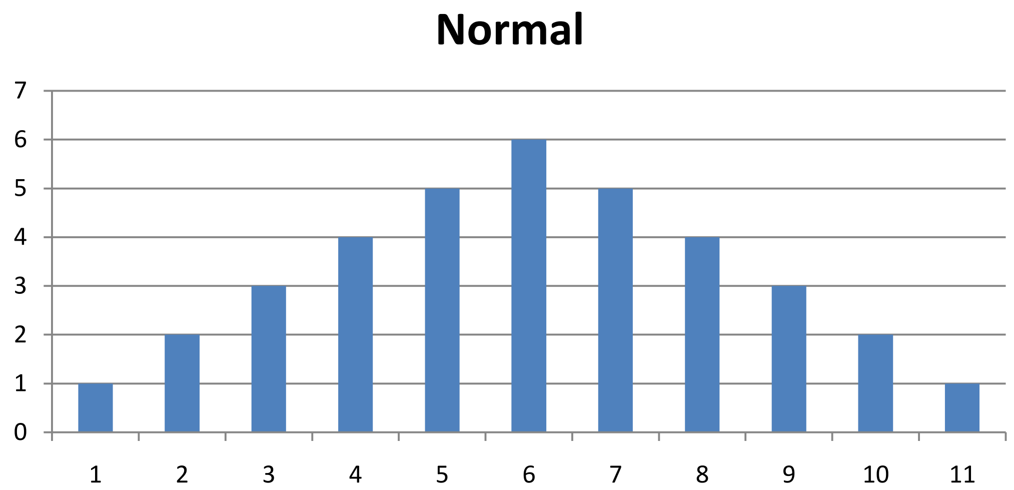
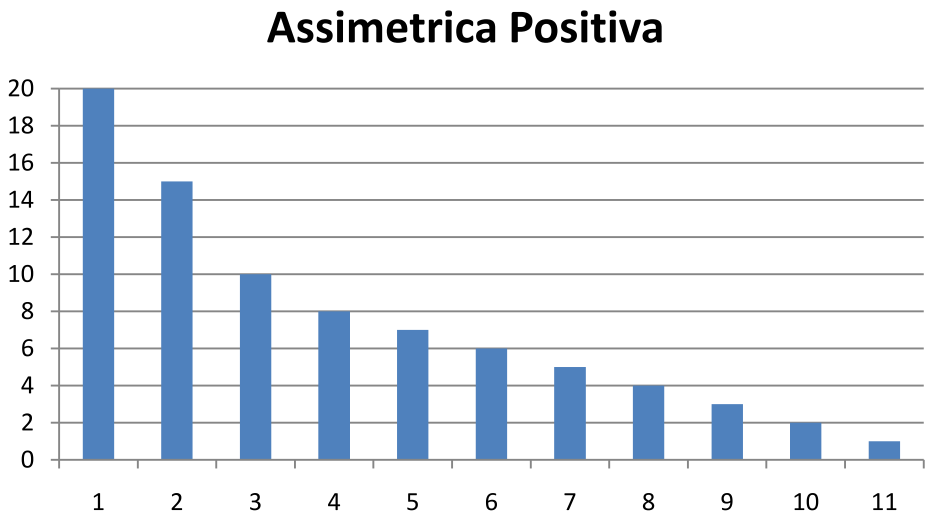
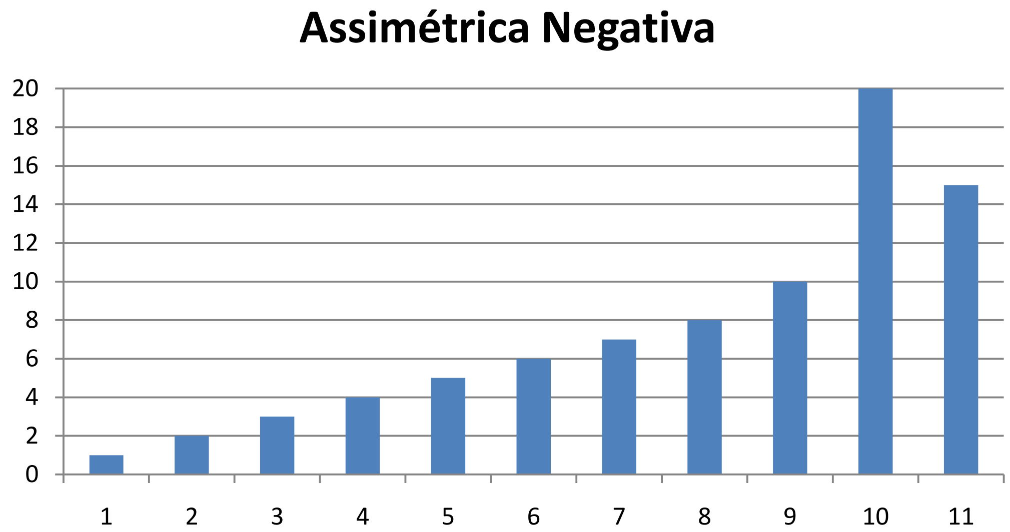
Distribuições de Probabilidade



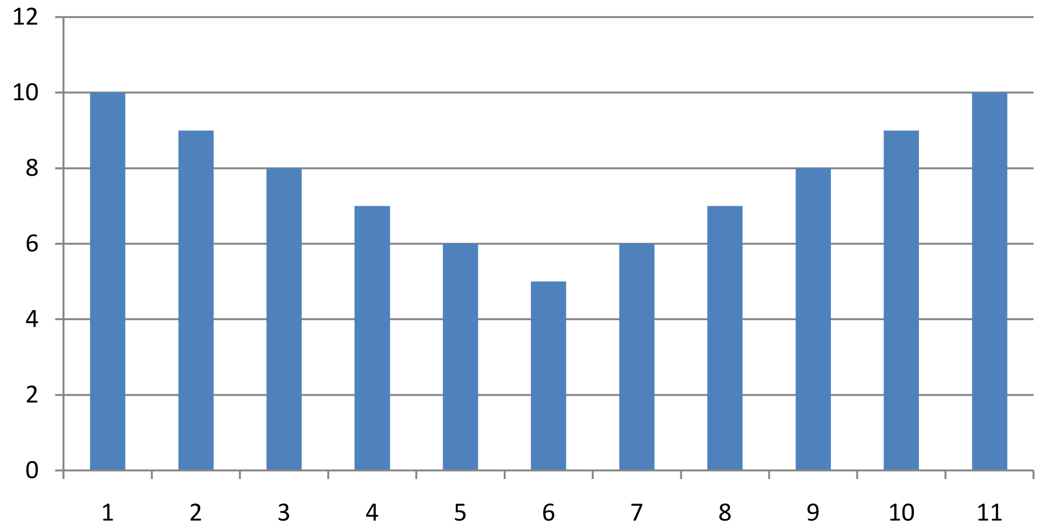
Exemplos: Temperatura do ar



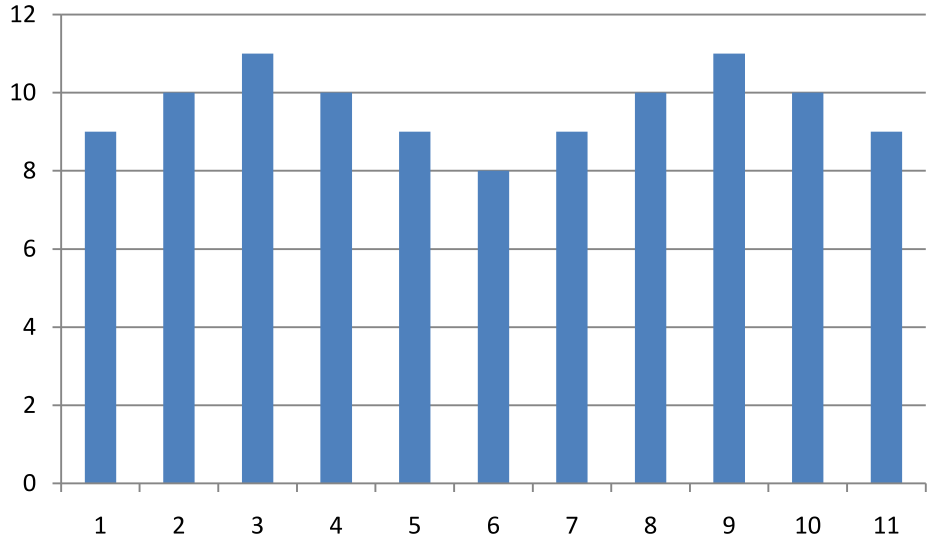
Exemplos: Precipitação



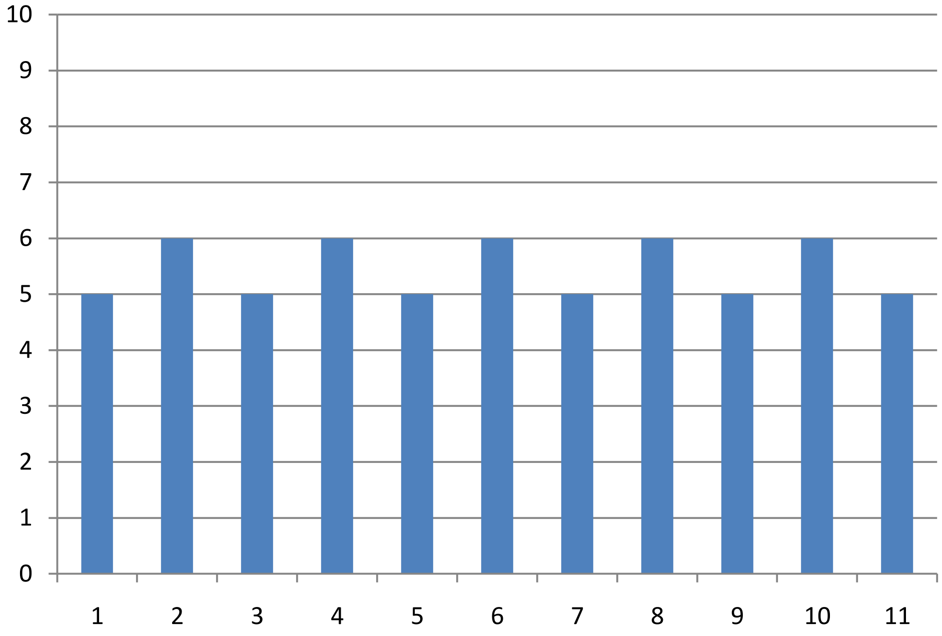
Exemplos: Temperatura de uma nuvem



Exemplos: Insolação



Exemplos: Concentração de aerossóis, gotículas



10
9
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
10
11

Variável Aleatória

– A variável aleatória é uma variável que tem um valor único (determinado aleatoriamente) para cada resultado de um experimento. A palavra aleatória indica que em geral só conhecemos aquele valor depois do experimento ser realizado (Triola, 1998).

Exemplos:

- *número de alunos que não compareceram a aula de estatística num determinado dia;*
- *altura de um adulto do sexo masculino selecionado aleatoriamente.*

- **Variável aleatória discreta:** é aquela que assume valores inteiros e finitos.

- **Variável aleatória contínua:** é aquela que pode assumir inúmeros valores num intervalo de números reais e é medida numa escala contínua.

Distribuição de Probabilidades

- A **distribuição de probabilidades** associa uma probabilidade a cada resultado numérico de um experimento, ou seja, dá a probabilidade de cada valor de uma variável aleatória.
- *Por exemplo, no lançamento de um dado cada face tem a mesma probabilidade de ocorrência que é 1/6.*

- Como os valores das distribuições de probabilidades são probabilidades, e como as variáveis aleatórias devem tomar um de seus valores, temos as duas regras a seguir que se aplicam a qualquer distribuição de probabilidades:

1) A soma de todos os valores de uma distribuição de probabilidades deve ser igual a 1

$\sum P(x) = 1$, onde x toma todos os valores possíveis

2) A probabilidade de ocorrência de um evento deve ser

$0 \leq P(x) \leq 1$ para todo x

No exemplo do lançamento de um dado, como todas as faces têm a mesma probabilidade de ocorrência que é 1/6, logo temos que $\sum P(x) = 1$, ou seja:

$$\sum P(x) = 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1$$

- A distribuição de probabilidades pode ser representada por um histograma de probabilidades, similar ao histograma de freqüências ***aonde a escala vertical representa probabilidades, em lugar das freqüências relativas.***

- Da mesma forma que antes, podemos calcular algumas características da distribuição:

- Média → $\mu=\sum x \ P(x)$
- Variância → $\sigma^2=\sum[(x-\mu)^2 P(x)] = [\sum x^2 P(x)] - \mu^2$
- Desvio-Padrão → $\sigma=([\sum x^2 P(x)] - \mu^2)^{1/2}$

- Média: indica o valor médio que esperaríamos ter se pudéssemos repetir as provas infinitamente. Não obtemos o valor que esperamos ocorrer com maior freqüência.
- Desvio Padrão: indica quanto a distribuição de probabilidades se dispersa em torna da média. Um grande desvio-padrão reflete dispersão considerável, enquanto que um desvio-padrão menor traduz menor variabilidade, com valores relativamente mais próximos da média.

A média de uma variável aleatória discreta é o resultado médio teórico de um número infinito de provas

Sendo que o valor esperado de uma variável aleatória discreta é denotado por E (Esperança) e representa o valor médio dos resultados:

$$E = \sum x \, P(x)$$

onde $E=\mu$. Isto é, a média de uma variável aleatória discreta coincide com seu valor esperado.

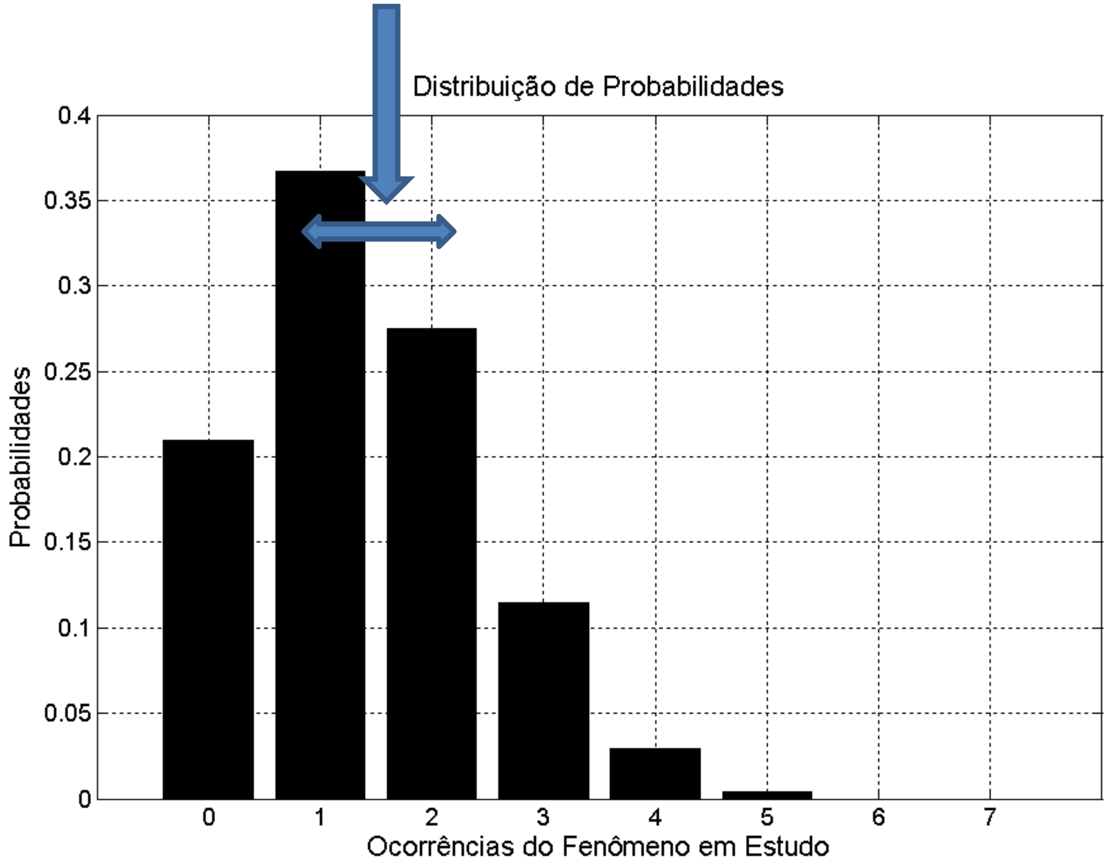
Cálculo da média, variância e desvio-padrão para uma distribuição de probabilidades.

X	P(x)	x P(x)	x^2	x^2 P(x)
0	0,210	0,000	0	0,000
1	0,367	0,367	1	0,367
2	0,275	0,550	4	1,100
3	0,115	0,345	9	1,035
4	0,029	0,116	16	0,464
5	0,004	0,020	25	0,100
6	0	0,000	36	0,000
7	0	0,000	49	0,000

$$\mu=\sum x \ P(x) = 1{,}398 = 1{,}4$$

$$\sigma^2=[\sum x^2 P(x)] - \mu^2 = 3{,}066-1{,}398^2 =1{,}111596 = 1{,}1$$

$$\sigma=(1{,}111596)^{1/2} = 1{,}054323 = 1{,}1$$



$$\mu = \sum x \; P(x) = 1{,}398 = 1{,}4$$

$$\sigma^2 = [\sum x^2 P(x)] - \mu^2 = 3{,}066 - 1{,}398^2 = 1{,}111596 = 1{,}1$$

$$\sigma = (1{,}111596)^{1/2} = 1{,}054323 = 1{,}1$$

Emprego das distribuições de probabilidade teóricas

- **Compacidade:** representação de um grande volume de dados. Uma distribuição teórica bem ajustada à série de dados pode caracterizar as propriedades da mesma.

- **Alisamento e interpolação:** os dados reais estão sujeitos a variações na amostragem que podem levar a falha de dados ou a dados errôneos nas distribuições empíricas. Logo pode-se verificar se um dado é real e de pode ocorrer ou não, para tanto podemos calcular a probabilidade de ocorrência.

- **Extrapolação:** estimar a probabilidade de eventos extremos a variação de um conjunto de dados particular exige a suposição de eventos ainda não observados. Isso pode ser realizado com a imposição de um modelo de probabilidade (isto é, uma distribuição teórica) ajustado a série de dados.

Distribuição Discreta e Contínua

- A **distribuição discreta** descreve quantidades aleatórias (dados de interesse) que podem assumir valores particulares e os valores são finitos. Por exemplo, a ocorrência de tempestades com granizo.

- A **distribuição contínua** representa quantidades aleatórias contínuas que podem tomar um número infinito de valores. Por exemplo: A temperatura, a pressão, a precipitação ou qualquer elemento medido numa escala contínua é uma variável aleatória contínua.

Distribuição de Probabilidade

- Quando conseguimos calcula as probabilidade de ocorrência de todos os possíveis valores de uma variável aleatória.
- O que implica que podemos calcular a probabilidade de uma determinada variável aleatória ocorrer

Distribuições contínuas

- A maioria das variáveis atmosféricas podem assumir valores contínuos.

Por exemplo: A temperatura, a precipitação, a altura geopotencial, e a velocidade do vento.

- Existem duas funções associadas a cada variável contínua X:

  1) **função densidade de probabilidade**, f(X);
  2) **função cumulativa de probabilidade**, ou **função de distribuição de probabilidade**, F(X).

- Logo a Função f(x) relaciona a probabilidade de ocorrência de um valor da variável aleatória:
  - $P(X = x) = f(x)$

  - Por exemplo no lançamento de dado:
  - $P(X=3) = f(3) = 1/6$

- Por outro lado, temos que a função f(X) é aquela cuja integral de X = a até X = b (b $\geq$ a) dá a probabilidade de que X assuma valores compreendidos no intervalo (a, b), ou seja,

$$P(a \leq X \leq b) = \int_a^b f(X)\, dX$$

- Já a função cumulativa de probabilidade F(b) é tal que:

$$F(b) = \Pr ob(X \leq b) = \int_{-\infty}^{b} f(X)\, dX$$

Exemplo:

- Se a nossa função densidade de probabilidade fosse descrita por uma distribuição exponencial, teríamos que:

$$f(X) = \lambda e^{-X\lambda} dX \quad \text{com x > 0}$$

Já a função cumulativa de probabilidade

$$F(X) = \int_0^x \lambda e^{-X\lambda} dX = 1 - e^{-X\lambda}$$

Qualquer função definida no campo real só pode ser considerada como uma função densidade de probabilidade caso as seguintes condições sejam satisfeitas:

1) f(X) > 0   para todo X e

2) $F(X) = \int_{-\infty}^{\infty} X \, dX = 1$

Dessa maneira, a probabilidade de uma variável X assuma os valores no intervalo (a,b) é:

$$P(a \leq X \leq b) = \int_{a}^{b} f(X) \, dX = F(b) - F(a)$$

Distribuições

Contínuas

- Normal
- Gama
- Valores Extremos
- Exponencial

Discretas

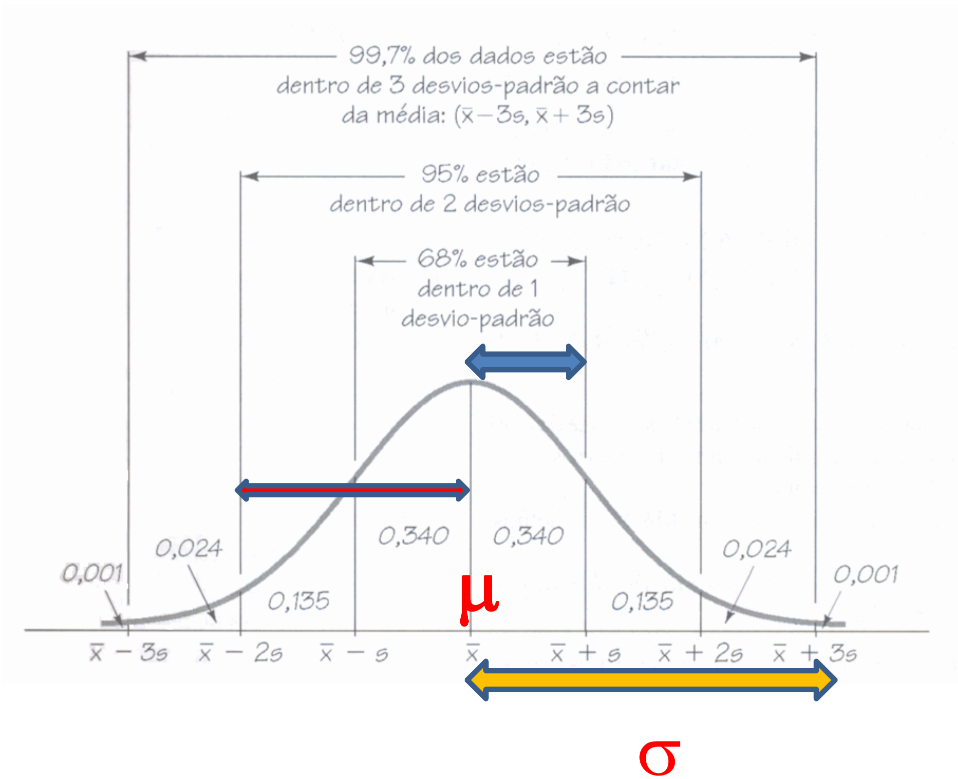
- Binomial
- Poisson
- Geométrica

Distribuição Normal

Uma variável aleatória contínua é representada por uma distribuição normal se a sua distribuição for:

- Simétrica
- A forma gráfica é similar a um sino

Distribuição Normal



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

$$\text{para } -\infty < x < +\infty$$

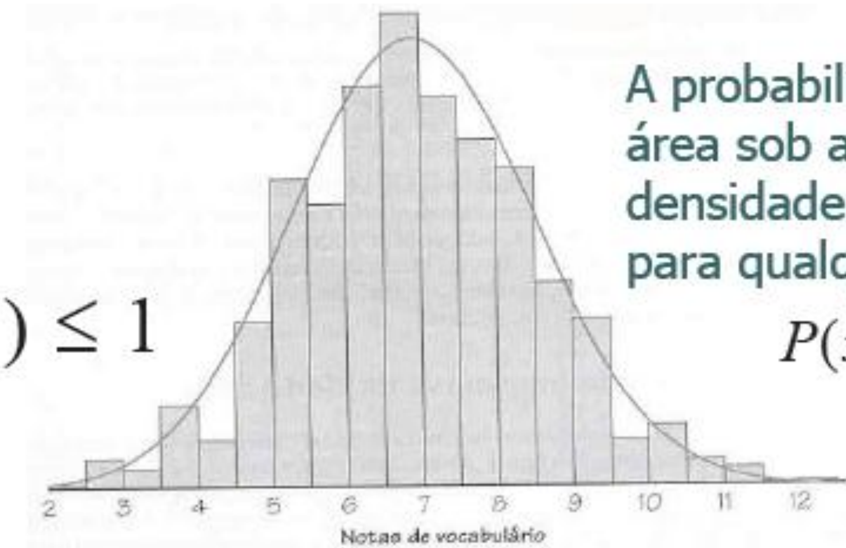
onde μ é a média e;

σ é desvio-padrão da população

A função densidade pode ser compreendia como uma extensão natural do histograma.

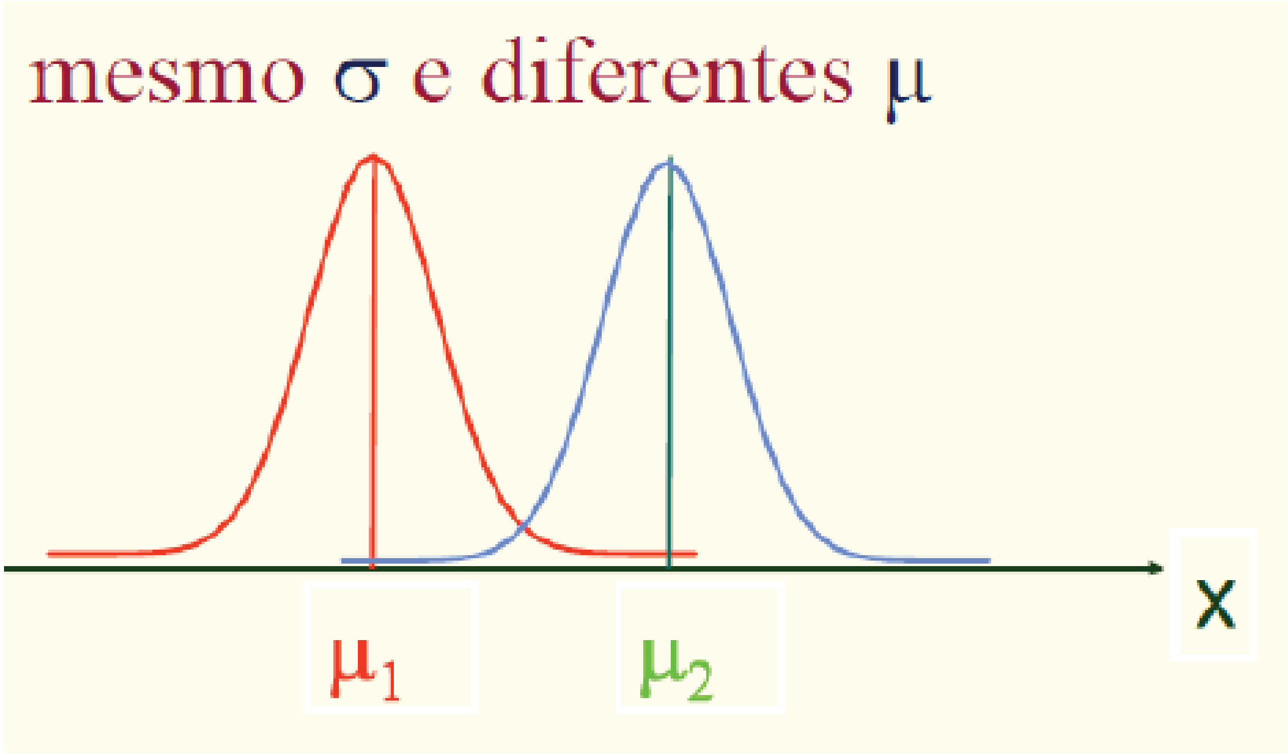
$$\int_{-\infty}^{+\infty} P(x)dx = 1$$

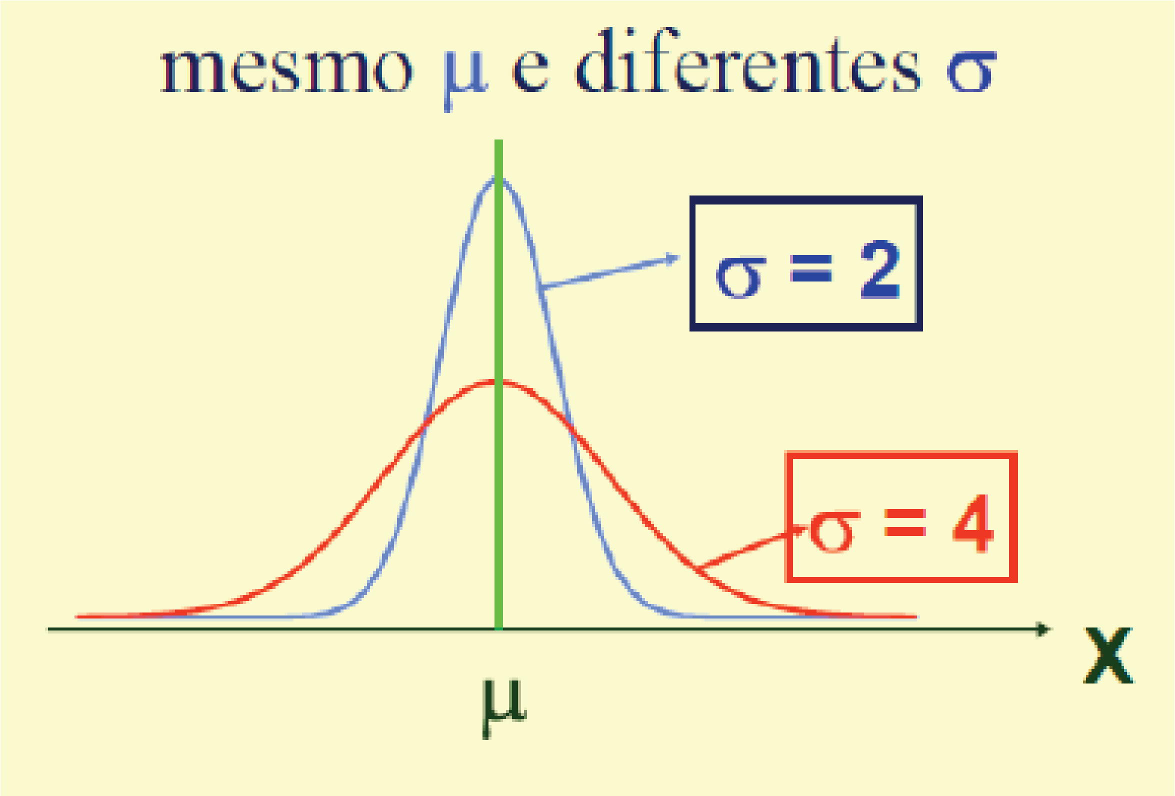
$$0 \leq P(x) \leq 1$$



A probabilidade é a área sob a curva de densidade. Portanto, para qualquer P(x):

$$P(x) \geq 0$$


mesmo σ e diferentes μ
μ1
μ2
X


mesmo μ e diferentes σ
σ = 2
σ = 4
X
μ

- Cada para de parâmetros (μ; σ) define uma distribuição normal distinta;