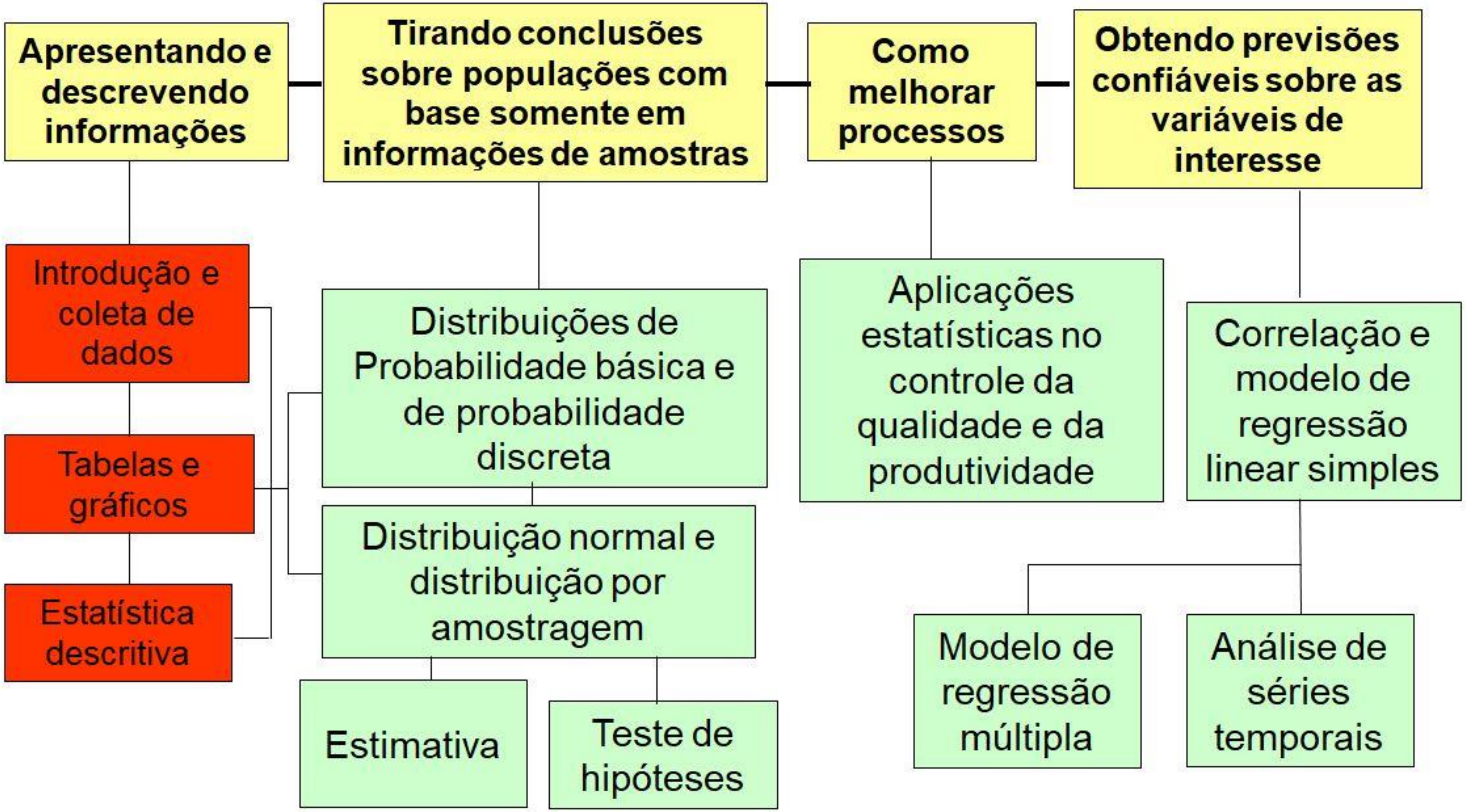
Núcleos de Desenvolvimento

Estatística Descritiva

Porque um gerente precisa conhecer estatística:

- Os gerentes precisam saber como apresentar e descrever informações de forma adequada;
- Os gerentes precisam saber como tirar conclusões a partir de grandes populações com base somente na informação obtida de amostras;
- Os gerentes precisam saber como melhorar processos;
- Os gerentes precisam saber como obter previsões confiáveis a partir de variáveis de interesse.

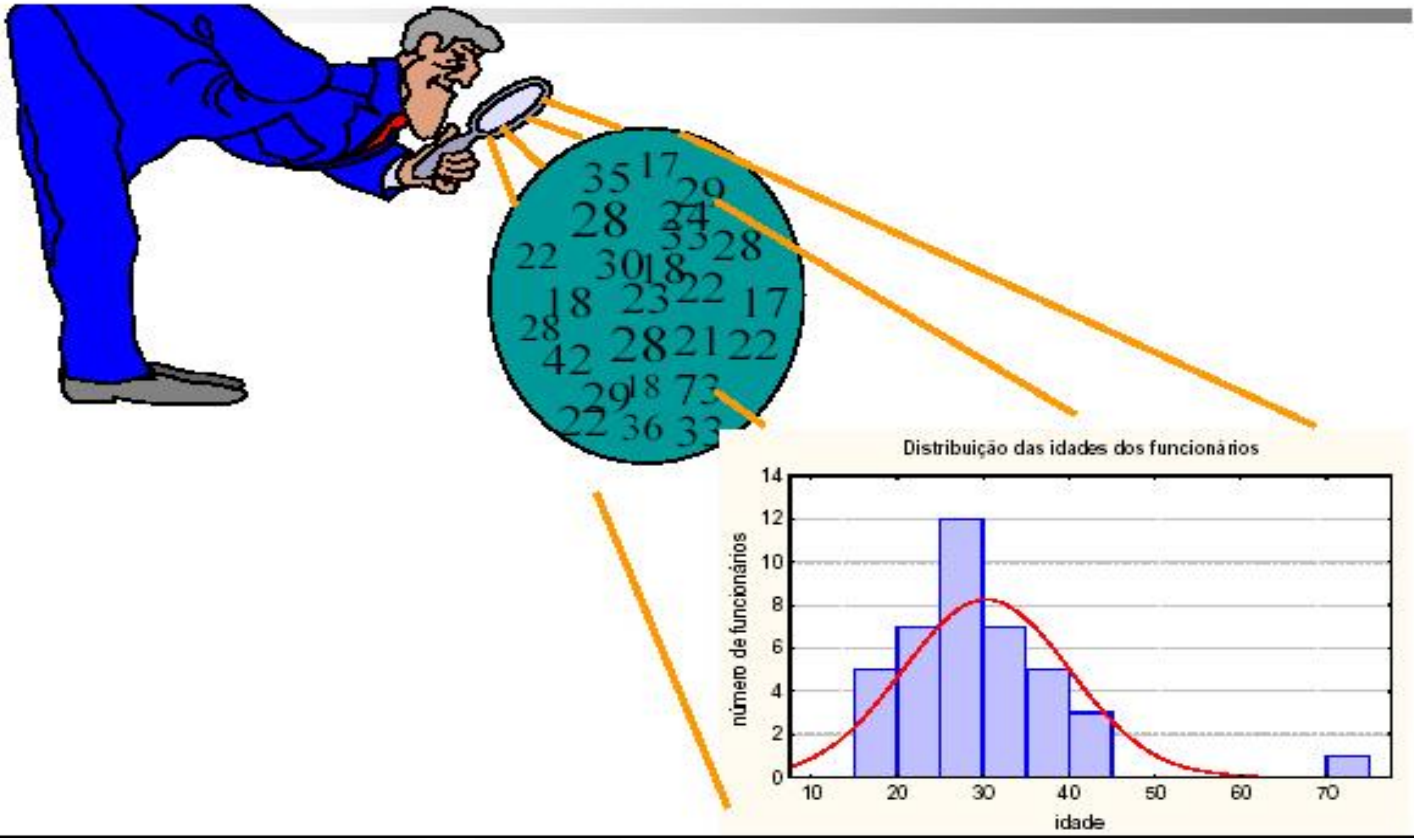
ESEG
FACULDADE DO GRUPO ETAPA

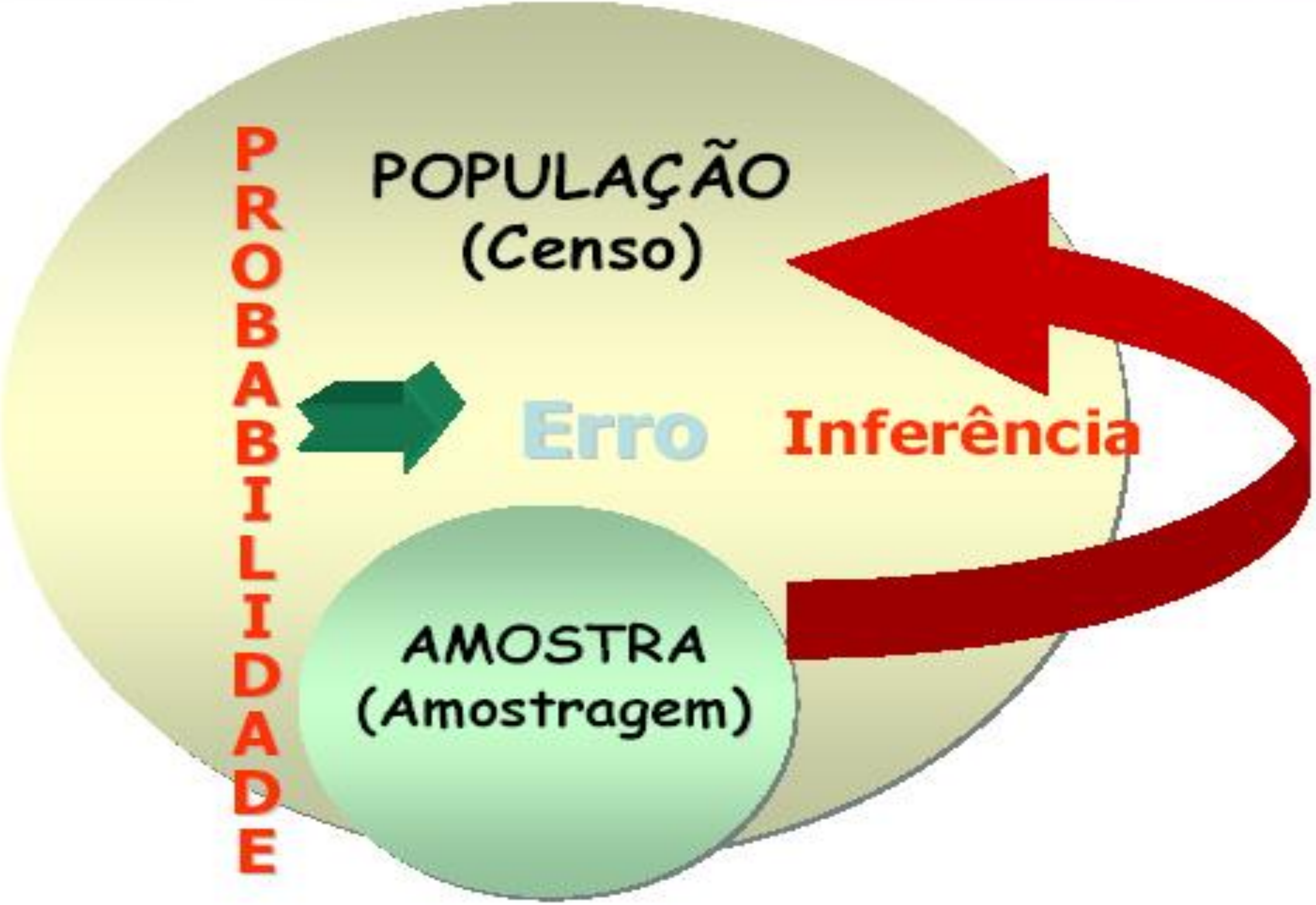


ESEG
FACULDADE DO GRUPO ETAPA

Análise exploratória de dados

técnicas para extrair informações de conjuntos de dados





PROBABILIDADE
POPULAÇÃO
(Censo)
Erro
Inferência
AMOSTRA
(Amostragem)

EXPLORAÇÃO DE DADOS

- Interesse da Estatística: características passíveis de representação **numérica**.

- Obtidas através de **medições e contagens.**

- Características = **variáveis.**

- Dois tipos de variáveis: **qualitativas e quantitativas.**

- VARIÁVEL
  - QUALITATIVA
    - NOMINAL
    - ORDINAL
  - QUANTITATIVA
    - DISCRETA
    - CONTÍNUA

Nominal

Sexo
Religião
Estado civil
Curso

ORDINAL

Conceito
Grau de Instrução
Mês
Dia da semana

DISCRETAS

Número de faltas
Número de irmãos
Número de acertos

CONTÍNUAS

Altura
Área
Peso
Volume

QUALITATIVA - ORDINAL

As características são ordenadas
(de maneira crescente ou decrescente) em
situações para as quais a posição associada é
importante

Exemplo: Ao se verificar o desempenho de uma pessoa
ou atividade há cinco categorias. Para facilidade de
codificação associa-se um número a cada
desempenho:
(5) ótimo, (4) bom , (3) regular, (2) ruim e (1) péssimo
Um (4) indica melhor desempenho que um (3) etc.

Variáveis quantitativas

- Quando o resultado da observação é um número, decorrente de um processo de mensuração ou contagem.
- Exemplos: número de filhos; salário mensal; altura; peso; idade; tamanho da família etc.

Variável quantitativa discreta

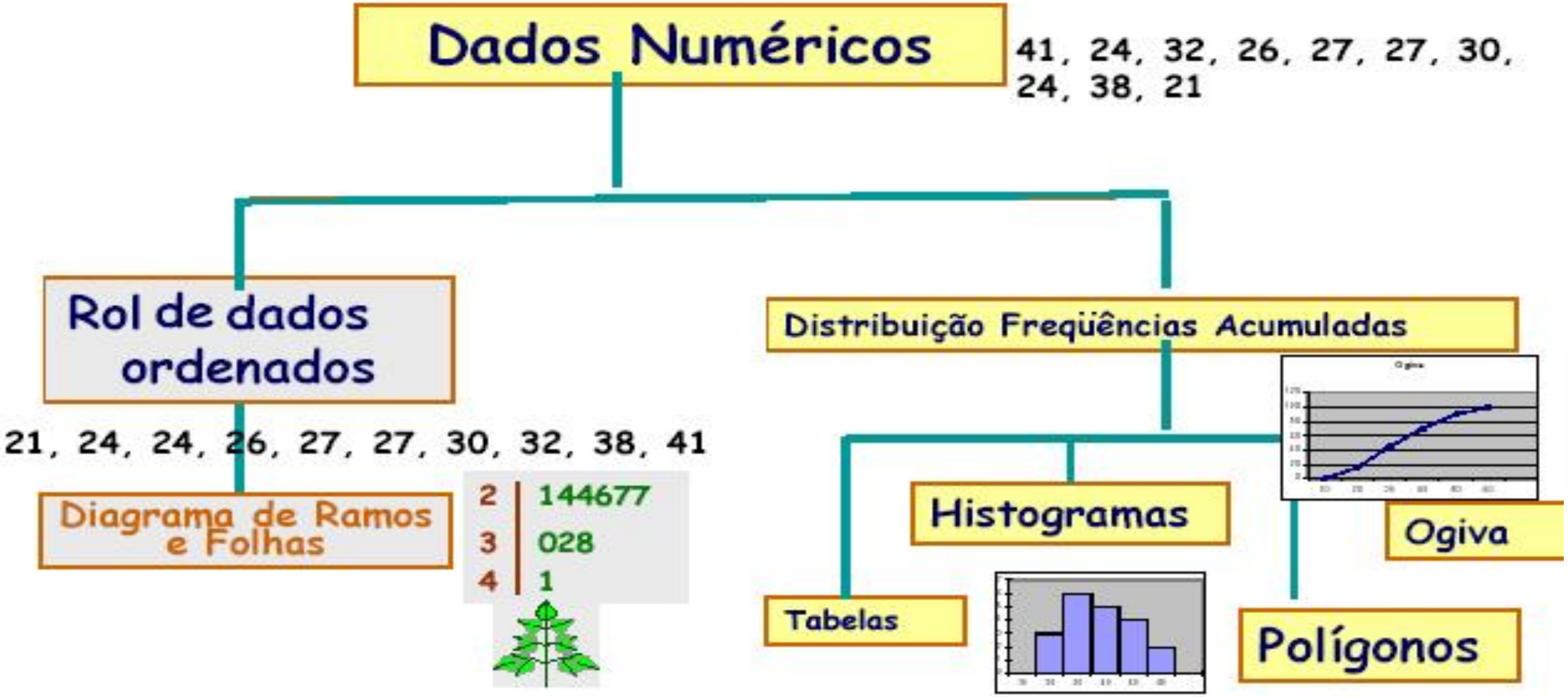
- Quando os resultados possíveis da observação formam um conjunto finito ou enumerável de números e que resultam, freqüentemente, de uma contagem.
- Exemplos: número de filhos (0, 1, 2, ...); tamanho da família (1, 2, 3, ...)

Variável quantitativa contínua

Quando os possíveis valores formam um intervalo ou uma união de intervalos de números reais e que resultam, normalmente, de uma mensuração.

- **Exemplos:** salário mensal; altura; peso.

ESEG
FACULDADE DO GRUPO ETAPA


Dados Numéricos
41, 24, 32, 26, 27, 27, 30, 24, 38, 21
Rol de dados ordenados
21, 24, 24, 26, 27, 27, 30, 32, 38, 41
Diagrama de Ramos e Folhas
2 | 144677
3 | 028
4 | 1
Distribuição Freqüências Acumuladas
Histogramas
Ogiva
Tabelas
Polígonos

MEDIDAS ESTATÍSTICAS

- Valores que resumem todo o conjunto de dados.
- Podem ser divididas em:
- \- Medidas de Posição
- \- Medidas de Dispersão
- \- Medidas de Assimetria
- \- Medidas de Curtose

ESEG
FACULDADE DO GRUPO ETAPA

Medidas de posição: média, mediana, moda, quartis e decis.

Medidas de dispersão: desvio padrão, variância e coeficiente de variação.

Média Aritmética

Idéia mais comum : soma / número de valores

Propriedades:

- pode ser sempre calculada
- é única
- sensível a todos os valores
- média (x + constante) = constante + média (x)
- soma dos desvios em torno da média é zero

Dados não agrupados

Ex 3, 5, 10, 15, 2, 1, 16

Média - População

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

ESEG
FACULDADE DO GRUPO ETAPA

Mediana

- Divide um conjunto ordenado de dados em partes iguais
- Não é afetada por valores extremos

Passos :

- ordenar valores
- verificar se há número de argumentos

  par - média dos dois valores do meio

  ímpar - valor do meio

Moda

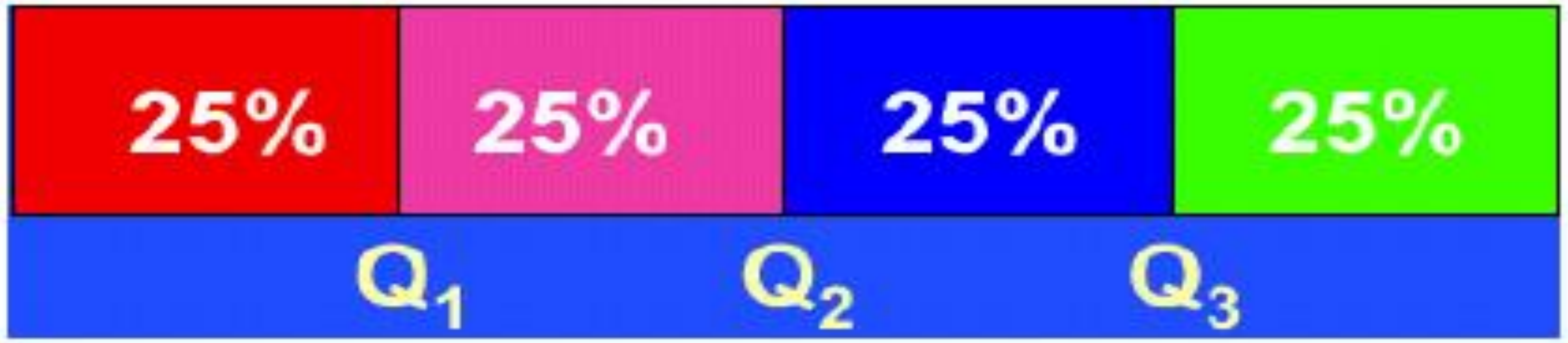
- Valor que ocorre com maior freqüência
- Podem existir séries com mais de uma moda
- Não é afetada pelos valores extremos

Quartis e Média Quartílica

Quartis : **não são** medidas de tendência central

Divide os dados ordenados em quadro conjuntos

Posição do quartil **i** :



Média Quartílica

- Medida de tendência central
- Média do 1° e 3° quartil.
- Não afetada por valores extremos

$$\text{média quartílica} = \frac{Q_1 + Q_3}{2}$$

ESEG
FACULDADE DO GRUPO ETAPA

Medidas de dispersão

- Necessárias para expressar a variabilidade de um conjunto de dados
- Indicam se os valores estão próximos ou separados
  - Amplitude
  - Desvio Médio Absoluto
  - Desvio Padrão e Variância
  - Coeficiente de variação (variação relativa)

Média como ponto de referência

Amplitude total (*intervalo*)

Medida mais simples de calcular e fácil de entender

Foca o maior e o menor valor

Pode ser expressa

- diferença entre maior e menor
- maior e menor valor do grupo

Exemplo :

Números	Diferença	Mín e Máx
14, 3, 17, 4, 8, 73, 36, 48	73- 3 = 70	de 3 a 73

ESEG
FACULDADE DO GRUPO ETAPA

Variância - População

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}$$

Desvio Padrão - População

σ raiz positiva da variância

Coeficiente de Variação

Coeficiente de dispersão relativa

Razão entre desvio e média

- permitir análise conjunta da média e do desvio
- é melhor aplicar na bolsa ou na poupança ?

$$CV = \frac{\sigma}{\mu}$$

população

$$CV = \frac{S}{\overline{X}}$$

amostra

Regras empíricas

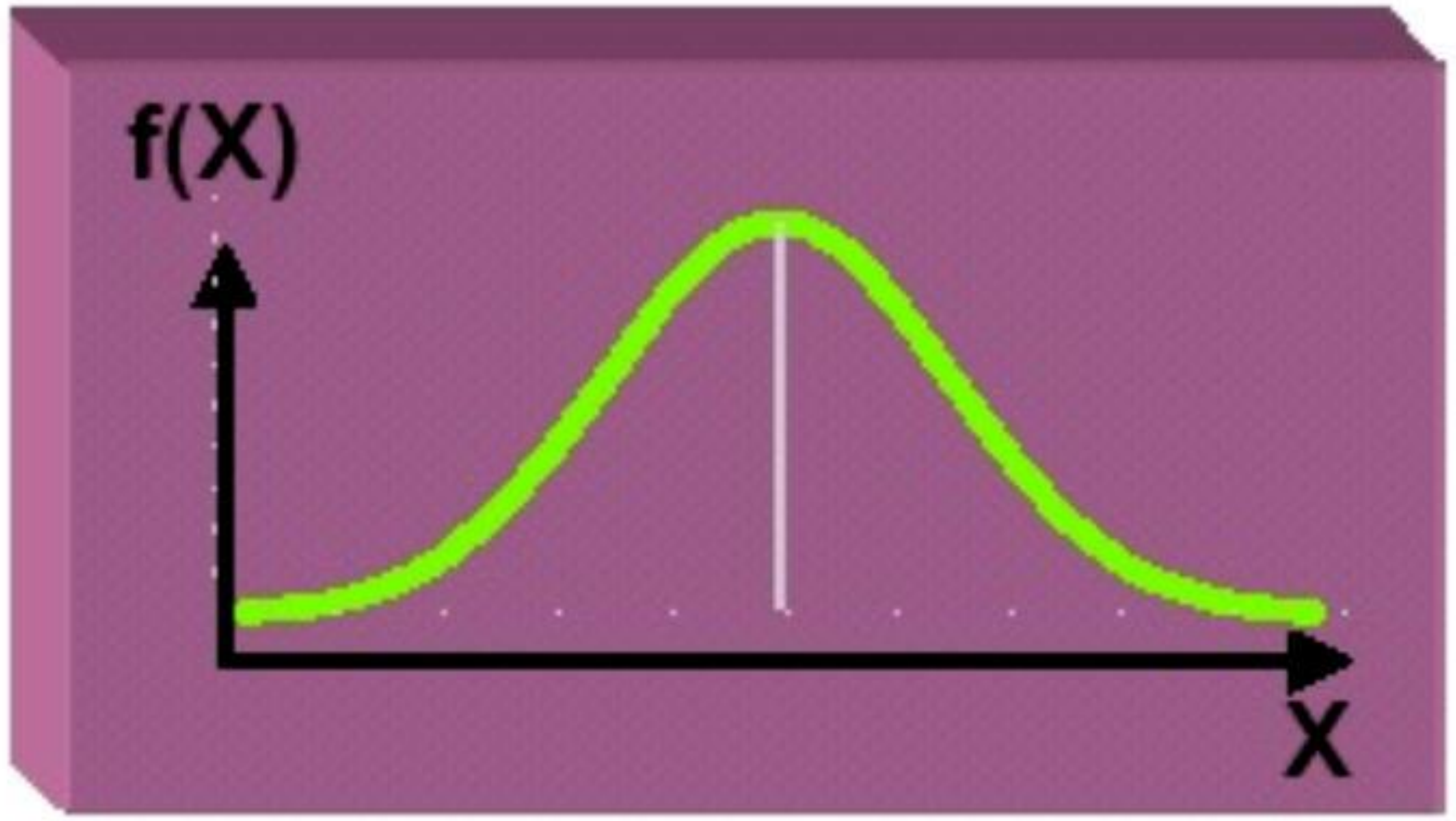
$CV < 15\%$ Tem-se baixa dispersão

$15\% < CV < 30\%$ Tem-se média dispersão

$CV > 30\%$ Tem-se elevada dispersão

Distribuição Normal

1. Curva em forma de "sino"
2. Média =Moda=Mediana
3. A variável aleatória assume infinitos valores



Média
Mediana
Moda

ESEG
FACULDADE DO GRUPO ETAPA

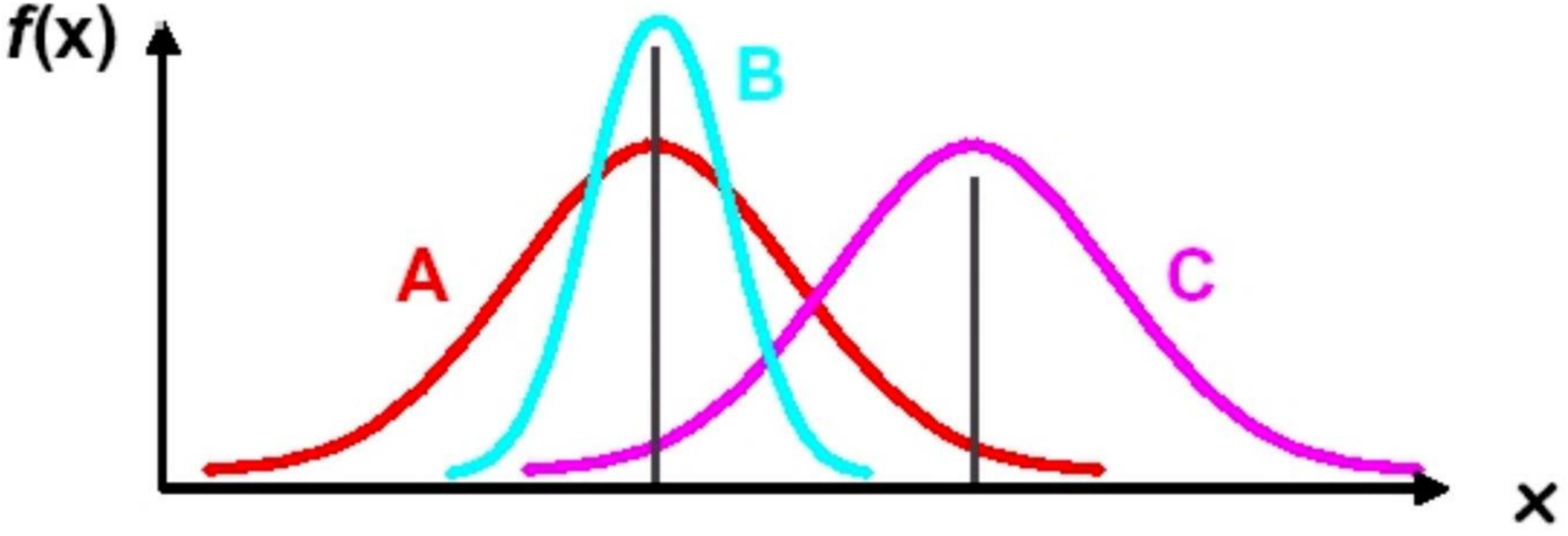
Distribuição Normal

Função de densidade de probabilidade

$$f(x) = \frac{1}{\sqrt{2\pi}\;\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

f(X)	=	freqüência da variável aleatória *X*
π	=	3.14159; *e* = 2.71828
σ	=	desvio padrão da população
x	=	valores da variável ($-\infty < x < \infty$)
μ	=	media da população

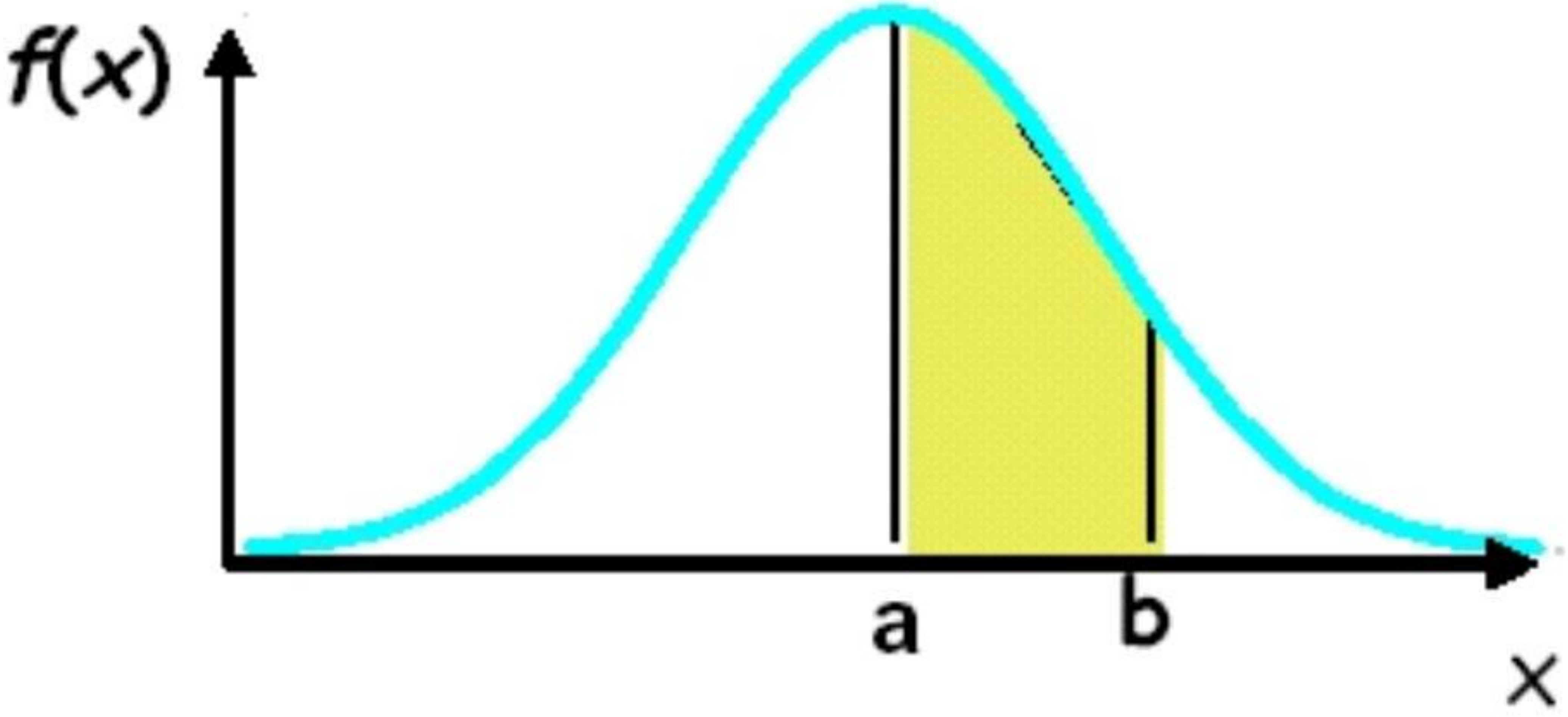
Efeitos Modificando a Média e o Desvio Padrão (μ e σ)



Probabilidade na Distribuição Normal

Probabilidade é a área sob a curva

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

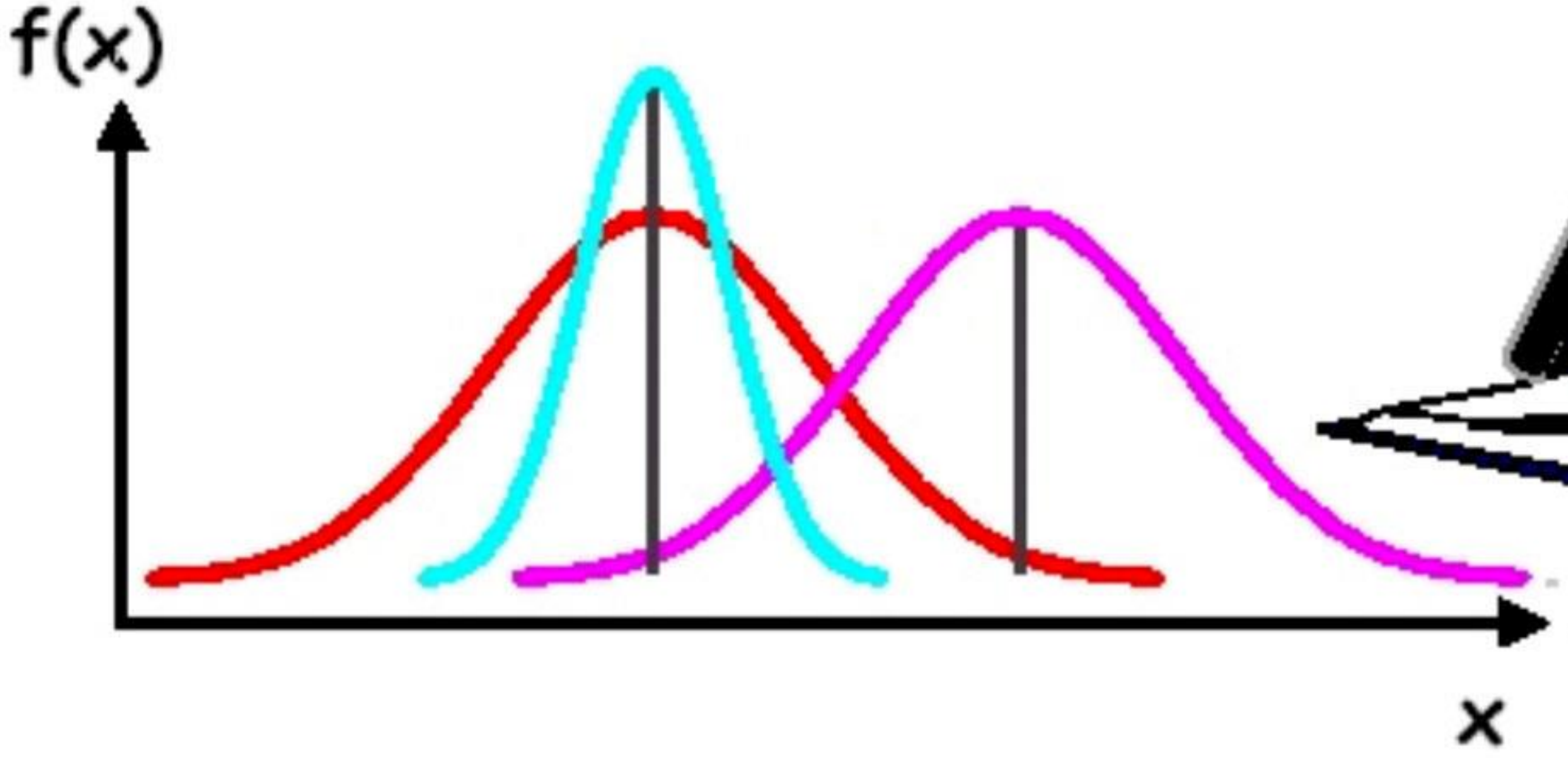


Tabelas da Distribuição Normal

Distribuição Normal

Diferentes médias e diferentes desvios padrões

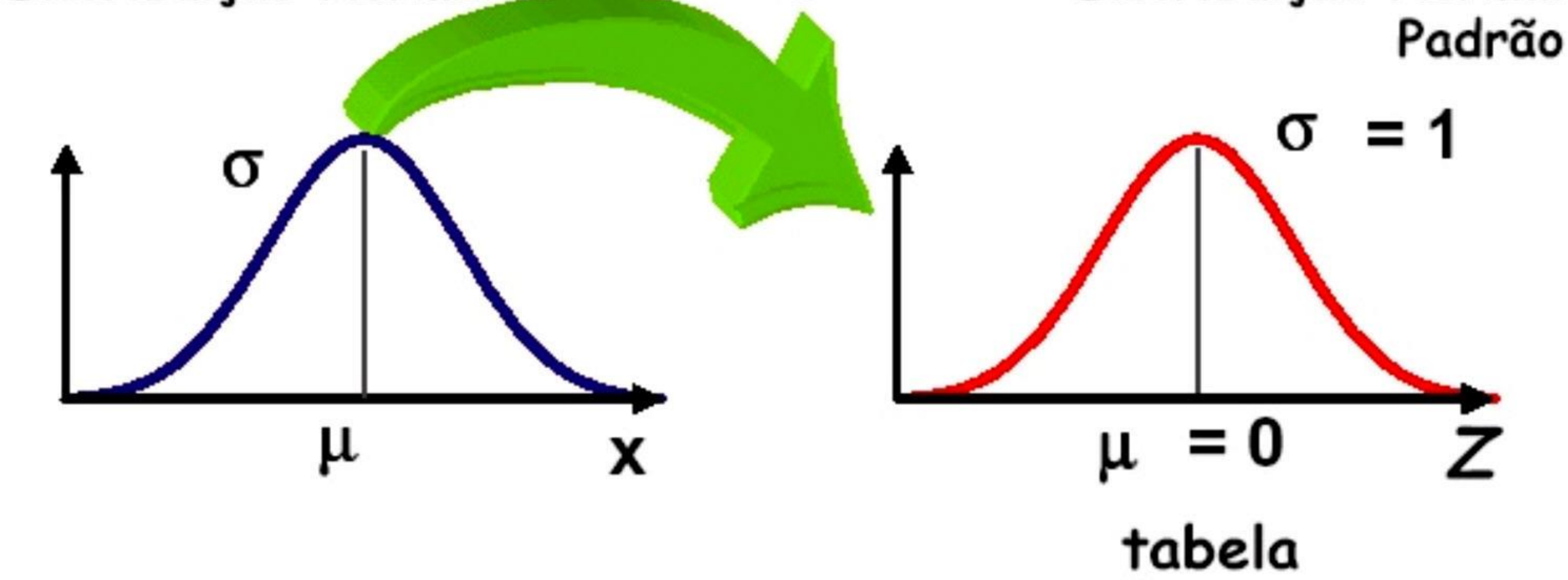
Cada distribuição normal requer uma tabela.



Existem infinitos valores!

Distribuição Normal Padrão

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



Exemplo

Seja $X : n(\mu = 80, \sigma = 5)$

$$z = \frac{x - \mu}{\sigma} = \frac{80{,}6 - 80}{5} = 0{,}12$$

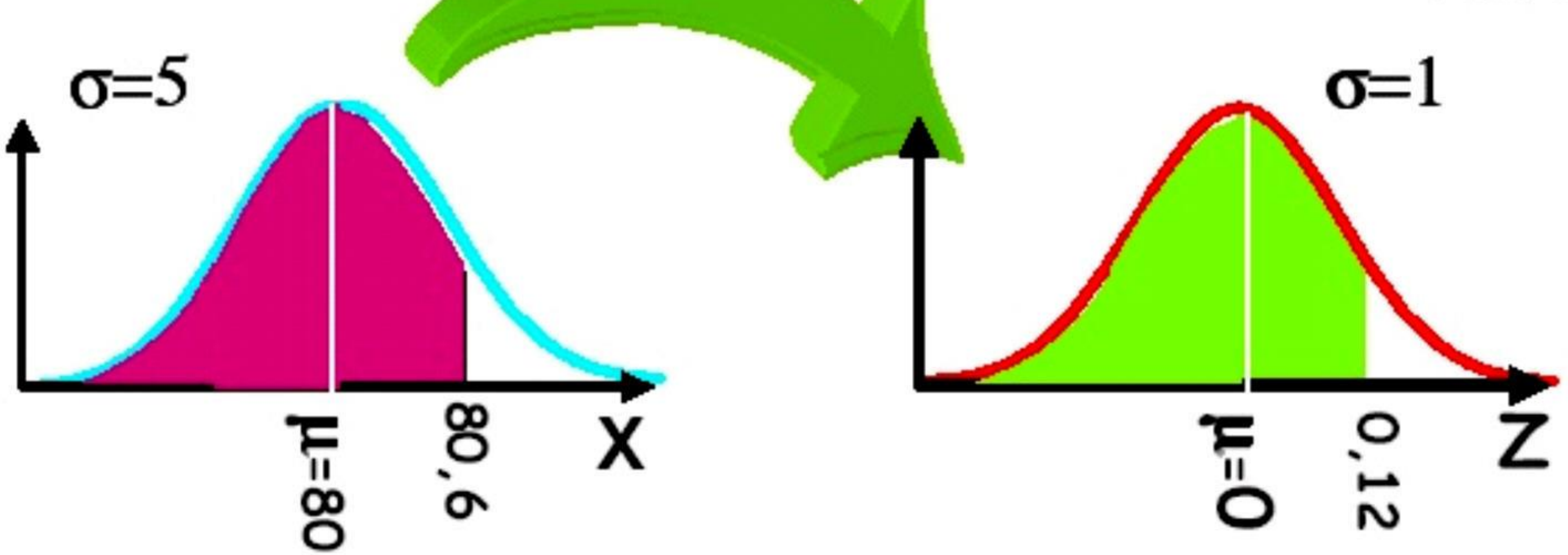
Determine

$P(X < 80{,}6)$

$P(X<80{,}6) = 54{,}78\%$

Distribuição Normal

Distribuição Normal Padrão



Obrigado pela atenção!

ESEG
FACULDADE DO GRUPO ETAPA