx

Exemplo:

Seja a função $f(x) = 2x$, no intervalo $S_X = [0,1]$. Verifique se f (x) é uma função densidade de probabilidade.



Primeira condição: $f(x) \geq 0,\ \forall\ x \in S_X$

Como a função é linear, são necessários dois pontos para traçar a reta.

Por conveniência esses pontos são os limites do intervalo S_X.

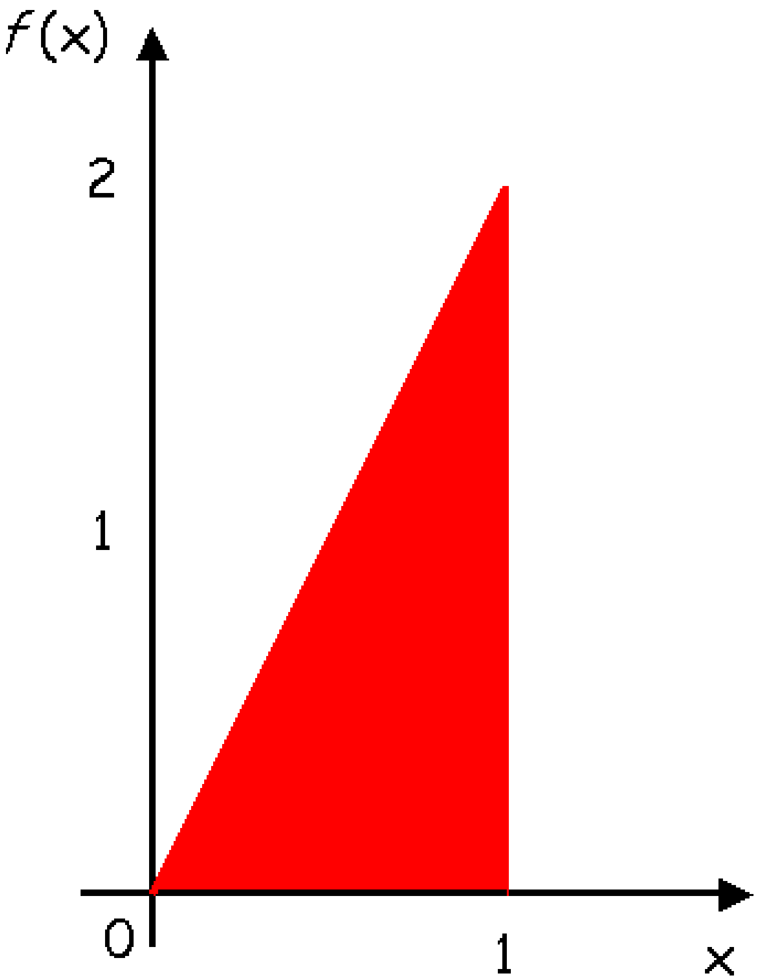
$$f(x) = 2x$$

$$f(x=0) = 2 \times 0 = 0$$

$$f(x=1) = 2 \times 1 = 2$$

Todos os valores da função f (x) são não negativos no intervalo de 0 a 1.

Segunda condição: $\int_{S_X} f(x)dx = 1$

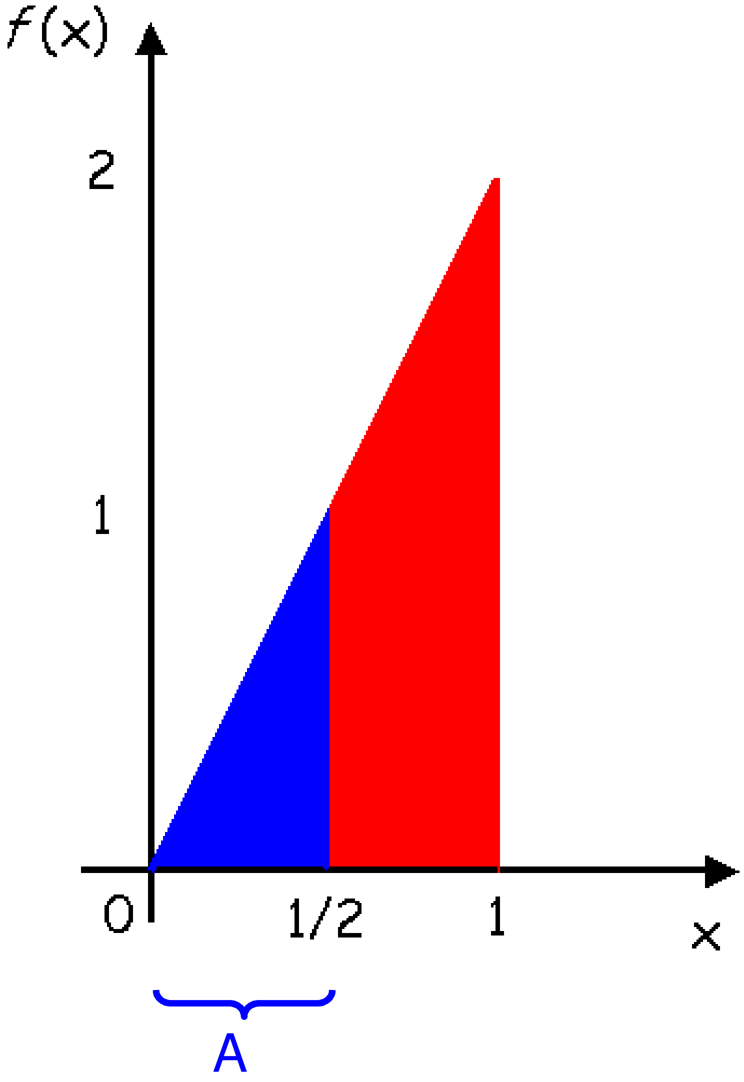


Área: $\frac{b \times h}{2} = \frac{1 \times 2}{2} = 1$

A área sob a função f (x) no intervalo S_X, que equivale a $P(X \in S_X)$, é igual a 1.

A função f (x) = 2x, no intervalo S_X =[0, 1] é uma função densidade de probabilidade!!

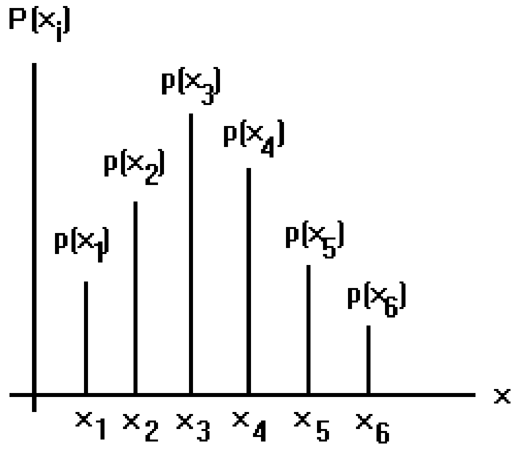
Seja A=[0, 1/2]. Qual é a probabilidade de ocorrer o evento A?

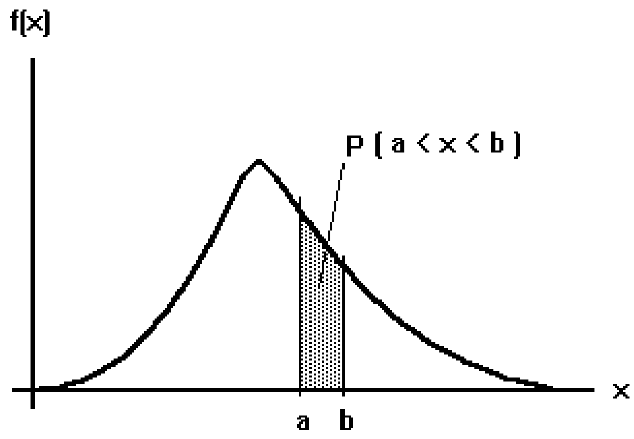


Probabilidade = área

$$\text{Área:} \quad \frac{b \times h}{2} = \frac{1/2 \times 1}{2} = \frac{1}{4}$$

$$P(0 \leq X \leq 1/2) = 1/4$$

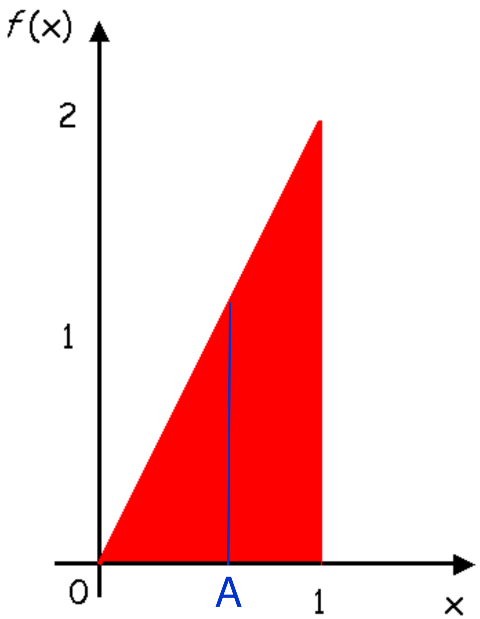




Importante!!!

No caso de variáveis contínuas, as representações $a \leq x \leq b$, $a \leq x < b$, $a < x \leq b$ e $a < x < b$ são todas equivalentes, pois a probabilidade num ponto, por definição, é nula.

Seja o evento $A = \{x;\ x = a\}$. Então,



$$P(A) = \int_A f(x)dx = \int_a^a f(x)dx = F(a) - F(a) = 0$$

2. Função de distribuição ou probabilidade acumulada

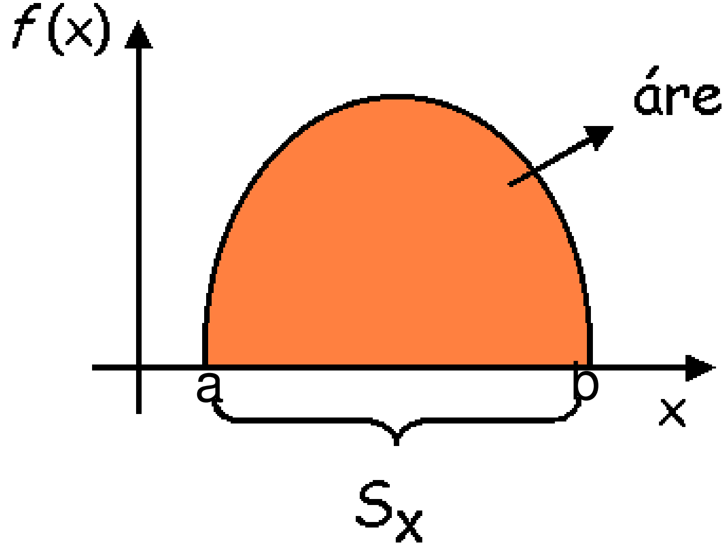
Definição: Seja X uma variável aleatória contínua e S_X o seu espaço amostral. A função de distribuição, denotada por **F(x)** ou **P(X≤x)**, é a **função que associa a cada ponto $x \in S_X$ a probabilidade P(X≤x)**. Desta forma, tem-se

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \quad \text{, para } S_X = (-\infty;+\infty).$$

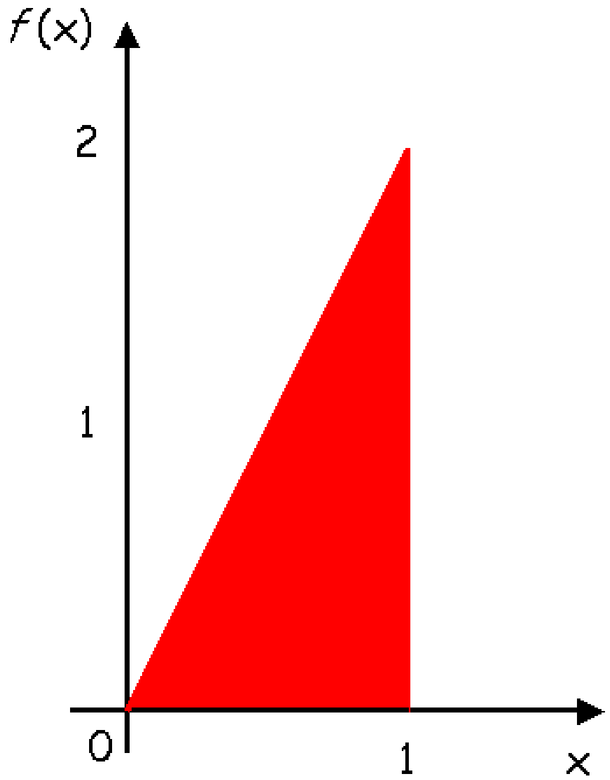
Se $S_X = [a, b]$, então

$$F(a) = P(X \leq a) = 0$$

$$F(b) = P(X \leq b) = 1$$



Exemplo: $f(x) = 2x, \ S_X = [0,1]$



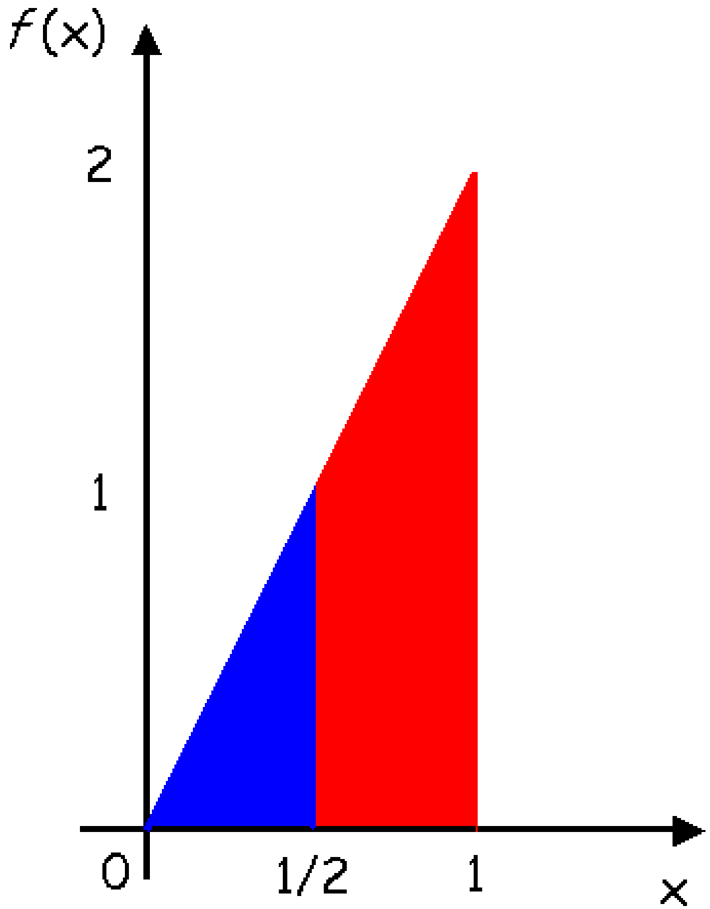
$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

$$= \int_{0}^{x} 2t\, dt$$

$$= 2\left[\frac{t^2}{2}\right]_0^x$$

$$F(x) = x^2$$

Exemplo: $f(x) = 2x, \ S_X = [0,1]$



$$F(x) = x^2$$

$$F(0) = P(X \leq 0) = 0^2 = 0$$

$$F(1) = P(X \leq 1) = 1^2 = 1$$

$$F(1/2) = P(X \leq 1/2) = (1/2)^2 = 1/4$$

$P(A) = F(1/2) = 1/4$

$P(B) = F(1) - F(1/2) = 1 - 1/4 = 3/4$

A=[0, 1/2]

B=[1/2, 1]

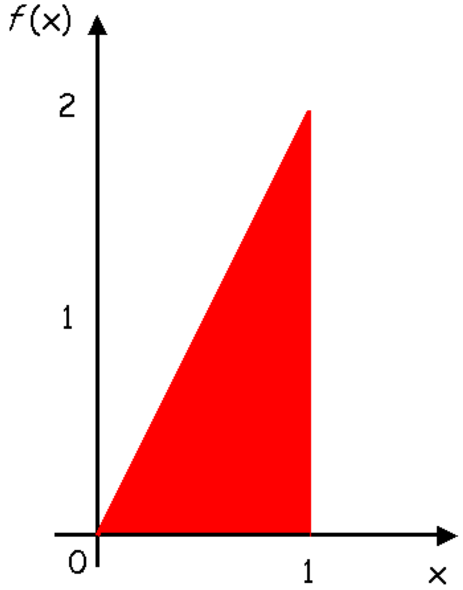
Medidas descritivas

Média ou valor esperado

Definição: Seja X uma variável aleatória contínua e S_X o seu espaço amostral. O valor esperado de X, denotado por **E(X)** ou μ , será dado por

$$E(X) = \mu = \int_{S_X} x\, f(x)dx$$

Exemplo: $f(x) = 2x, \; S_X = [0,1]$



$$E(X) = \mu = \int_{S_x} x\, f(x)\, dx$$

$$= \int_0^1 x\, 2x\, dx$$

$$= \int_0^1 2x^2\, dx = 2\int_0^1 x^2\, dx = 2\left[\frac{x^3}{3}\right]_0^1 = \frac{2}{3}$$

Medidas descritivas

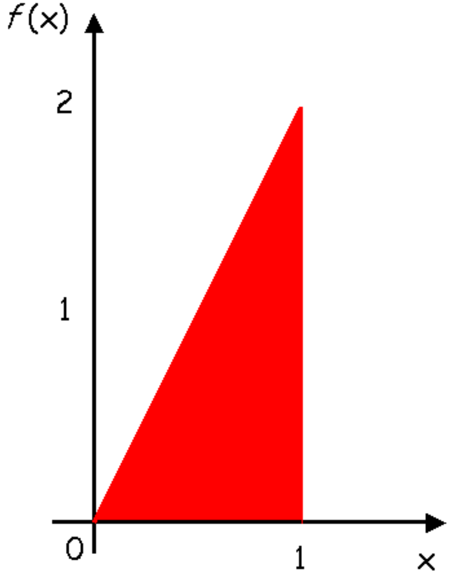
Variância

Definição: Seja X uma variável aleatória contínua e S_X o seu espaço amostral. A variância de X, denotada por **V(X)** ou **σ^2**, será dada por

$$V(X) = \sigma^2 = E(X-\mu)^2 = \int_{S_X} (x-\mu)^2 f(x)dx \qquad \text{(Fórmula de definição)}$$

$$V(X) = \sigma^2 = E(X^2) - \mu^2 = \left[\int_{S_X} x^2 f(x)dx\right] - \mu^2 \qquad \text{(Fórmula prática)}$$

Exemplo: $f(x) = 2x, \; S_X = [0,1]$



$$V(X) = \sigma^2 = E(X^2) - \mu^2$$

$$= \left[\int_0^1 x^2\, 2x\, dx\right] - \mu^2 = \left[\int_0^1 2x^3\, dx\right] - \left(\frac{2}{3}\right)^2$$

$$= \left\{2\left[\frac{x^4}{4}\right]_0^1\right\} - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18}$$

Exercício:

Determinar a média e a variância da vac cuja fdp é dada por:

$f(x) = 3x^2$ se $-1 \leq x \leq 0$

$\mu = -3/4$ e $\sigma^2 = 3/80$

Solução:

$$\mu = E(X) = \int_{-1}^{0} x\, f(x)\, dx = \int_{-1}^{0} x\,(3x^2)\, dx = \int_{-1}^{0} (3x^3)\, dx$$

$$= 3\left[\frac{x^4}{4}\right]_{-1}^{0} = -3/4$$

$$\sigma^2 = V(X) = \int_{-1}^{0} x^2 f(x)\, dx - \mu^2 = \int_{-1}^{0} x^2 (3x^2)\, dx - \left(-3/4\right)^2$$

$$= \int_{-1}^{0} 3x^4\, dx - \left(9/16\right) = 3\left[\frac{x^5}{5}\right]_{-1}^{0} - \left(9/16\right) = \frac{3}{5} - \frac{9}{16} = \frac{3}{80}$$

Distribuições de probabilidade de variáveis contínuas

⇨ É difícil identificar o tipo de distribuição de probabilidade de uma variável contínua

- pesquisa bibliográfica
- observação do campo de variação da variável

Existem vários tipos de distribuições contínuas:

⇨ **Distribuição Uniforme**

⇨ **Distribuição Exponencial**

⇨ **Distribuição Normal**

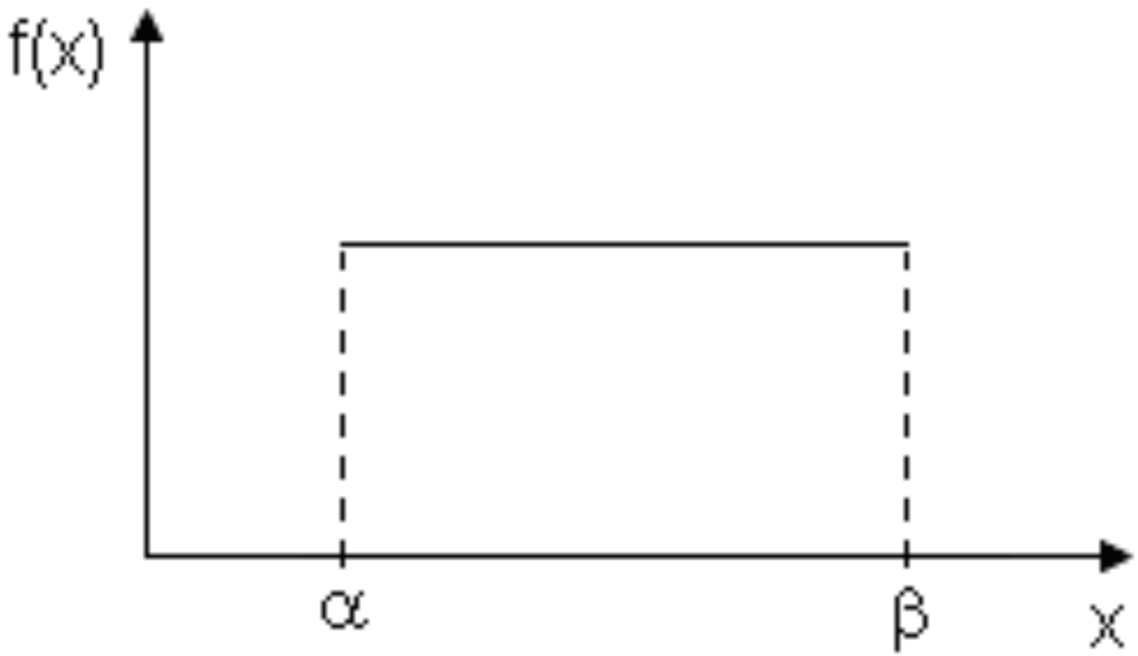
⇨ **Distribuição χ^2**

⇨ **Distribuição t**

⇨ **Distribuição F**

1. Distribuição Uniforme

Definição: Seja X uma variável aleatória contínua que assume valores no intervalo [α, β]. Se a probabilidade de X assumir valores num subintervalo é a **mesma** que para qualquer outro subintervalo de mesmo comprimento, então, esta variável tem distribuição uniforme.

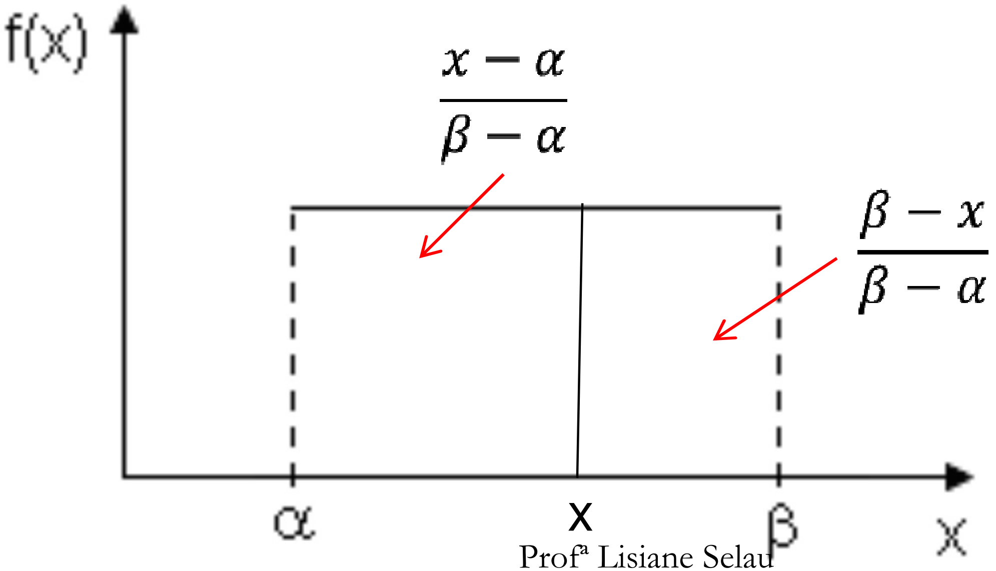


Função densidade de probabilidade

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \text{para } \alpha \leq x \leq \beta \\ 0, & \text{em caso contrário} \end{cases}$$

Função de probabilidade acumulada

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx = \begin{cases} 0, & \text{se } x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha}, & \text{se } \alpha \leq x \leq \beta \\ 1, & \text{se } x > \beta \end{cases}$$



Parâmetros

A distribuição uniforme tem dois parâmetros:

α: menor valor para o qual a variável X está definida

β: maior valor para o qual a variável X está definida

$$X \sim U(\alpha, \beta)$$

Medidas descritivas

- Média ou valor esperado: $$\boxed{E(X) = \mu = \frac{(\beta + \alpha)}{2}}$$

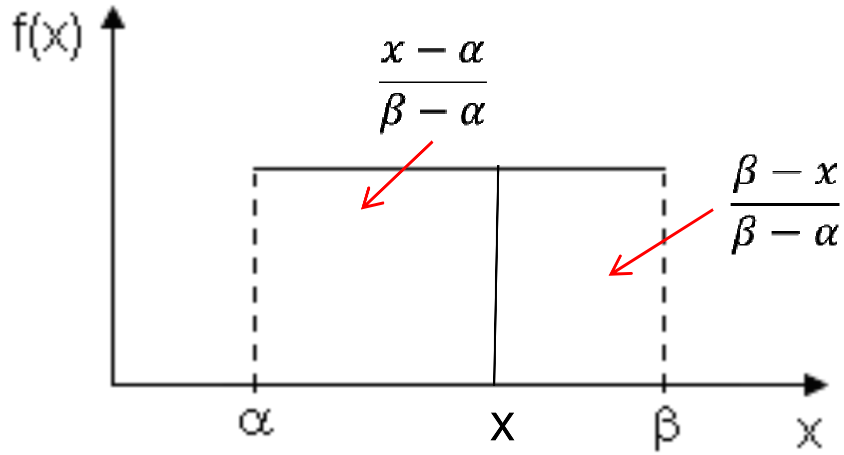
- Variância: $$\boxed{V(X) = \sigma^2 = \frac{(\beta - \alpha)^2}{12}}$$

RESUMO - Distribuição Uniforme

Descrição probabilística de uma variável aleatória contínua que assume valores no intervalo $[\alpha, \beta]$, cuja probabilidade de assumir valores num subintervalo é a mesma que para qualquer outro subintervalo de mesmo comprimento.

Função dens. probabilidade

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \text{para } \alpha \leq x \leq \beta \\ 0, & \text{em caso contrário} \end{cases}$$



Parâmetros

α: menor valor para o qual a variável X está definida

β: maior valor para o qual a variável X está definida

Medidas descritivas

$$E(X) = \mu = \frac{(\beta + \alpha)}{2} \qquad V(X) = \sigma^2 = \frac{(\beta - \alpha)^2}{12}$$

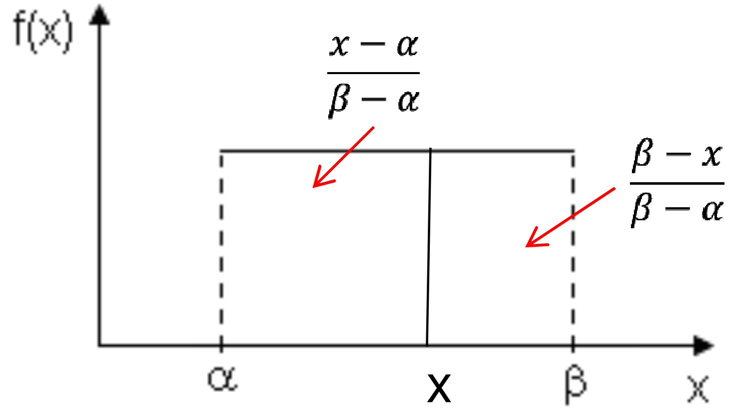
Exemplo: Seja X uma variável aleatória contínua com distribuição uniforme no intervalo [5, 10]. Determinar as probabilidades:

a) P(X < 7)

b) P(X > 8,5)

c) P(8 < x < 9)

$$\boxed{P(X < x) = F(x) = \frac{x-\alpha}{\beta-\alpha}}$$



Utilizando a função de distribuição acumulada:

a) $$P(X < 7) = \frac{x-\alpha}{\beta-\alpha} = \frac{7-5}{10-5} = \frac{2}{5} = 0{,}4$$

b) $$P(X > 8{,}5) = \frac{\beta-x}{\beta-\alpha} = \frac{10-8{,}5}{10-5} = \frac{1{,}5}{5} = 0{,}3$$

c) $$P(8 < X < 9) = P(X < 9) - P(X < 8) = \frac{x_2-\alpha}{\beta-\alpha} - \frac{x_1-\alpha}{\beta-\alpha} = \frac{9-5}{10-5} - \frac{8-5}{10-5} = \frac{1}{5} = 0{,}2$$

Exercício: Uma variável X é uniformemente distribuída no intervalo [10, 20]. Determine:

a) valor esperado e variância de X

$$\boxed{\alpha = 10 \quad \text{e} \quad \beta = 20}$$

$$E(X) = \mu = \frac{(\beta + \alpha)}{2} = \frac{(20 + 10)}{2} = 15$$

$$V(X) = \sigma^2 = \frac{(\beta - \alpha)^2}{12} = \frac{(20 - 10)^2}{12} = 8{,}33$$

b) $P(12{,}31 < X < 16{,}50)$

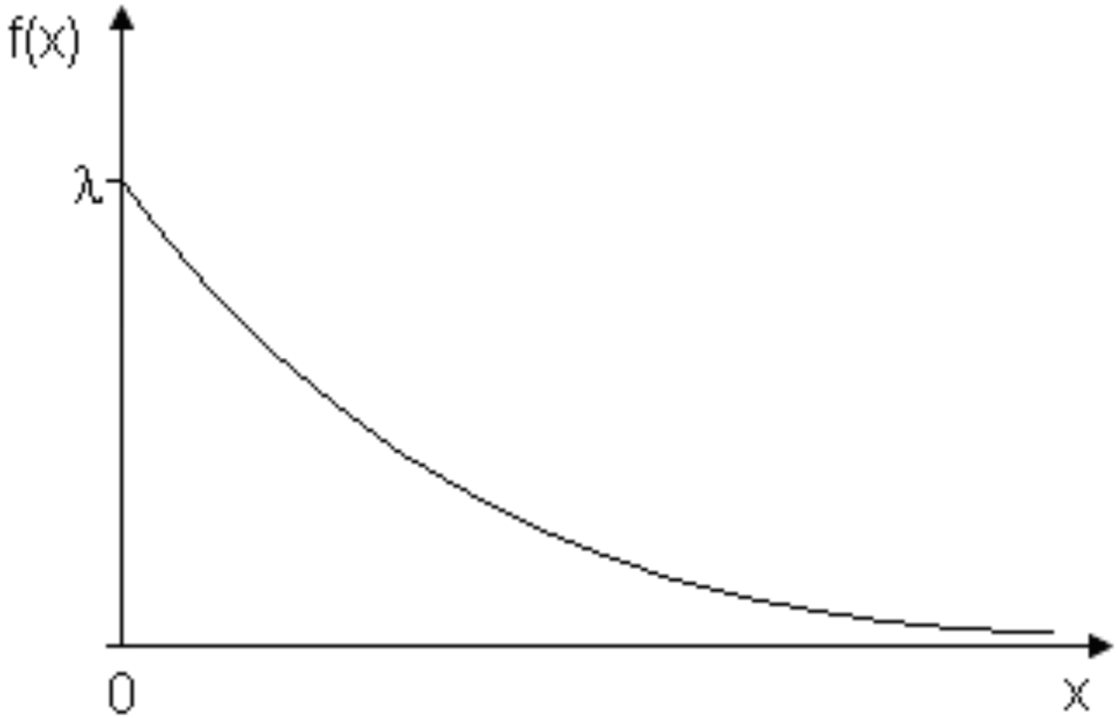
$$\boxed{P(X < x) = F(x) = \frac{x - \alpha}{\beta - \alpha}}$$

$$P(12{,}31 < X < 16{,}50) = F(16{,}50) - F(12{,}31)$$

$$= \frac{16{,}50 - 10}{10 - 10} - \frac{12{,}431 - 10}{20 - 10} = \frac{16{,}50 - 12{,}31}{10} = 0{,}4190$$

2. Distribuição Exponencial

Definição: Seja X uma variável aleatória contínua que só assume valores **não negativos**. Se esta variável é o **tempo** decorrido entre ocorrências sucessivas de um processo de Poisson, então ela tem distribuição exponencial.



Na distribuição de Poisson, a variável aleatória é definida como o número de ocorrências (sucessos) em determinado período de tempo, sendo a média das ocorrências no período definida como λ.

Na distribuição exponencial, a variável aleatória é definida como o tempo entre duas ocorrências, sendo a média de tempo entre ocorrências igual a $1/\lambda$.

Por exemplo, se a média de atendimentos no caixa de uma loja é de $\lambda = 6$ clientes/min, então o tempo médio entre atendimentos é $1/\lambda = 1/6$ de minuto ou 10 segundos.

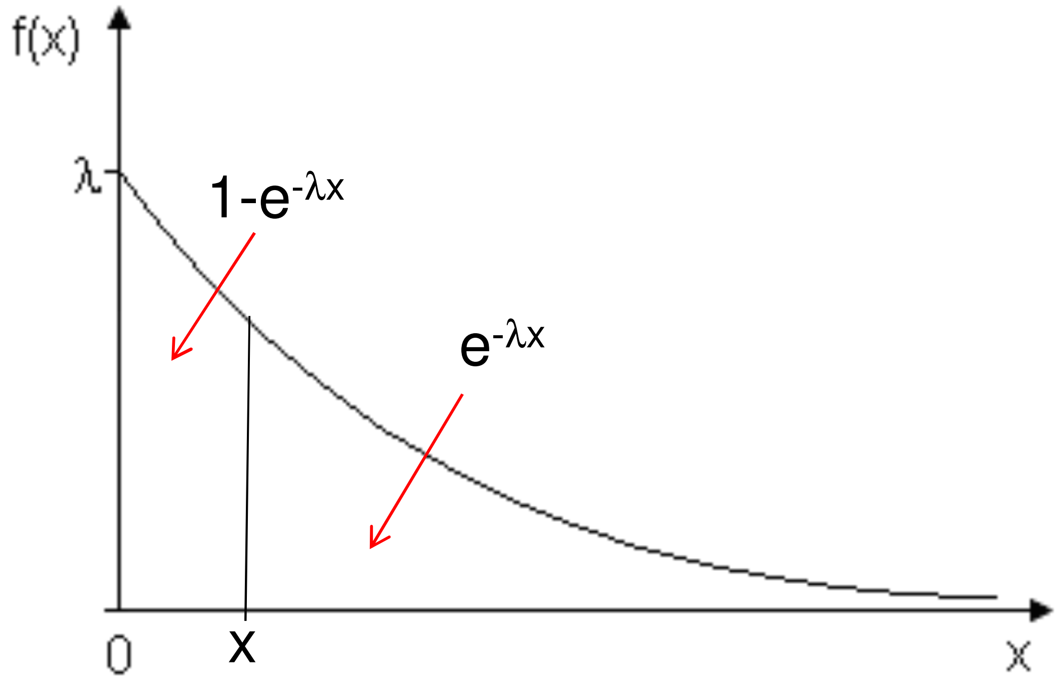
A distribuição exponencial é muito utilizada no campo da confiabilidade para a modelagem do tempo até a ocorrência de falha em componentes eletrônicos, bem como do tempo de espera em sistemas de filas.

Função densidade de probabilidade

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{para } x>0 \\ 0, & \text{em caso contrário} \end{cases}$$

Função de probabilidade acumulada

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx = \begin{cases} 0, & \text{se } x < 0 \\ 1 - e^{-\lambda x}, & \text{se } x \geq 0 \end{cases}$$



Parâmetros

A distribuição exponencial tem apenas um parâmetro:

λ: número médio de ocorrências em determinado período de tempo (λ>0)

$$X \sim Exp(\lambda)$$

Medidas descritivas

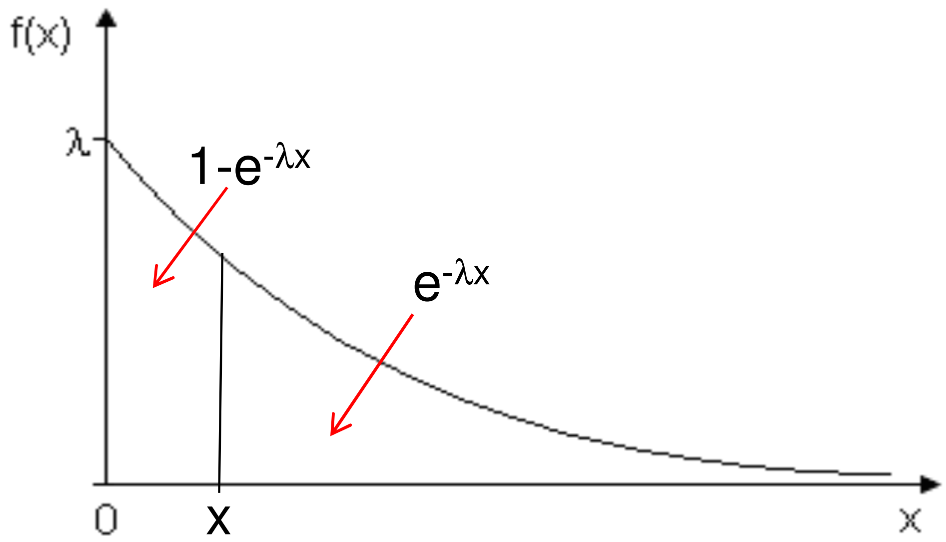
- Média ou valor esperado: $\boxed{E(X) = \mu = \dfrac{1}{\lambda}}$

- Variância: $\boxed{V(X) = \sigma^2 = \dfrac{1}{\lambda^2}}$

RESUMO - Distribuição exponencial

Descrição probabilística de uma variável aleatória contínua que é o tempo decorrido entre ocorrências sucessivas de um processo de Poisson.

Função dens. probabilidade



$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{para } x>0 \\ 0, & \text{em caso contrário} \end{cases}$$

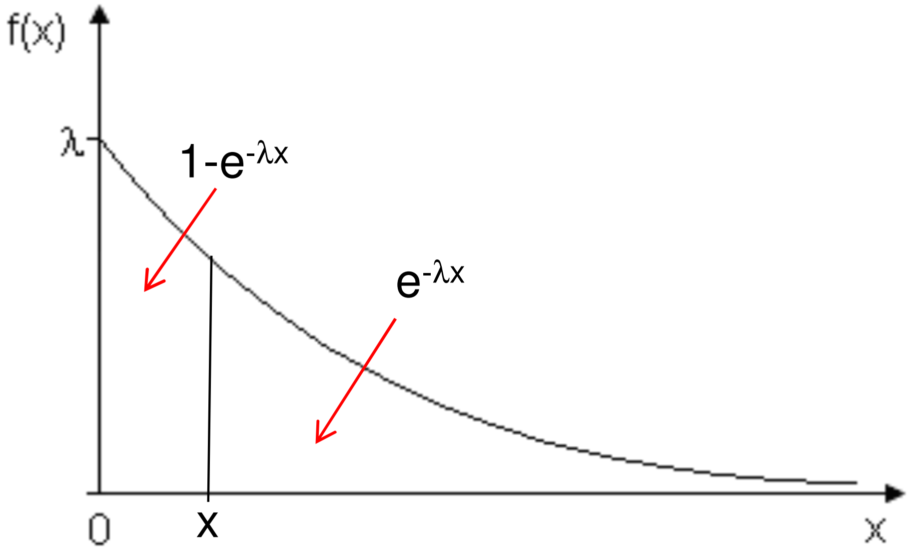
Parâmetros

λ: número médio de ocorrências em determinado período de tempo ($\lambda>0$)

Medidas descritivas

$$E(X) = \mu = \frac{1}{\lambda} \qquad V(X) = \sigma^2 = \frac{1}{\lambda^2}$$

Exemplo: Os tempos até a falha de um dispositivo eletrônico seguem o modelo exponencial, com uma taxa de falha λ = 0,012 falhas/hora. Indique qual a probabilidade de um dispositivo escolhido ao acaso sobreviver a 50 horas? E a 100 horas?



X: tempo até a falha (horas)

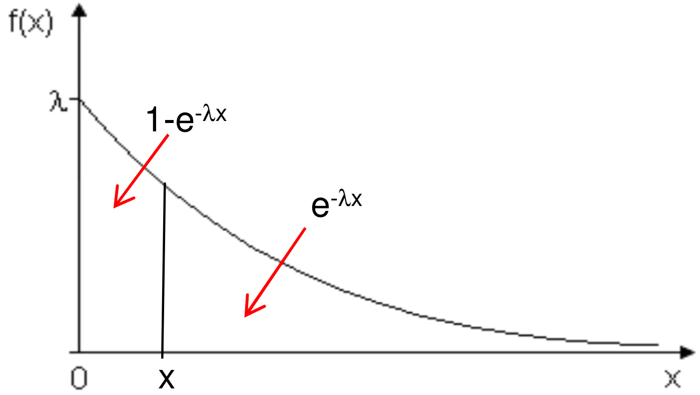
X ~ Exp(0,012)

$$P(X > 50) = e^{-0{,}012 \text{x} 50} = e^{-0{,}6} = 0{,}5488$$

$$P(X > 100) = e^{-0{,}012 \text{x} 100} = e^{-1{,}2} = 0{,}3012$$

Exercício: Suponha que um componente eletrônico tenha um tempo de vida X (em unidades de 1000 horas) que segue uma distribuição exponencial de parâmetro $\lambda = 1$. Suponha que o custo de fabricação do item seja 2 reais e que o preço de venda seja 5 reais. O fabricante garante devolução total se $X < 0{,}90$. Qual o lucro esperado por item?

$$X \sim Exp(1)$$



A probabilidade de um componente durar menos de 900 horas é dada por:

$$P(X < 0{,}9) = F(0{,}9) = 1 - e^{-0{,}9} = 0{,}5934$$

Assim, o lucro do fabricante será uma variável aleatória discreta Y com a seguinte distribuição:

$Y = y$	-2	3	Σ
$P(Y = y)$	0,5934	0,4066	1

Então o lucro esperado será:

$$E(Y) = -2 \times 0{,}5934 + 3 \times 0{,}4066 = R\$\ 0{,}03$$

3. Distribuição Normal

É importante tanto no aspecto teórico como nas aplicações. Essa importância se deve a um conjunto de aspectos:

⇨ É útil para descrever uma grande quantidade de fenômenos naturais físicos, ambientais, etc.

⇨ Muitas variáveis não normais podem ser aproximadas como normais após transformações simples.

⇨ Propriedades matemáticas.

⇨ Distribuições de um grande número de variáveis aleatórias convergem para a distribuição normal.

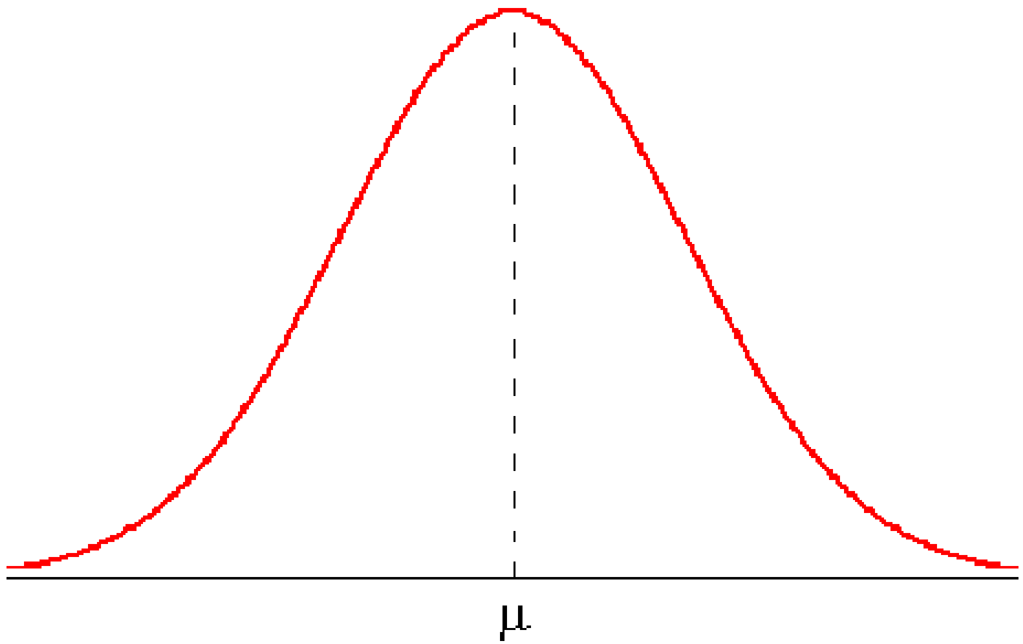
⇨ Uma grande quantidade de métodos e procedimentos de inferência estatística são derivados tendo-a como pressuposição básica.

O conjunto de métodos desenvolvidos para tratar variáveis que têm distribuição normal forma a chamada **Estatística Clássica** ou **Estatística Paramétrica**.

Distribuição normal

Definição: É uma distribuição teórica de frequências, onde a maioria das observações se situa em torno da **média** (centro) e diminui gradual e simetricamente no sentido dos extremos.

A distribuição normal é representada graficamente pela curva normal (curva de Gauss) que tem a forma de sino e é **simétrica** em relação ao centro, onde se localiza a média μ.



Função densidade de probabilidade

De modo geral, se X é uma variável contínua que tem distribuição normal, sua função densidade de probabilidade será:

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

parâmetros

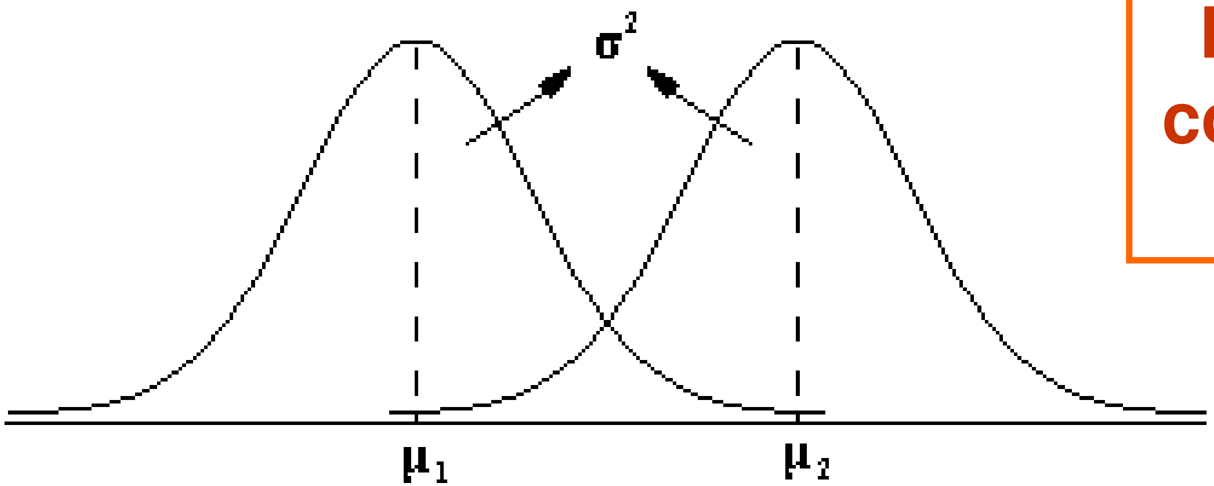
Parâmetros

A distribuição normal tem dois parâmetros:

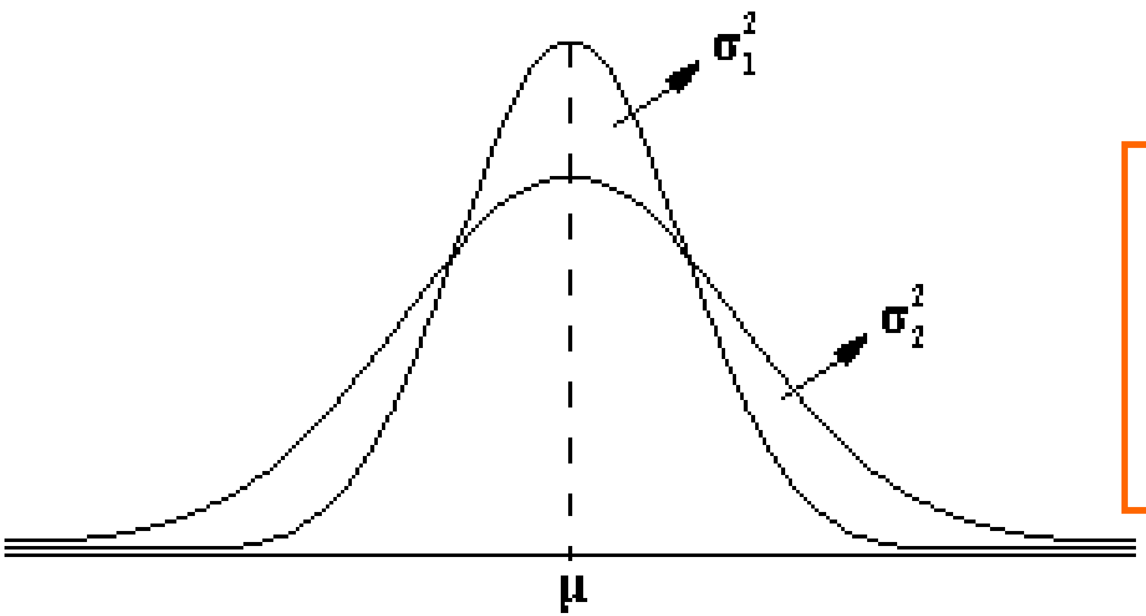
μ = média (determina o centro da distribuição)
σ^2 = variância (determina a dispersão da distribuição)

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

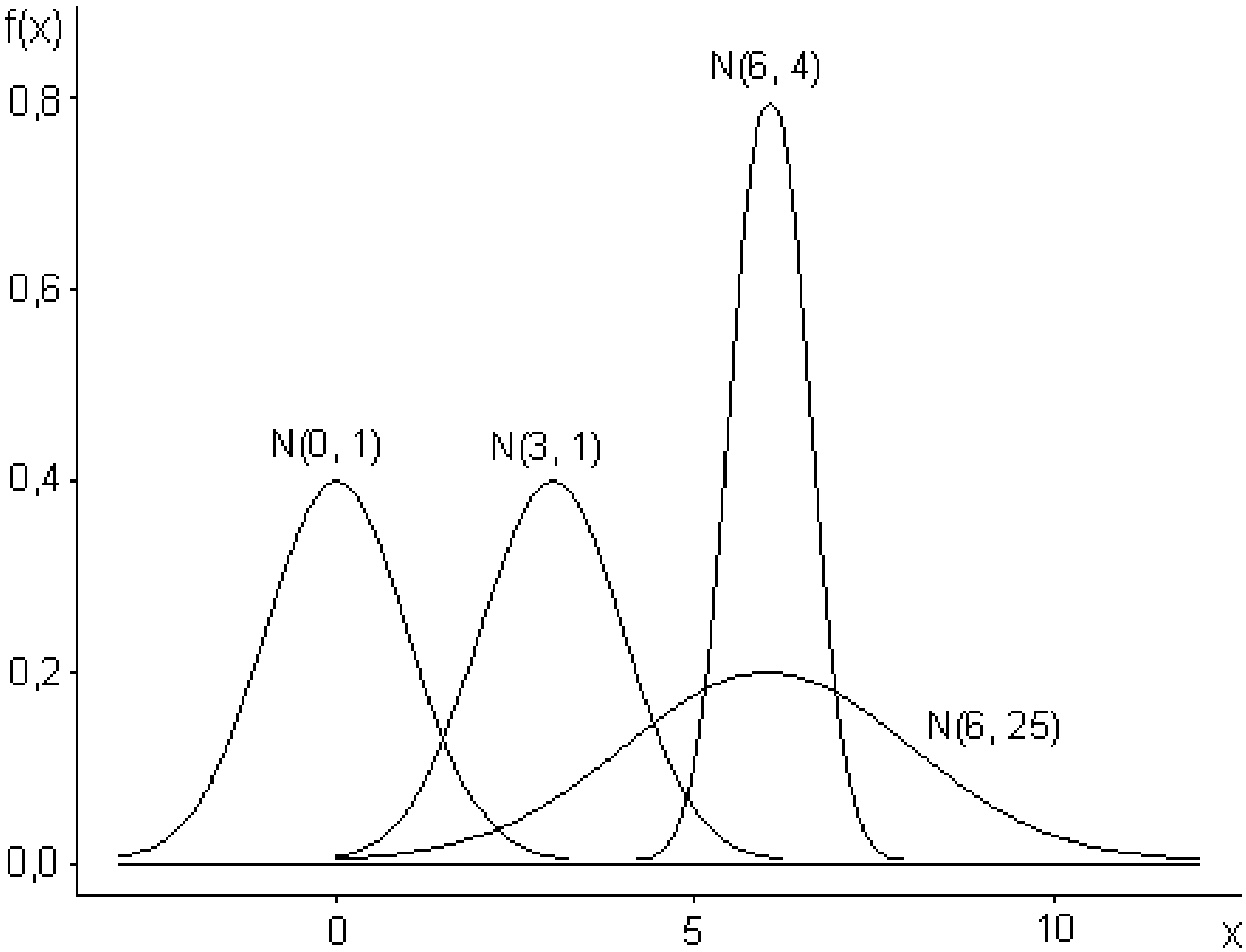
X tem distribuição normal com parâmetros μ e σ^2



Populações normais com médias diferentes e mesma variância



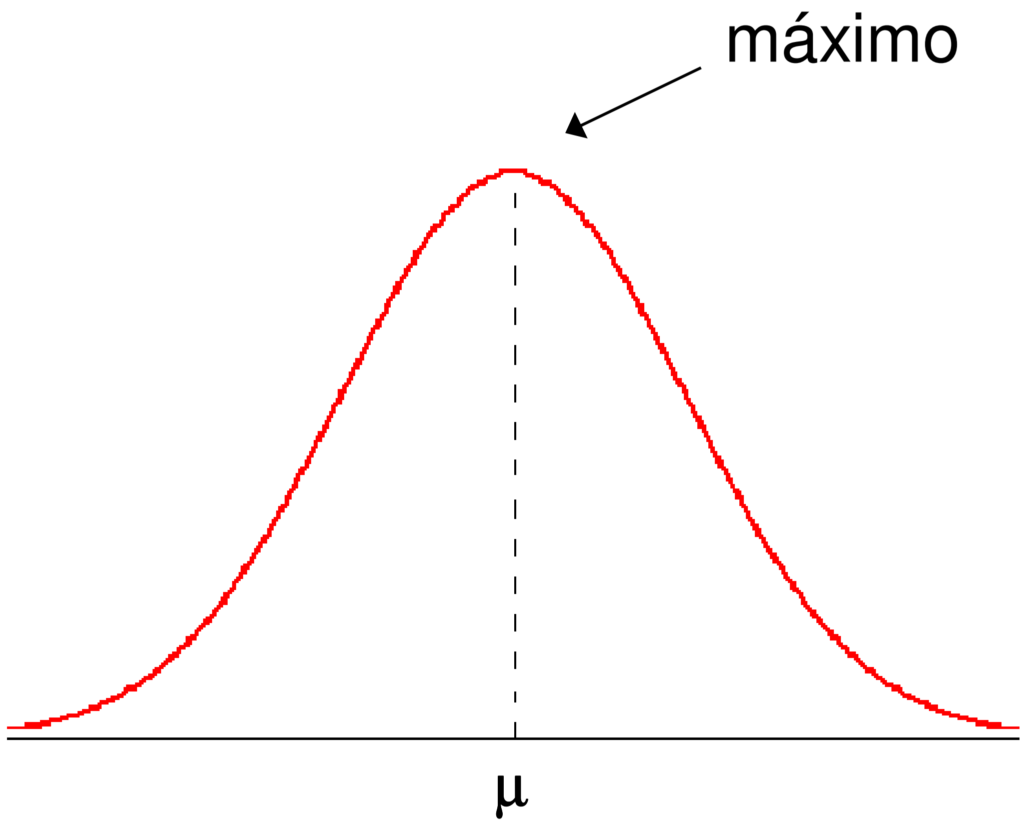
Populações normais com variâncias diferentes e mesma média



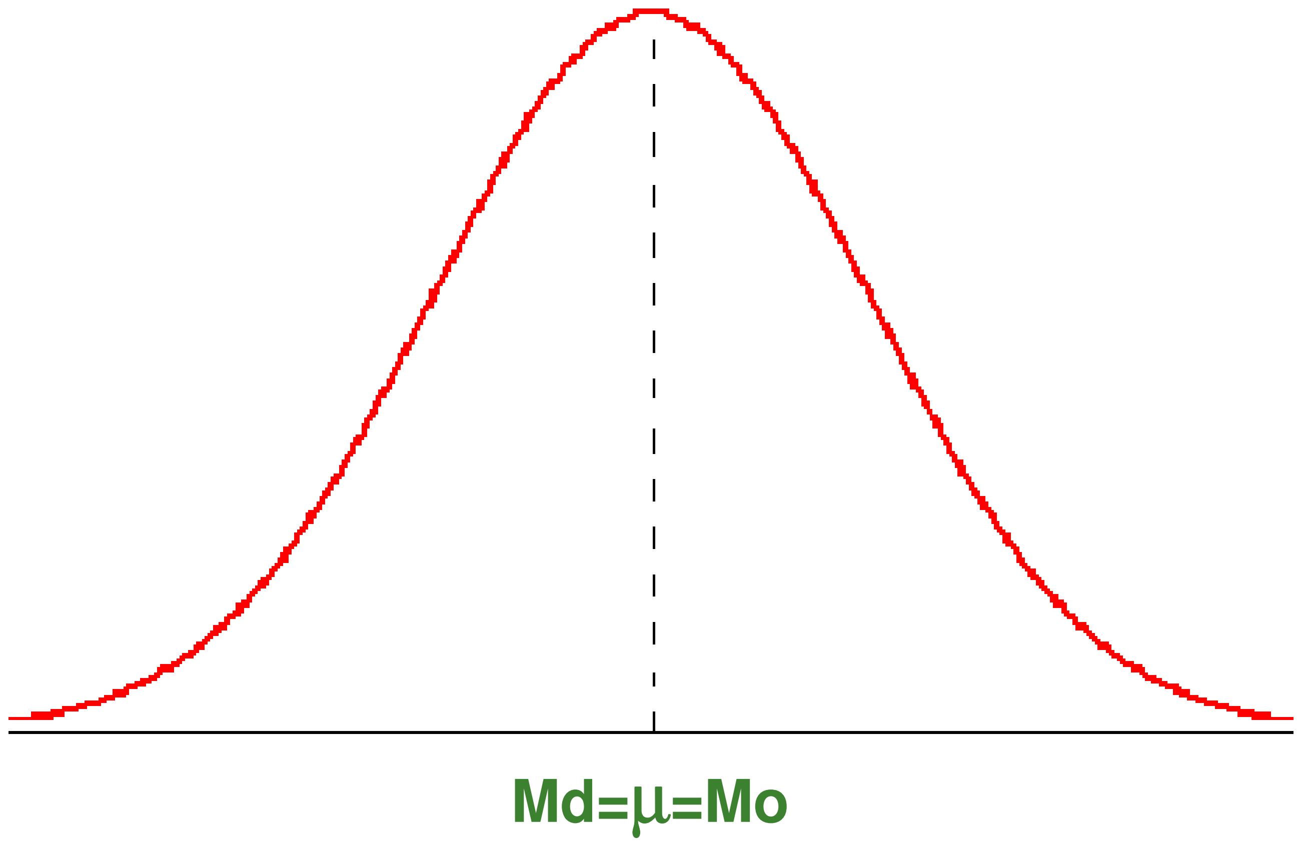
Existe um número infinito de curvas normais

Propriedades da distribuição normal

1. O máximo da função densidade de probabilidade se dá no ponto **$x=\mu$** .

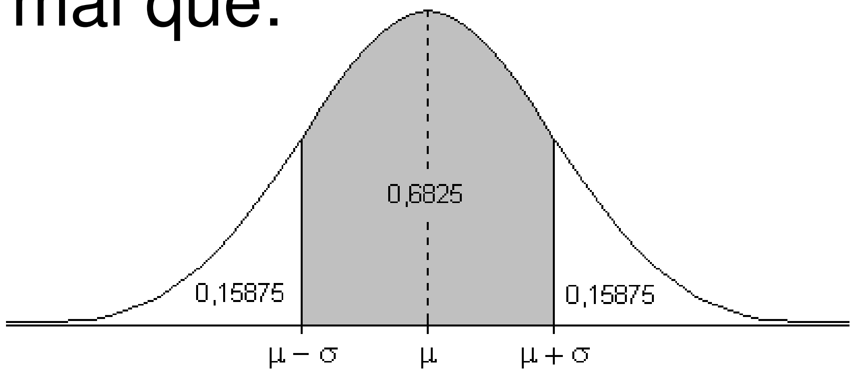


2. A distribuição é simétrica em relação ao centro onde coincidem a média, a moda e a mediana.

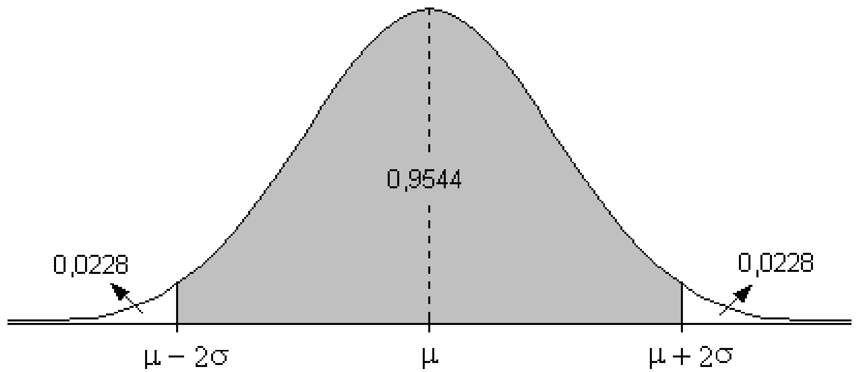


3. Verifica-se na distribuição normal que:

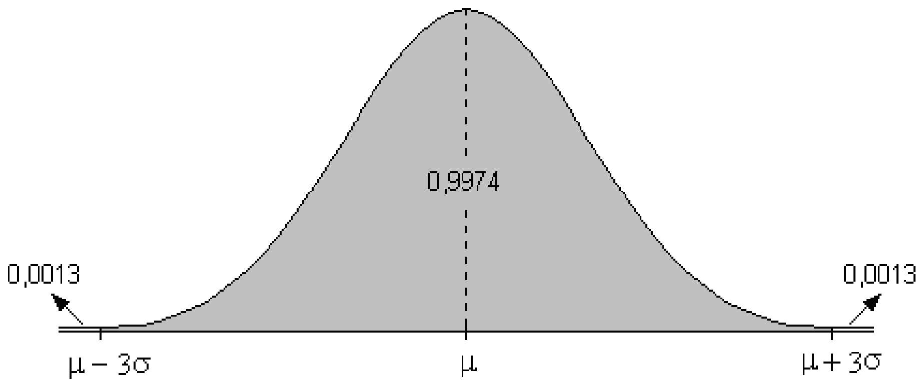
$P(\mu-\sigma < X < \mu+\sigma) = 0,6825$



$P(\mu-2\sigma < X < \mu+2\sigma) = 0,9544$

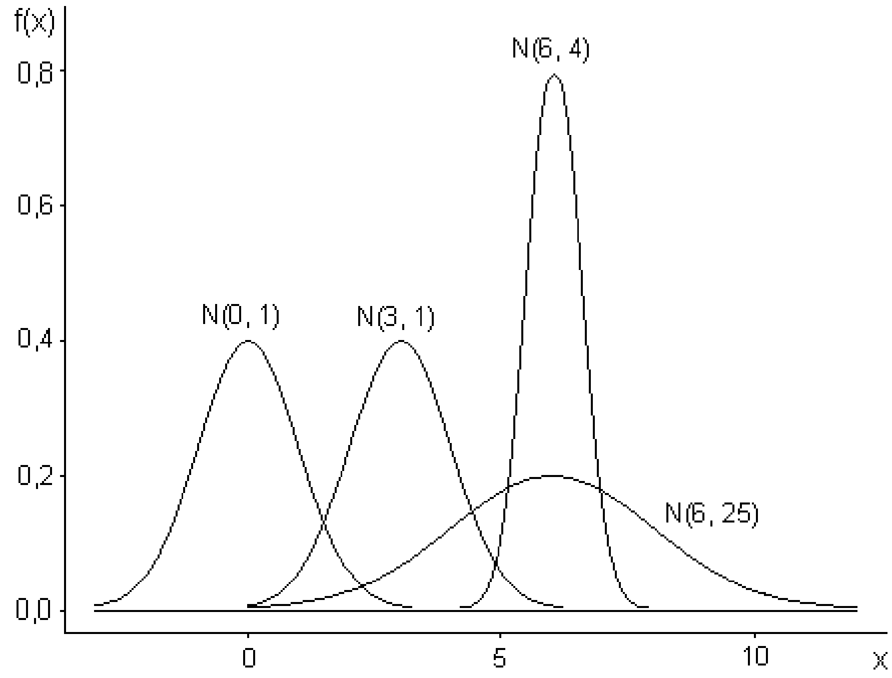


$P(\mu-3\sigma < X < \mu+3\sigma) = 0,9974$



Cálculo de áreas

⇨ Para cada valor de μ e de σ, existe uma distribuição normal diferente



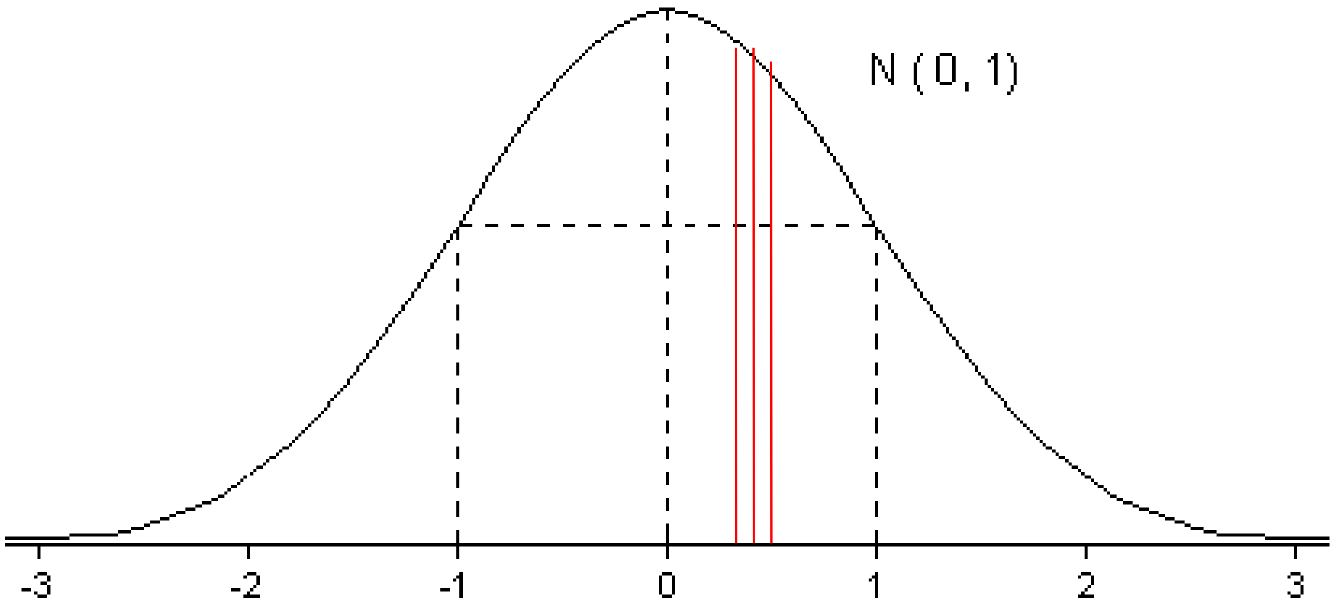
⇨ O cálculo de áreas sob a curva normal, deverá ser feito sempre em função dos valores particulares de μ e σ

⇨ Para evitar a trabalhosa tarefa de calcular as áreas foi determinada uma distribuição normal padrão ou reduzida

⇨ As áreas sob a distribuição normal padrão foram calculadas e apresentadas numa tabela

Distribuição normal padrão

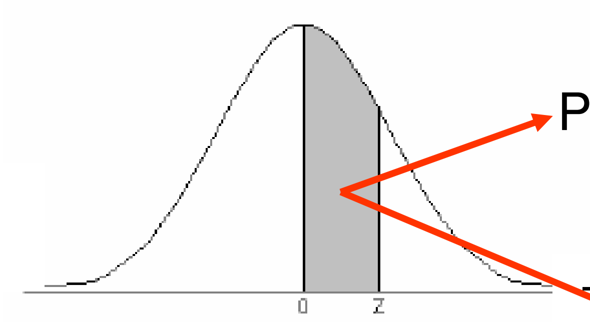
Definição: é a distribuição normal de uma variável Z que tem média igual a zero (**μ=0**) e desvio padrão igual a um (**σ=1**).



A curva normal padrão foi dividida em pequenas tiras, cujas áreas foram calculadas e apresentadas numa tabela.

Na tabela da distribuição normal padrão, podemos encontrar as áreas correspondentes aos intervalos de 0 a z.

Tabela - Área sob a curva normal padrão de 0 a z, $P(0 \le Z \le z)$.



$P(0 < Z < z)$

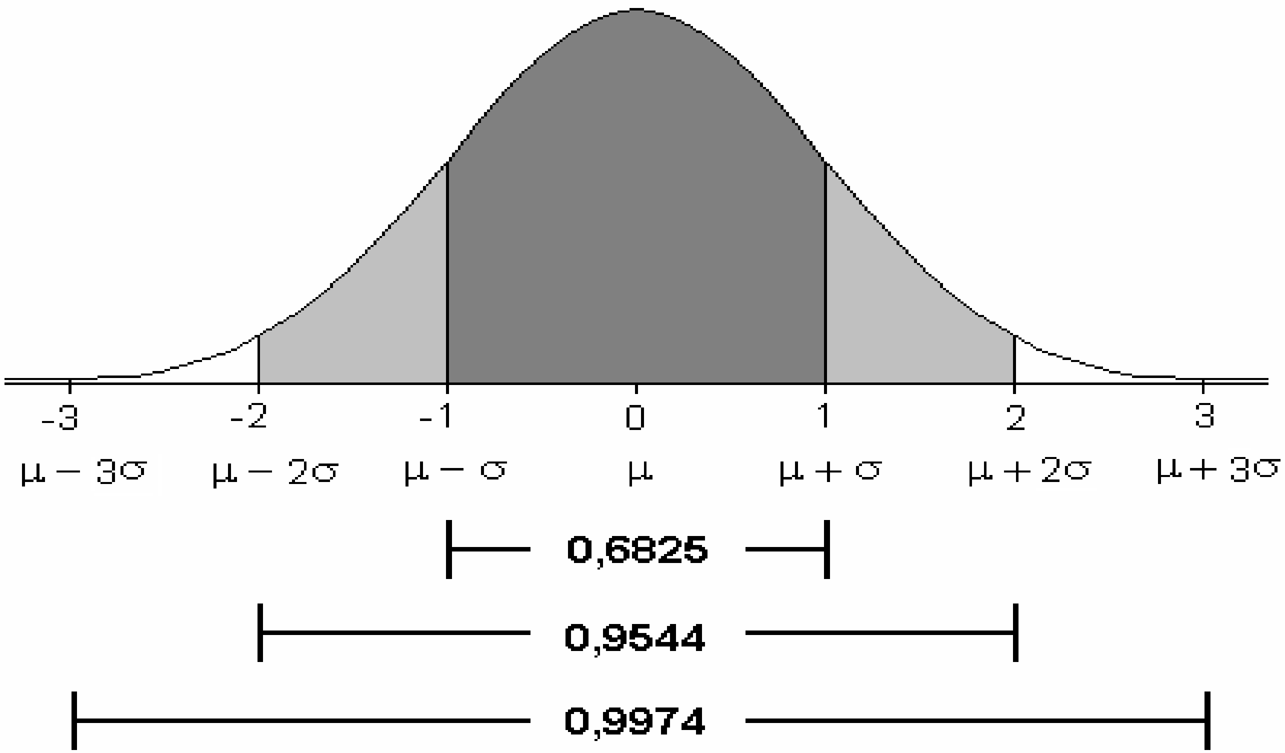
$P(0 < Z < 0{,}62) = ?$

$P(0 < Z < 0{,}62) = 0{,}2324$

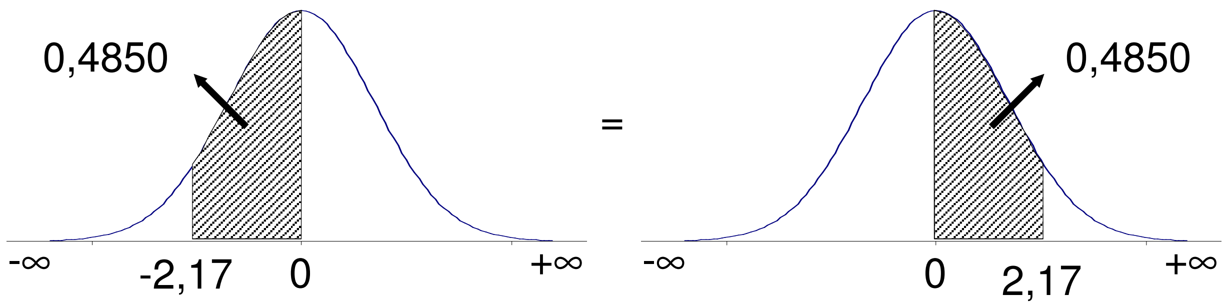
Z	0	1	2	3	4
0,0	0,0000	0,0040	0,0080	0,0120	0,0160
0,1	0,0398	0,0438	0,0478	0,0517	0,0557
0,2	0,0793	0,0832	0,0871	0,0910	0,0948
0,3	0,1179	0,1217	0,1255	0,1293	0,1331
0,4	0,1554	0,1591	0,1628	0,1664	0,1700
0,5	0,1915	0,1950	0,1985	0,2019	0,2054
0,6	0,2258	0,2291	$\boxed{0,2324}$	0,2357	0,2389
0,7	0,2580	0,2612	0,2642	0,2673	0,2704
0,8	0,2881	0,2910	0,2939	0,2967	0,2996
0,9	0,3159	0,3186	0,3212	0,3238	0,3264
1,0	0,3413	0,3438	0,3461	0,3485	0,3508
1,1	0,3643	0,3665	0,3686	0,3708	0,3729

⇨ Os valores negativos não são apresentados na tabela porque a curva é simétrica; assim, as áreas correspondentes a esses valores são exatamente iguais às dos seus simétricos positivos, por exemplo **$P(-1<Z<0)=P(0<Z<1)$**.

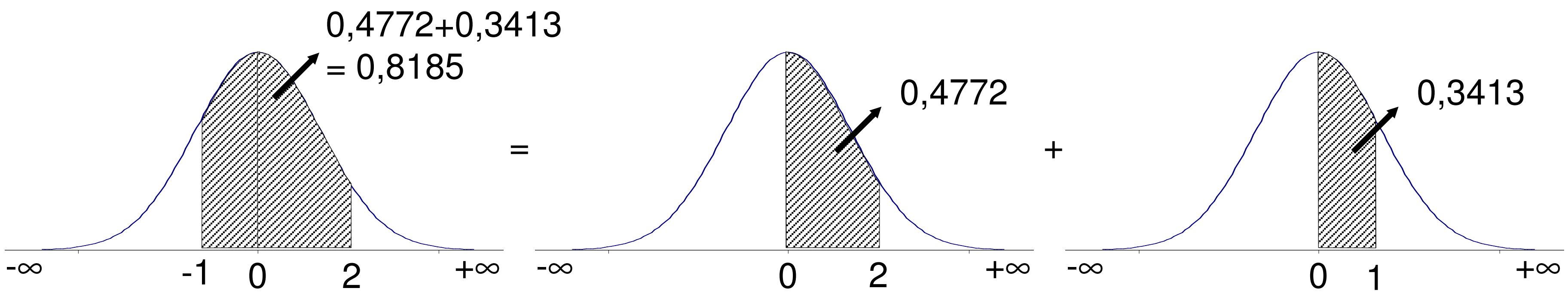
⇨ Na tabela da distribuição normal padrão, os valores de Z vão de 0 a 3,99 . Este limite é estabelecido com base nas propriedades da distribuição normal.



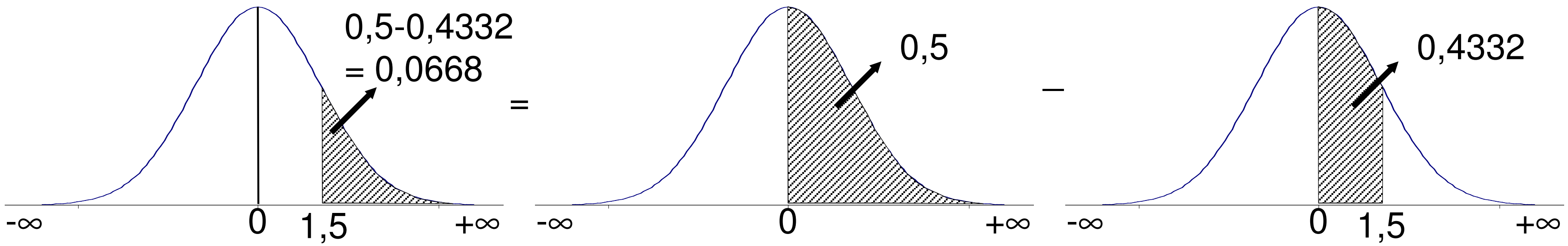
$P(-2{,}17 < Z < 0) = ?$



$P(-1 < Z < 2) = ?$



$P(Z > 1{,}5) = ?$



Exercício proposto:

Seja Z uma N(0,1). Determinar as seguintes probabilidades:

(a) $P(Z < 2,23)$

(b) $P(Z > -1,45)$

(c) $P(-2 < Z \leq 2)$

(d) $P(-1 \leq Z \leq 1)$

(e) $P(0 < Z < 1,73)$

(f) $P(Z > 0,81)$

(g) $P(-1,25 \leq Z \leq -0,63)$

(a) 0,9871

(b) 0,9265

(c) 0,9544

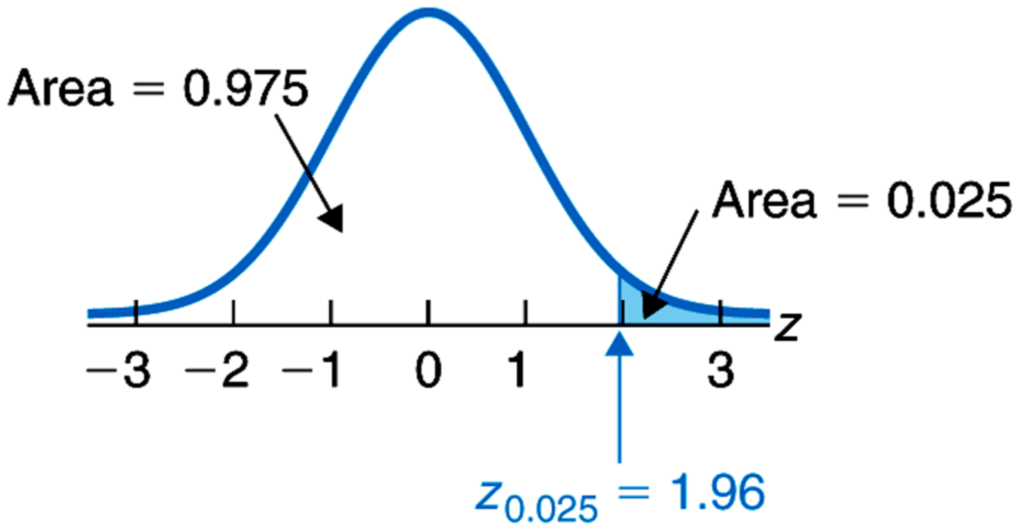
(d) 0,6826

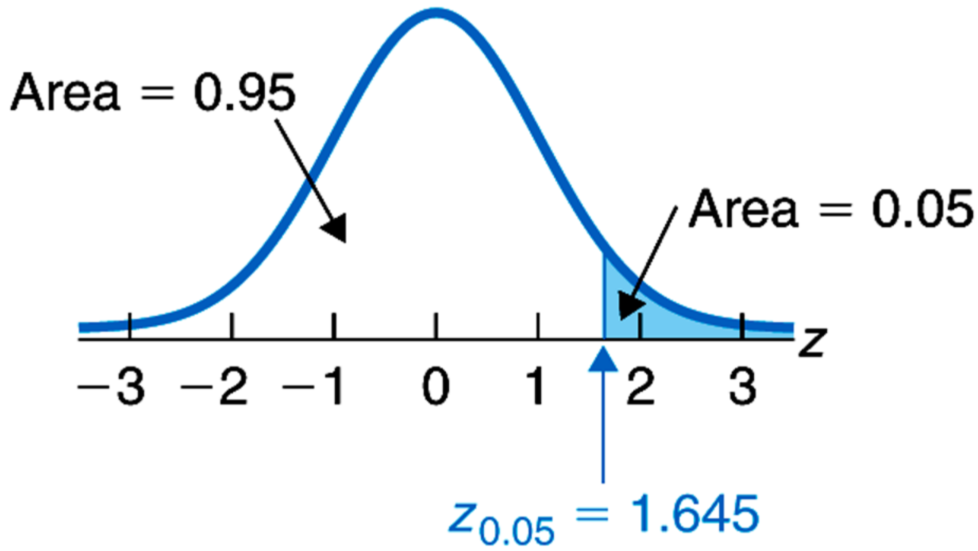
(e) 0,4582

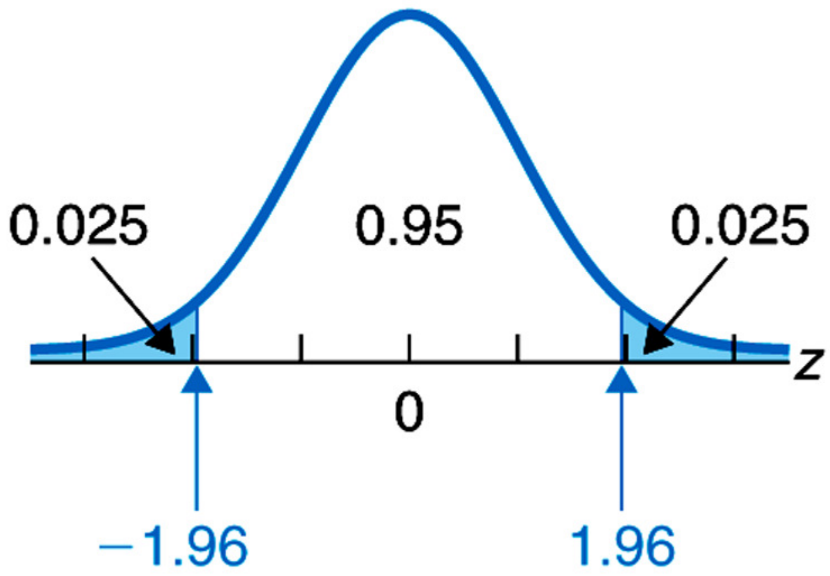
(f) 0,2090

(g) 0,1587

Alguns valores importantes







⇨ Através da distribuição normal padrão é possível estudar qualquer variável X que tenha distribuição normal, com quaisquer valores para μ e σ.

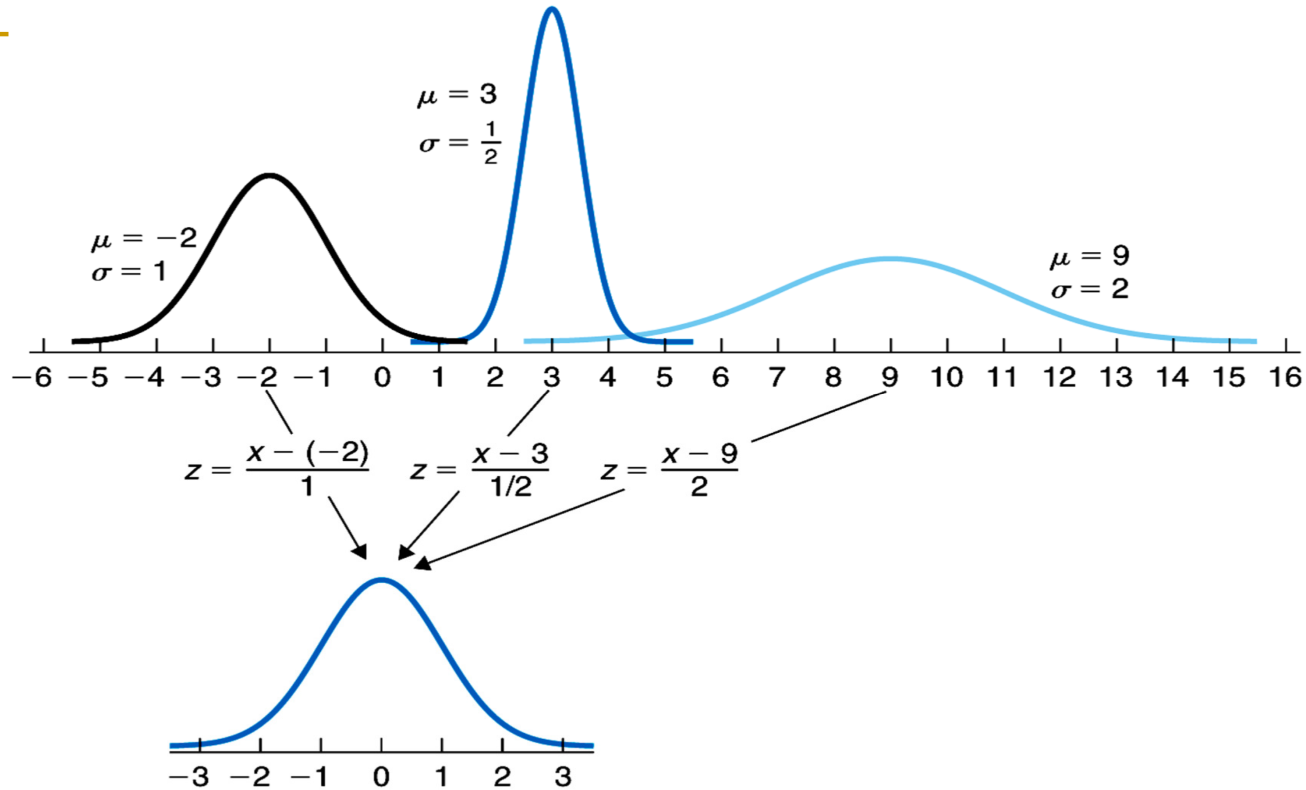
⇨ Para utilizarmos os valores da tabela, devemos **padronizar** a variável X, ou seja, transformar X em Z.

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

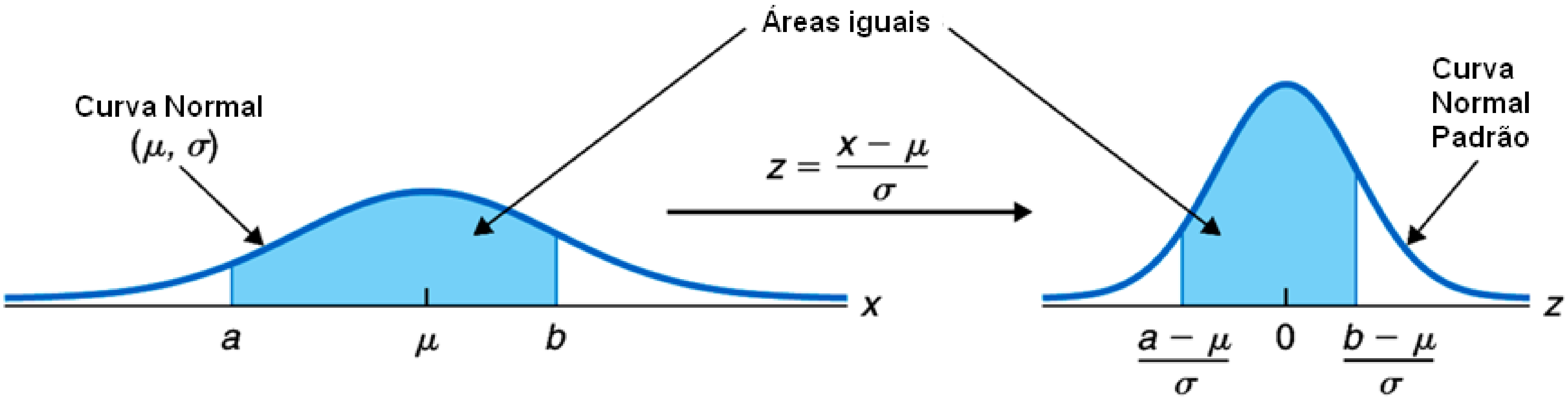
transformar → $Z = \dfrac{X - \mu}{\sigma}$

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

⇨ Após a transformação, procuramos na tabela a área compreendida entre **0** e **z**, que corresponderá a área entre μ e **x**.



A transformação muda as variáveis, mas não altera a área sob a curva.

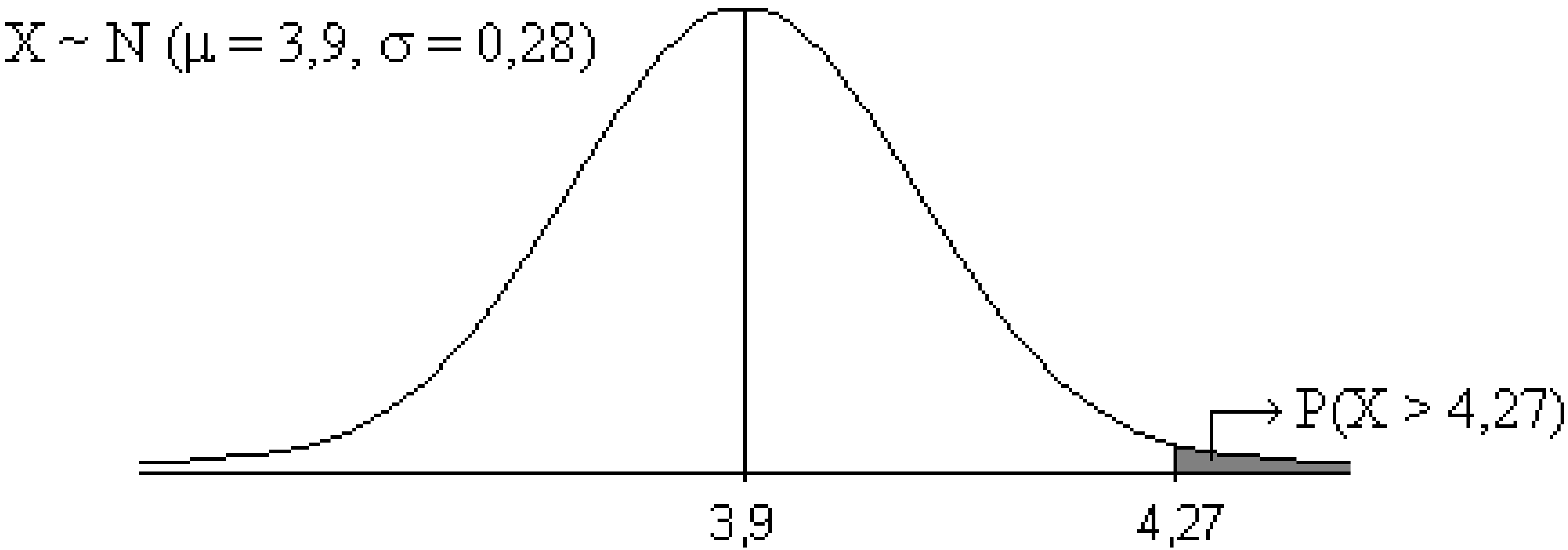


Exemplo:

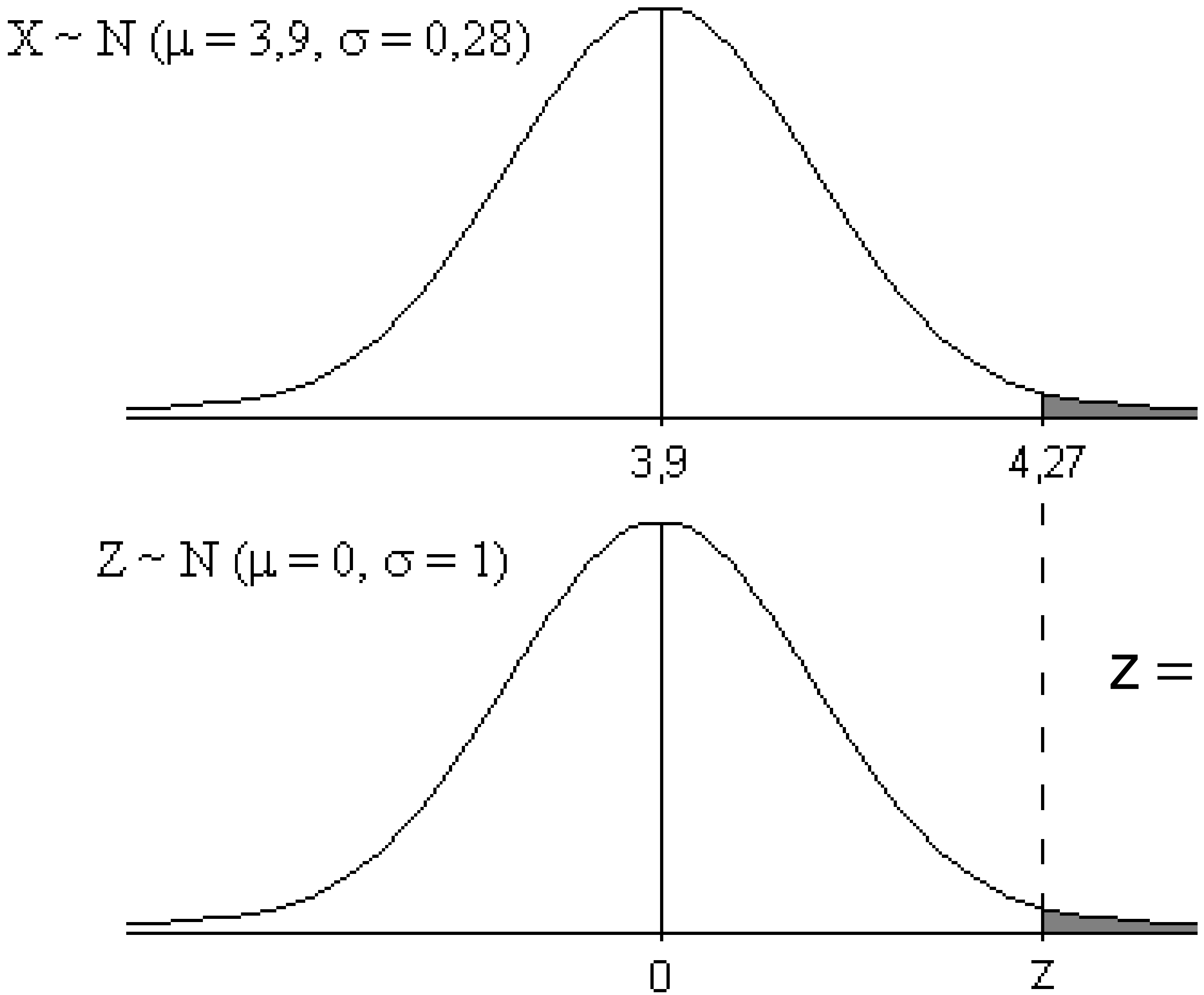
Sabendo que as notas de 450 alunos estão normalmente distribuídas, com média μ = 3,9 e desvio padrão σ = 0,28, determine:

a) a probabilidade de um aluno ter nota maior que 4,27;

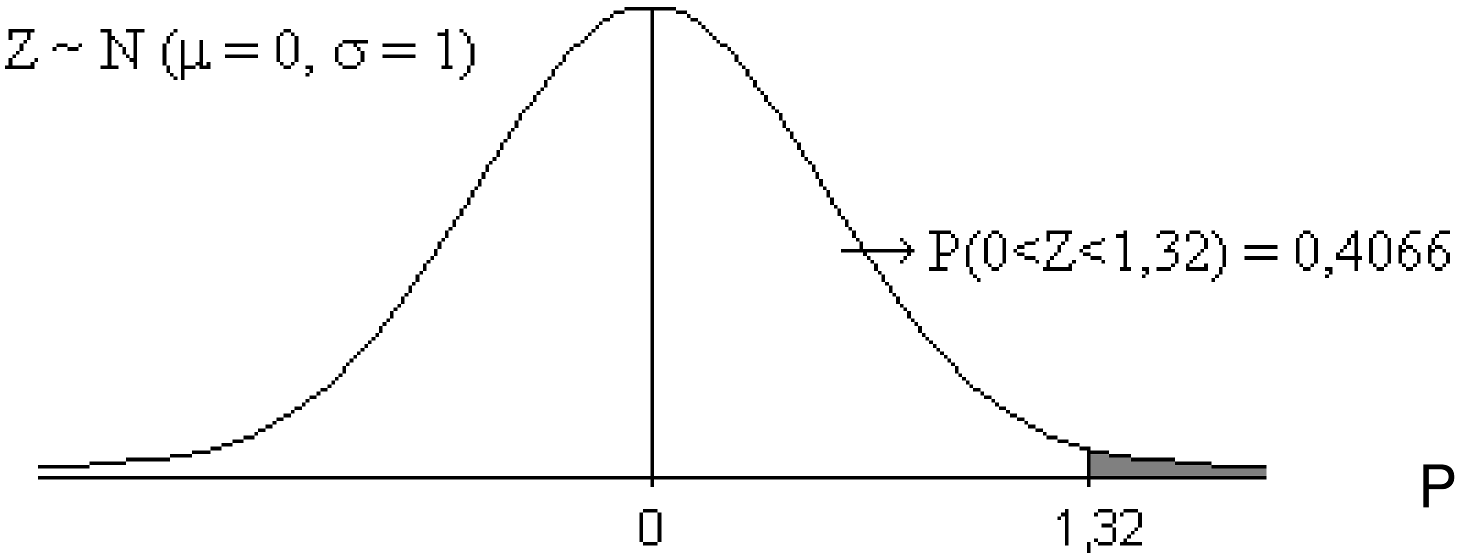
b) o número de alunos que têm nota superior a 4,27.



Para encontrar essa área, vamos utilizar a tabela da distribuição normal padrão. Inicialmente, fazemos a transformação da variável X para a variável Z.



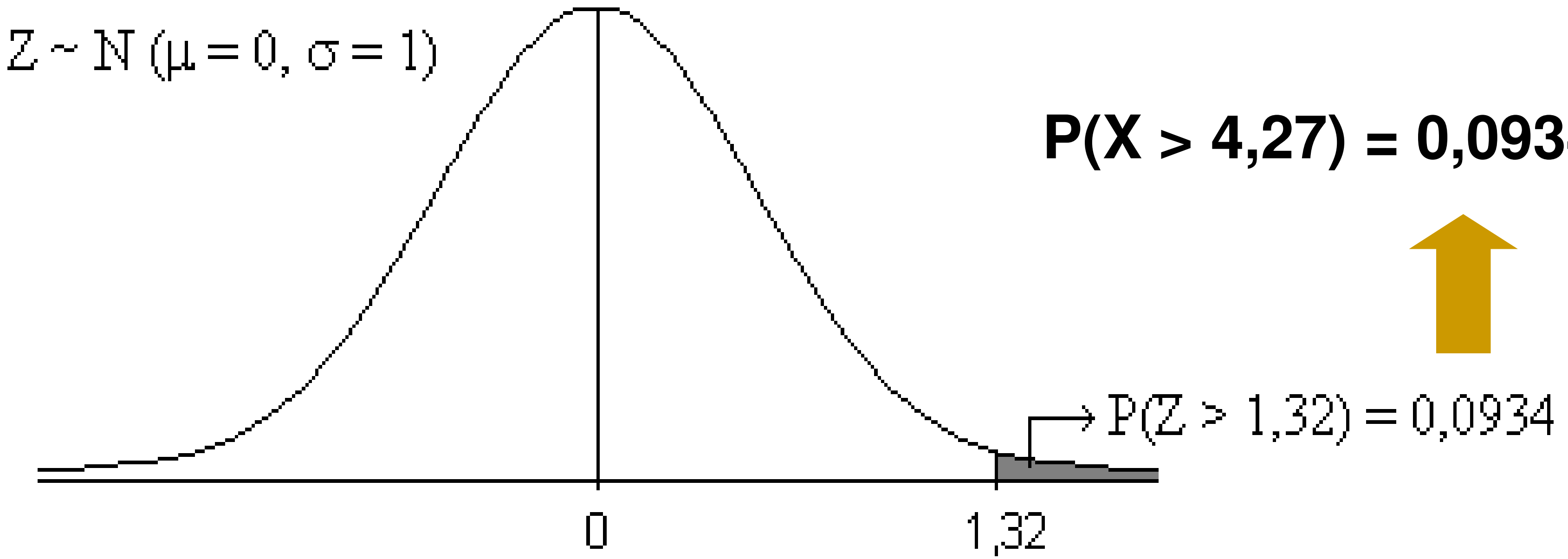
$$z = \frac{4,27 - 3,9}{0,28} = 1,32$$



P(Z > 1,32)

= P(Z > 0) – P(0<Z<1,32)

= 0,5 – 0,4066 = 0,0934



P(X > 4,27) = 0,0934

b) o número de alunos que têm nota superior a 4,27.

No item (a), vimos que este percentual é de 9,34%. Sendo assim, através de uma regra de três simples, podemos determinar quantos estudantes correspondem a 9,34% de uma população de 450 estudantes.

Esse valor pode ser obtido facilmente multiplicando o tamanho da população pela probabilidade de ocorrer uma nota maior que 4,27.

Assim, temos:

$$450 \times 0{,}0934 = 42{,}03$$

Concluímos, então, que, dos 450 estudantes, 42 têm nota superior a 4,27.

Exemplo: Uma fábrica de carros sabe que os motores de sua fábrica tem duração normal com média de 150.000km e desvio padrão de 5.000km. Qual a probabilidade de que um carro escolhido ao acaso, tenha um motor que dure:

a) entre 140.000 e 160.000km?

b) menos de 170.000km?

c) Se o fabricante deseja oferecer uma garantia, tal que ele tenha que substituir no máximo 1% dos motores, qual deve ser o valor desta garantia?

Solução:

a) P(140.000<X<160.000) = P(-2<Z<2) =

= 0,4772+0,4772 = 0,9544 = 95,44%

b) P(X < 170.000) = P(Z < (170.000-150.000)/5.000)

= P(Z < 4) = 1 = 100%

c) Seja G o valor de garantia

P(X<G) = P(Z < (G-150)/5) = 0,01 = P(Z < -2,33)

$\frac{G-150}{5} = -2,33$ ⇒ Garantia tem que ser de 138.350km.

Exercício: Suponha que a estatura de recém-nascidos do sexo feminino é uma variável com distribuição normal de média μ = 48 cm e σ = 3 cm. Determine:

a) a probabilidade de um recém-nascido ter estatura entre 42 e 49 cm;

b) a probabilidade de um recém-nascido ter estatura superior a 52 cm;

c) o número que de recém-nascidos que têm estatura inferior à $\mu+\sigma$cm, dentre os 532 que nasceram numa determinada maternidade, no período de um mês.

a) 0,6065 **b) 0,0918** **c) 448**

Exercício:

O consumo de gasolina por Km rodado para certo tipo de carro tem distribuição normal com média de 100 ml com desvio padrão de 5 ml.

a) calcular a probabilidade de um carro consumir entre 92 e 106 ml.

b) sabe-se que 73,24% dos carros consumem menos que certa quantidade de gasolina qual é essa quantidade?

c) num grupo de 5 carros qual a probabilidade de dois consumirem mais que 107 ml?

a) 0,8301 **b) 103,1** **c) 0,0507**

Exercício: O diâmetro do eixo principal de um disco rígido segue a distribuição Normal com média 25,08 in e desvio padrão 0,05 in.

Se as especificações para esse eixo são 25,00 ± 0,15 in, determine o percentual de unidades produzidas em conformidades com as especificações.

$$P\{24,85 \le x \le 25,15\} = P\{x \le 25,15\} - P\{x \le 24,85\}$$

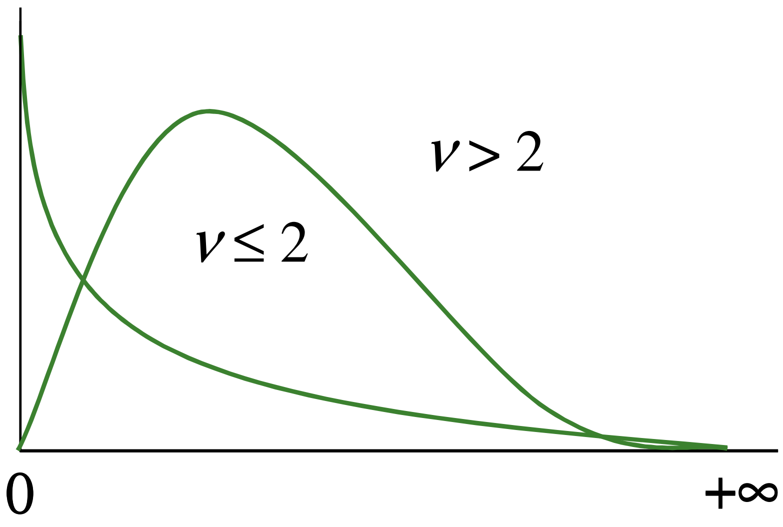
$$= P\left\{Z \le \frac{25,15 - 25,08}{0,05}\right\} - P\left\{Z \le \frac{24,85 - 25,08}{0,05}\right\}$$

$$= P\{Z \le 1,40\} - P\{Z \le -4,60\} = 0,9192 - 0,0000 = 0,9192$$

ou seja, 91,92% dentro das especificações e 8,08% fora das especificações.

4. Distribuição χ^2 (qui-quadrado)

$$f(x) = \frac{1}{2^{\nu/2}\,\Gamma(\nu/2)}\, x^{\frac{\nu}{2}-1} e^{\frac{x}{2}} \qquad x \geq 0$$



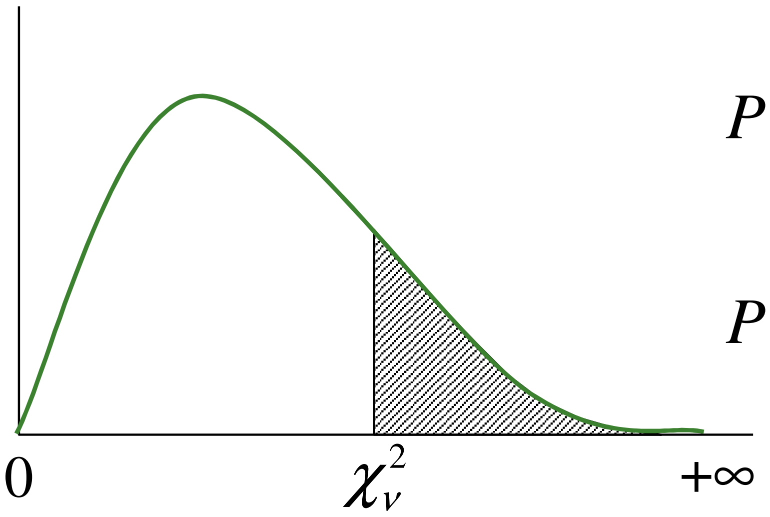
$$\left.\begin{array}{l} E(X) = \nu \\ V(X) = 2\nu \end{array}\right\} X \sim \chi^2_\nu$$

(lê-se: X tem distribuição qui-quadrado com ν graus de liberdade)

Propriedades:

a) se $Z \sim N(0,1)$ então $Z^2 \sim \chi^2_1$

b) se $X_i \sim \chi^2_1$ então $\sum_{i=1}^{n} X_i \sim \chi^2_n$



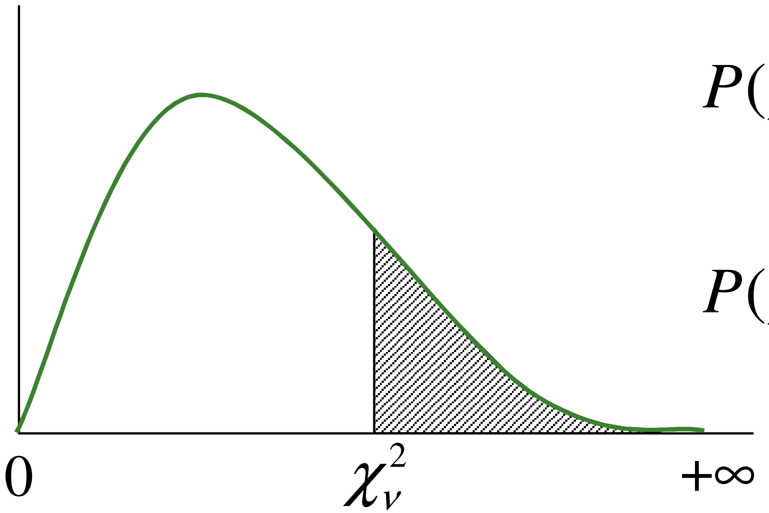
$$P(\chi^2_{10} < 3{,}25) = ?$$

$$P(\chi^2_{10} < 3{,}25) = 0{,}025$$

$$P(\chi^2_{10} > 23{,}21) = ?$$

$$P(\chi^2_{10} > 23{,}21) = 0{,}01$$

Graus de Liberdade	Nível de Significância (α)									
	Esquerda (q')					Direita (q)				
(ν)	0,005	0,01	$\boxed{0{,}025}$	0,05	0,1	0,1	0,05	0,025	$\boxed{0{,}01}$	0,005
1	0,00	0,00	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
2	0,01	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
3	0,07	0,11	0[illegible]2	0,35	0,58	6,25	7,81	9,35	11[illegible]34	12,84
4	0,21	0,30	0[illegible]8	0,71	1,06	7,78	9,49	11,14	13[illegible]28	14,86
5	0,41	0,55	0[illegible]3	1,15	1,61	9,24	11,07	12,83	15[illegible]09	16,75
6	0,68	0,87	1[illegible]4	1,64	2,20	10,64	12,59	14,45	16[illegible]81	18,55
7	0,99	1,24	1[illegible]9	2,17	2,83	12,02	14,07	16,01	18[illegible]48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	$\boxed{3{,}25}$	3,94	4,87	15,99	18,31	20,48	$\boxed{23{,}21}$	25,19



$$P(\chi^2_{10} < 3{,}25) = ?$$

$$P(\chi^2_{10} < 3{,}25) = 0{,}025$$

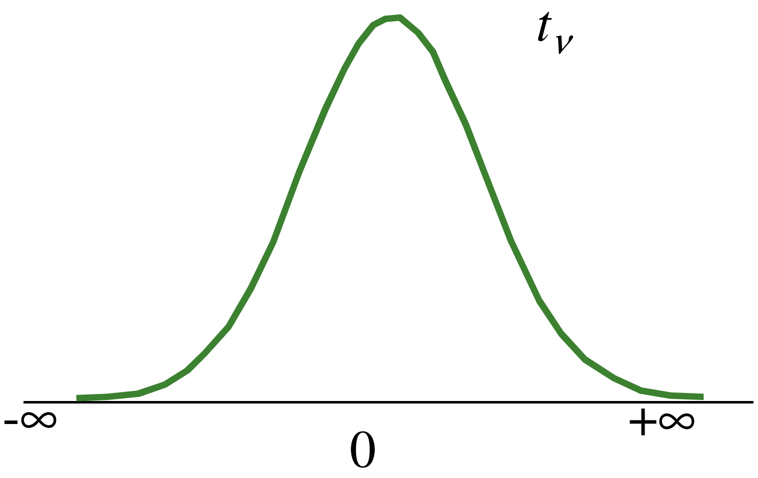
$$P(\chi^2_{10} > ?) = 0{,}025$$

$$P(\chi^2_{10} > 20{,}48) = 0{,}025$$

Graus de Liberdade	Nível de Significância (α)									
	Esquerda (q')					Direita (q)				
(ν)	0,005	0,01	0,025	0,05	0,1	0,1	0,05	0,025	0,01	0,005
1	0,00	0,00	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
2	0,01	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
3	0,07	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84
4	0,21	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86
5	0,41	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	0,68	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19

5. Distribuição *t-student*

$$f(x) = \frac{\Gamma[((\nu+1)/2)]}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} \qquad -\infty < x < \infty$$



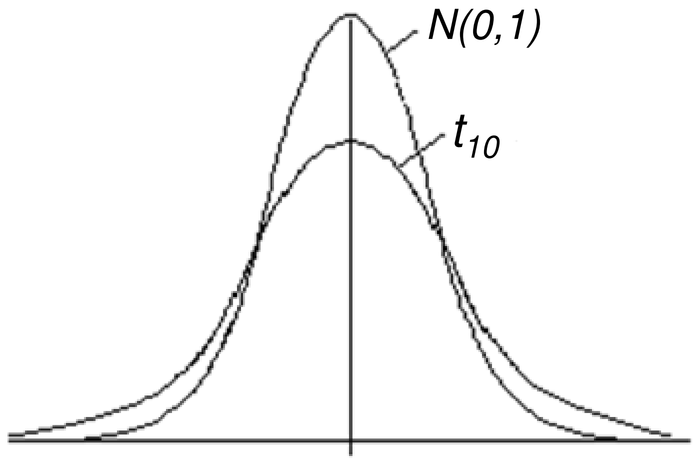
$$\left.\begin{aligned} E(X) &= 0 \\ V(X) &= \frac{\nu}{\nu-2} \end{aligned}\right\} X \sim t_\nu$$

(lê-se: X tem distribuição *t-student* com ν graus de liberdade)

Propriedades:

a) se Z $\sim N(0,1)$ e Q $\sim \chi^2_\nu$ então $\dfrac{Z}{\sqrt{\dfrac{Q}{\nu}}} \sim t_\nu$

b) se $\nu \to \infty$ então $t_\nu \sim N(0,1)$

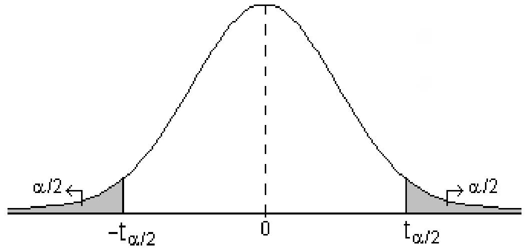


$$P(T_{10} > 2{,}764) = ?$$

$$P(T_{10} < -2{,}764) + P(T_{10} > 2{,}764) = 0{,}02$$

$$P(T_{10} > 2{,}764) = 0{,}01$$

Tabela II. Limites da distribuição t de Student.



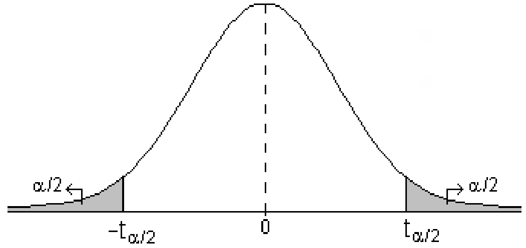
Graus de Liberdade (ν)	**Limites bilaterais**: $P(\lvert t \rvert > t_{\alpha/2})$ — Nível de Significância (α)							
	0,50	0,20	0,10	0,05	0,025	$\boxed{0{,}02}$	0,01	0,005
1	1,000	3,078	6,314	12,706	25,542	31,821	63,657	127,320
2	0,816	1,886	2,920	4,303	6,205	6,9[illegible]5	9,925	14,089
3	0,715	1,638	2,353	3,183	4,177	4,5[illegible]1	5,841	7,453
4	0,741	1,533	2,132	2,776	3,495	3,7[illegible]7	4,604	5,598
5	0,727	1,476	2,015	2,571	3,163	3,3[illegible]5	4,032	4,773
6	0,718	1,440	1,943	2,447	2,969	3,1[illegible]3	3,707	4,317
7	0,711	1,415	1,895	2,365	2,841	2,9[illegible]8	3,500	4,029
8	0,706	1,397	1,860	2,306	2,752	2,896	3,355	3,833
9	0,703	1,383	1,833	2,262	2,685	2,821	3,250	3,690
10	0,700	1,372	1,813	2,228	2,634	$\boxed{2{,}764}$	3,169	3,581
	0,25	0,10	0,05	0,025	0,0125	$\boxed{0{,}01}$	0,005	0,0025
Graus de Liberdade (ν)	Nível de Significância (α) — **Limites unilaterais**: $P(\lvert t \rvert > t_{\alpha})$							

$$P(T_{10} > ?) = 0{,}025 \qquad P(T_{10} < -?) + P(T_{10} > ?) = 0{,}05$$

$$P(T_{10} > 2{,}228) = 0{,}025 \qquad P(T_{10} < -2{,}228) + P(T_{10} > 2{,}228) = 0{,}05$$

$$P(T_{10} < -2{,}228) = 0{,}025$$

Tabela II. Limites da distribuição t de Student.



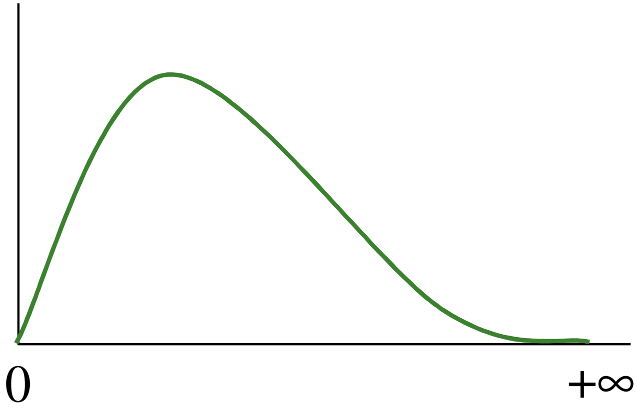
Graus de Liberdade (ν)	**Limites bilaterais**: $P(\lvert t \rvert > t_{\alpha/2})$							
	Nível de Significância (α)							
	0,50	0,20	0,10	$\boxed{0{,}05}$	0,025	0,02	0,01	0,005
1	1,000	3,078	6,314	12,706	25,542	31,821	63,657	127,320
2	0,816	1,886	2,920	4,303	6,205	6,965	9,925	14,089
3	0,715	1,638	2,353	3,1[illegible]3	4,177	4,541	5,841	7,453
4	0,741	1,533	2,132	2,7[illegible]6	3,495	3,747	4,604	5,598
5	0,727	1,476	2,015	2,5[illegible]1	3,163	3,365	4,032	4,773
6	0,718	1,440	1,943	2,4[illegible]7	2,969	3,143	3,707	4,317
7	0,711	1,415	1,895	2,3[illegible]5	2,841	2,998	3,500	4,029
8	0,706	1,397	1,860	2,306	2,752	2,896	3,355	3,833
9	0,703	1,383	1,833	2,262	2,685	2,821	3,250	3,690
10	0,700	1,372	1,813	$\boxed{2{,}228}$	2,634	2,764	3,169	3,581
	0,25	0,10	0,05	$\boxed{0{,}025}$	0,0125	0,01	0,005	0,0025
Graus de Liberdade (ν)	Nível de Significância (α)							
	Limites unilaterais: $P(\lvert t \rvert > t_{\alpha})$							

$$P(T_{10} > 2{,}764) = ?$$

$$P(T_{10} > 2{,}764) = 0{,}01$$

6. Distribuição *F (de Snedecor)*

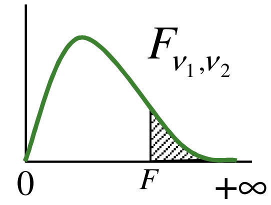
$$f(x) = \frac{\Gamma[(\nu_1+\nu_2)/2]}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} x^{\nu_1/2-1}\left(1+\frac{\nu_1}{\nu_2}x\right)^{-(\nu_1+\nu_2)/2} \qquad x \geq 0$$



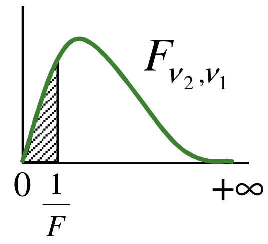
$$\left.\begin{aligned} E(X) &= \frac{\nu_2}{\nu_2-2} \\ V(X) &= \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \end{aligned}\right\} X \sim F_{\nu_1,\nu_2}$$

(lê-se: X tem distribuição F com ν_1 e ν_2 graus de liberdade)

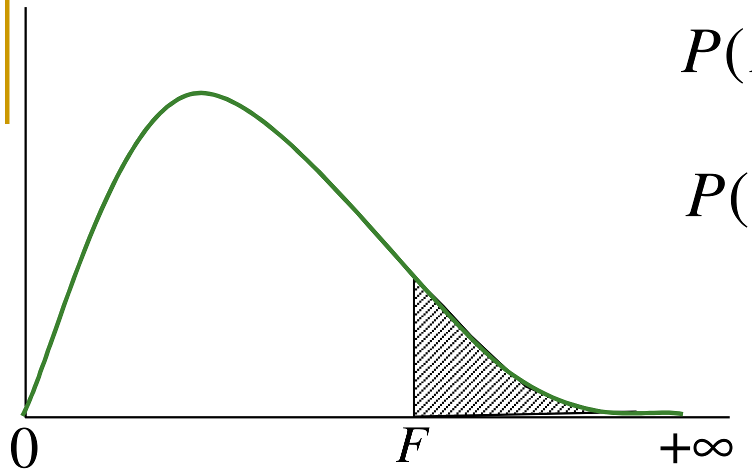
Propriedades:



a) se $U \sim \chi^2_{\nu_1}$ e $V \sim \chi^2_{\nu_2}$ então $\dfrac{U/\nu_1}{V/\nu_2} \sim F_{\nu_1,\nu_2}$



b) se $F \sim F_{\nu_1,\nu_2}$ então $\dfrac{1}{F} \sim F_{\nu_2,\nu_1} \Rightarrow P(F_{\nu_1,\nu_2} > F) = P(F_{\nu_2,\nu_1} < \dfrac{1}{F})$



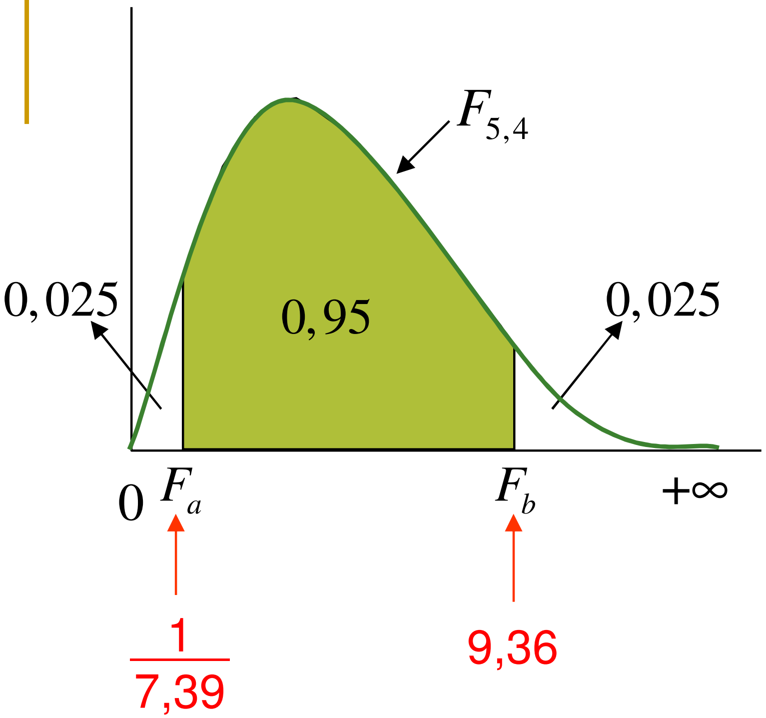
$$P(F_{5,4} > 15{,}52) = ?$$

$$P(F_{5,4} > ?) = 0{,}05$$

$$P(F_{5,4} > 15{,}52) = 0{,}01$$

$$P(F_{5,4} > 6{,}26) = 0{,}05$$

		ν_1																			
ν_2	α	1	2	3	4	5	6	7	8	9	10	11	12	15	20	24	30	40	60	120	Inf.
	0,05	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
1	0,025	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	976,7	984,9	984,9	993,1	997,2	1001,	1006,	1010,	1014,	1018,
	0,01	4052,	5000,	5403,	5625,	5764,	5859,	5928,	5982,	6022,	6056,	6082,	6106,	6157,	6209,	6235,	6261,	6287,	6313,	6339,	6366,
	0,001	4053*	5000*	5404*	5625*	5764*	5859*	5929*	5981*	6023*	6056*	6084*	6107*	6158*	6209*	6235*	6261*	6287*	6313*	6340*	6366*
	0,05	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
2	0,025	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,41	39,43	39,45	39,46	39,46	39,47	39,48	39,49	39,50
	0,01	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,41	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49	99,50
	0,001	998,5	999,0	999,2	999,2	999,3	999,3	999,4	999,4	999,4	999,4	999,4	999,4	999,4	999,4	999,5	999,5	999,5	999,5	999,5	999,5
	0,05	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
3	0,025	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,25	39,43	14,17	14,12	14,08	14,04	13,99	13,95	13,90
	0,01	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,13	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
	0,001	167,0	148,5	141,1	137,1	134,6	132,8	131,6	130,6	129,9	129,2	128,8	128,3	127,4	126,4	125,9	125,4	125,0	124,5	124,0	123,5
	0,05	[illegible]	[illegible]	[illegible]	[illegible]	$\boxed{6,26}$	6,16	6,09	6,04	6,00	5,96	5,93	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
4	0,025	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,66	8,56	8,51	8,46	8,41	8,36	8,31	8,26
	0,01	[illegible]	[illegible]	[illegible]	[illegible]	$\boxed{15,52}$	15,21	14,98	14,80	14,66	14,55	14,45	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
	0,001	74,14	61,25	56,18	53,44	51,71	50,53	49,66	49,00	48,47	48,05	47,70	47,41	46,76	46,10	45,77	45,43	45,09	44,75	44,40	44,05
	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	8,82	4,77	4,74	4,70	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
5	0,025	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,46	6,33	6,28	6,23	6,18	6,12	6,07	6,02
	0,01	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,96	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
	0,001	47,18	37,12	33,20	31,09	29,75	28,84	28,16	27,64	27,24	26,92	26,64	26,42	25,91	25,39	25,14	24,87	24,60	24,33	24,06	23,79



$$P(F_{5,4} > F_b) = 0{,}025$$

$$P(F_{5,4} < F_a) = 0{,}025 \;\Rightarrow P\left(F_{4,5} > \frac{1}{F_a}\right) = 0{,}025$$

7,39

		ν_1										
ν_2	α	1	2	3	4	5	6	7	8	9	10	11
	0,05	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,
1	0,025	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	976,
	0,01	4052,	5000,	5403,	5625,	5764,	5859,	5928,	5982,	6022,	6056,	608
	0,001	4053*	5000*	5404*	5625*	5764*	5859*	5929*	5981*	6023*	6056*	6084
	0,05	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,4
2	0,025	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,4
	0,01	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,4
	0,001	998,5	999,0	999,2	999,2	999,3	999,3	999,4	999,4	999,4	999,4	999,
	0,05	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,7
3	0,025	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,3
	0,01	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,1
	0,001	167,0	148,5	141,1	137,1	134,6	132,8	131,6	130,6	129,9	129,2	128,
	0,05	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,9
4	0,025	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,7
	0,01	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,4
	0,001	74,14	61,25	56,18	53,44	51,71	50,53	49,66	49,00	48,47	48,05	47,7
	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	8,82	4,77	4,74	4,7
5	0,025	[illegible]	[illegible]	[illegible]	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,5
	0,01	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,9
	0,001	47,18	37,12	33,20	31,09	29,75	28,84	28,16	27,64	27,24	26,92	26,6

Ajustamento de distribuições a dados reais

- A verificação do ajuste de um modelo estatístico a um determinado conjunto de valores de uma variável é uma das mais relevantes atribuições do método estatístico.
- Verificar o ajustamento significa verificar se o modelo pressuposto de fato pode ser utilizado para representar a distribuição de uma variável num determinado contexto.
- Diversos métodos estão disponíveis para essa verificação e, dentre eles, podemos citar:
  - métodos visuais, dentre os quais se destacam o histograma, o gráfico de probabilidade e o box-plot;
  - testes de aderência, dentre os quais podemos citar os testes de Kolmogorov-Smirnov e Shapiro-Wilks.