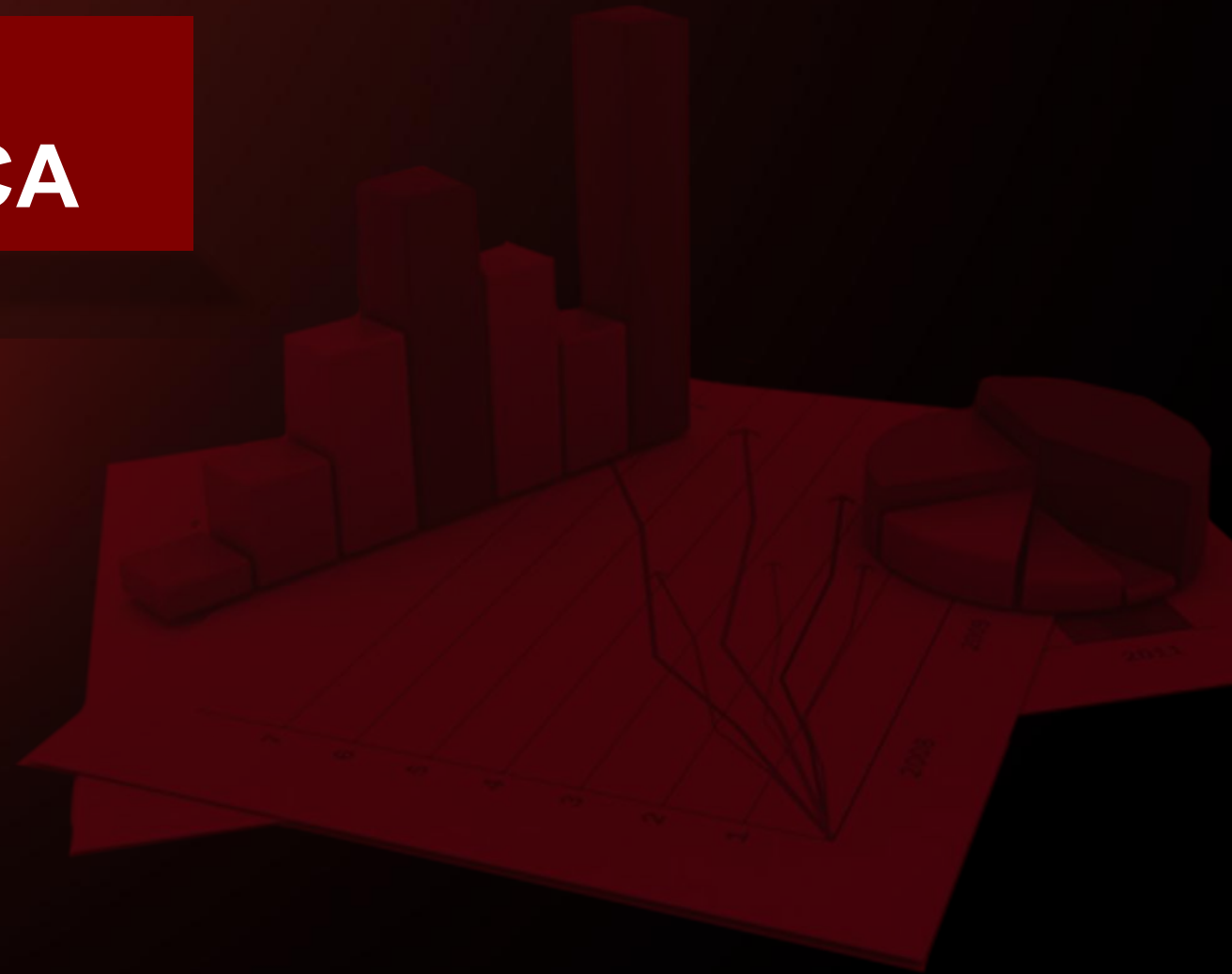


# MODELAGEM E INFERÊNCIA ESTATÍSTICA

## Regressão Linear: Introdução



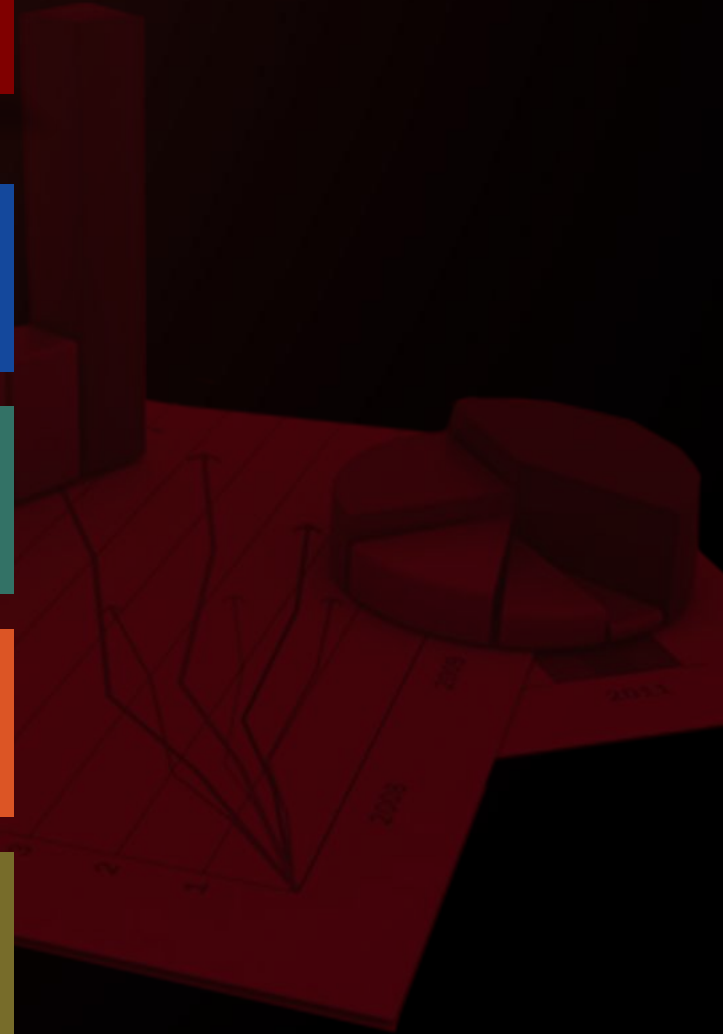
# O QUE VOU ESTUDAR HOJE?

Modelo de regressão linear simples.

Representações das variáveis

Lembrar conceitos relacionados aos parâmetros usados na distribuição normal.

Modelo probabilístico linear.



# O QUE É REGRESSÃO?

- Objetivo: Estudar a relação entre duas ou mais variáveis para estudar uma terceira.
- Tipos de variáveis: Determinísticas e Aleatórias.

## Variáveis determinísticas

$$y = 25 + 0,3 x$$

Aluguel de uma van R\$ 25 e 30 centavos a cada quilômetro andado

## Variáveis aleatórias

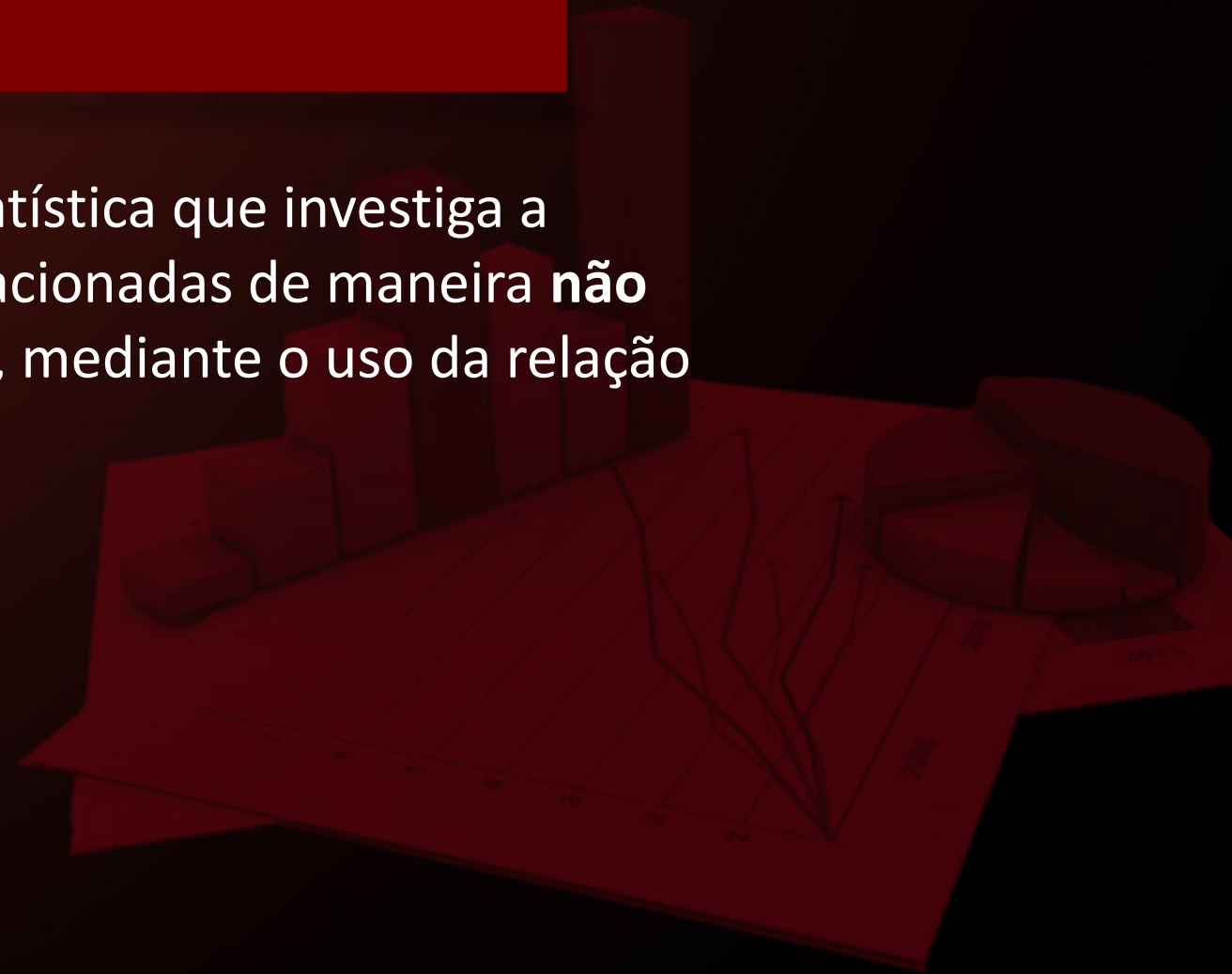
$x$  = idade de uma criança

$y$  = tamanho do vocabulário da criança

# O QUE É REGRESSÃO?

"A análise de regressão é a parte da estatística que investiga a relação entre duas ou mais variáveis relacionadas de maneira **não** determinística", (DEVORE, 2008, p. 455), mediante o uso da relação determinística

$$v = \beta_0 + \beta_1 x_1$$



# MODELO DE REGRESSÃO LINEAR SIMPLES

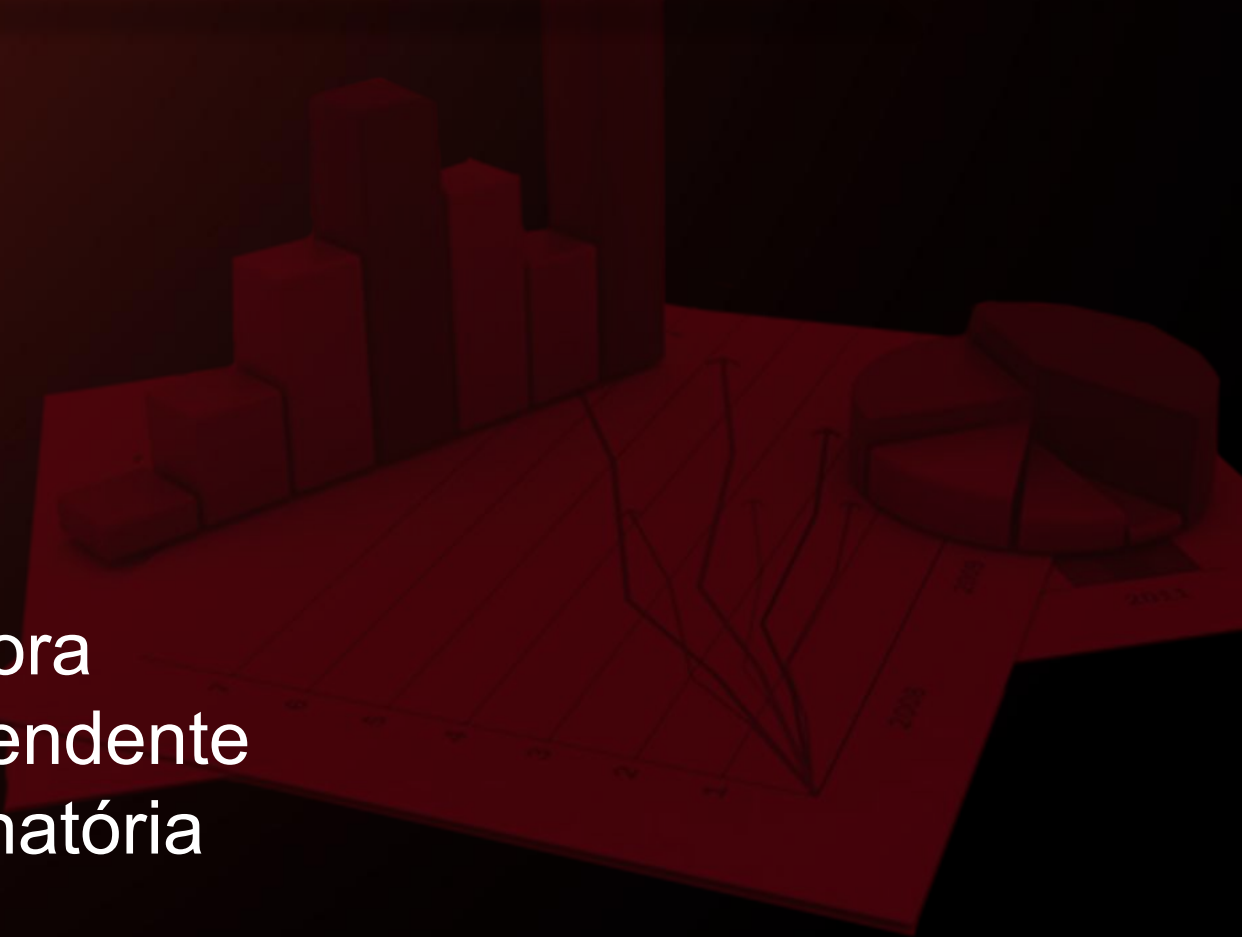
$$y = \beta_0 + \beta_1 x_1$$



Variável aleatória  
Variável dependente  
Variável resposta



Variável preditora  
Variável independente  
Variável explanatória

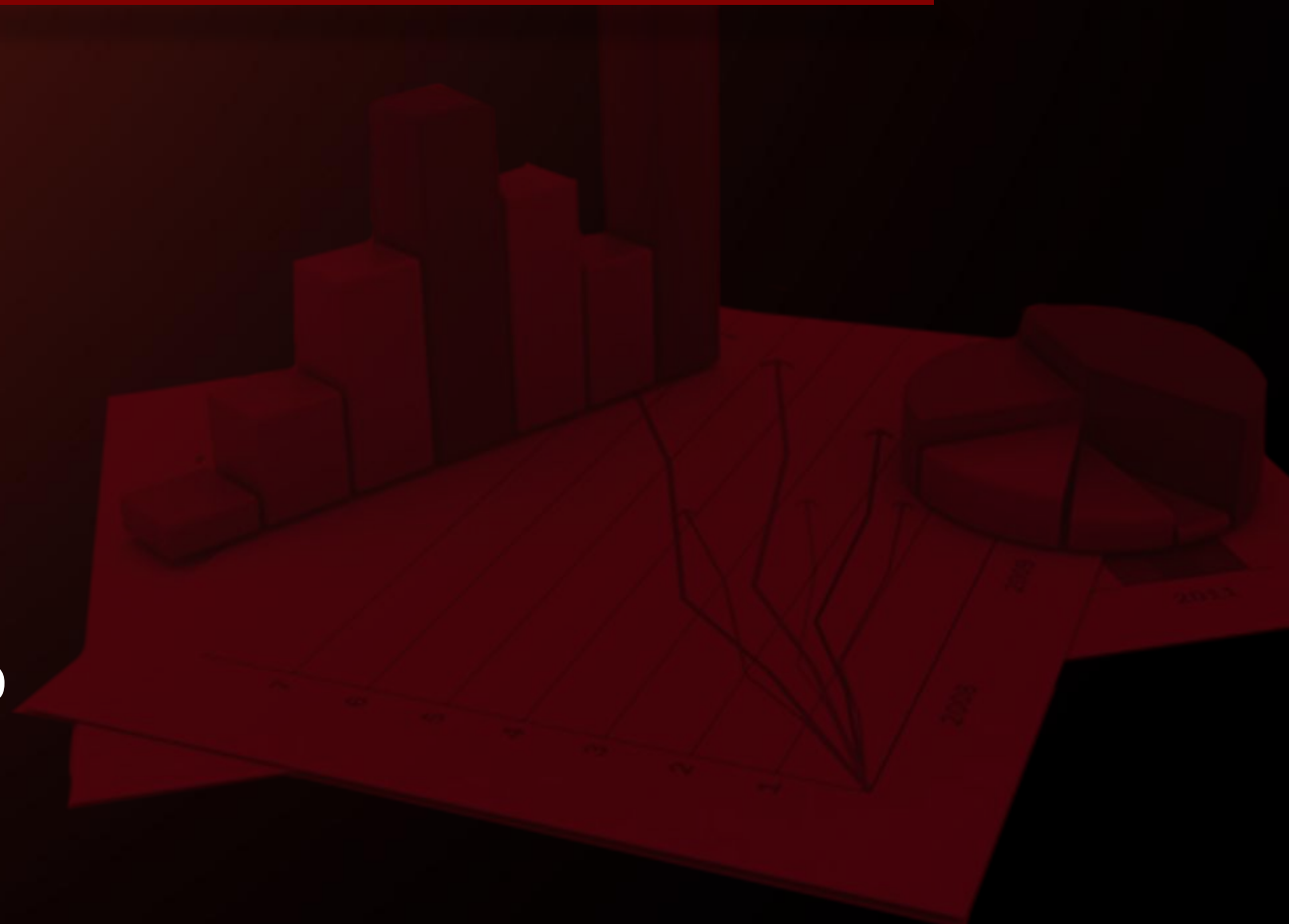


# MODELO DE REGRESSÃO LINEAR SIMPLES

$$y = \beta_0 + \beta_1 x_1$$

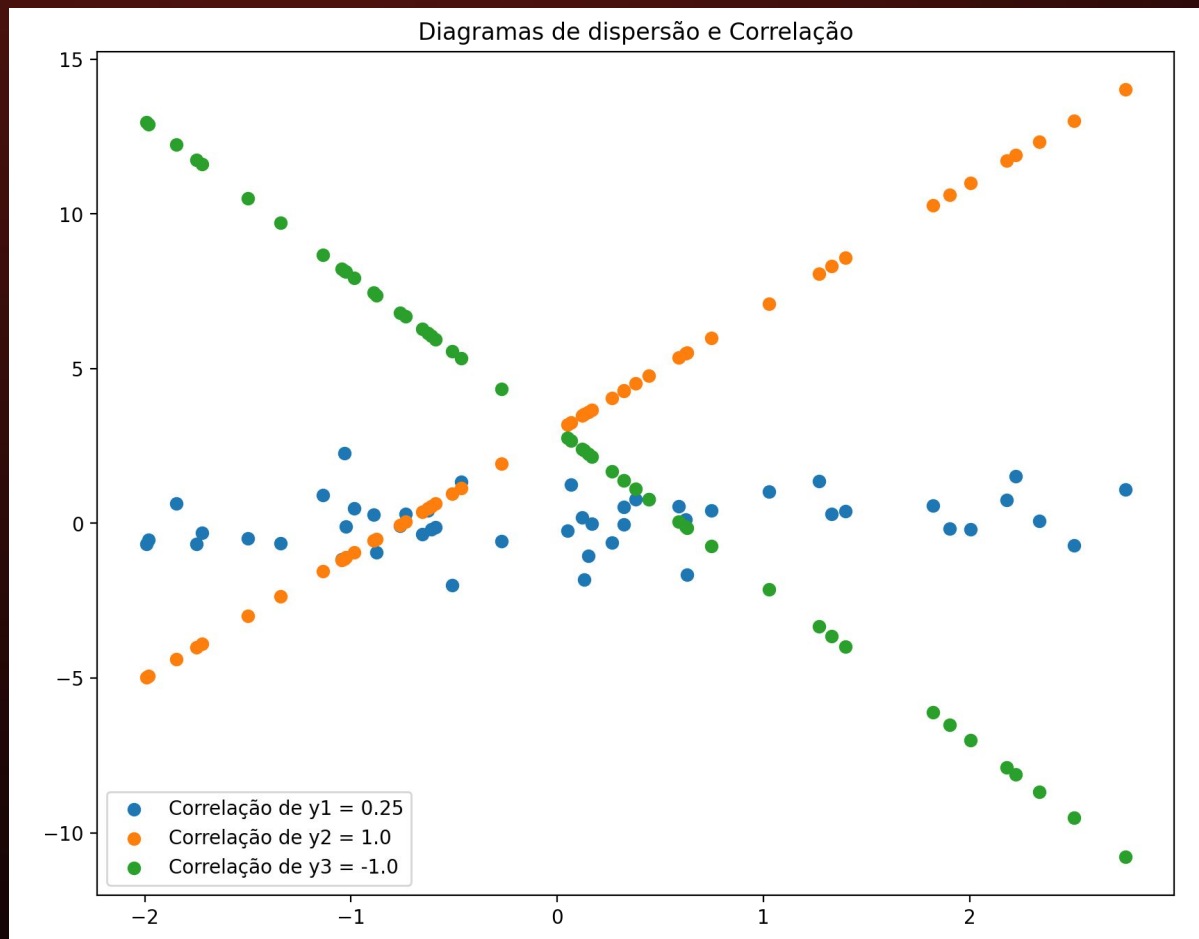
Intercepto

Inclinação  
+/-



# ORGANIZAÇÃO DE DADOS

$$y_1 = rand(x) ; y_2 = 3 + 4x ; y_3 = 3 - 4x$$



## EXEMPLO 1 (DEVORE, 2018, P. 456)

Problemas visuais e musculoesqueléticos associados ao uso de terminais de exibição visual (*Visual Display Terminals* – VDT) tornaram-se muito comuns nos últimos anos. Alguns pesquisadores concentraram-se na direção vertical do olhar como uma fonte de esforço e irritação ocular. Acredita-se que essa direção esteja intimamente relacionada com a Área da Superfície Ocular (ASO), de modo que é necessário um método para medi-la. Os dados representativos a seguir sobre  $y = \text{ASO (cm}^2\text{)}$  e  $x = \text{largura da fissura da pálpebra (isto é, a largura horizontal da abertura do olho, em cm)}$  foram reproduzidos do artigo “*Analysis of ocular surface area for comfortable VDT workstation layout*” (Ergonomics, 1996: 877-884). A ordem na qual as observações foram obtidas não foi dada, de maneira que, por conveniência, são relacionadas na ordem crescente dos valores  $x$



# EXEMPLO 1: DADOS

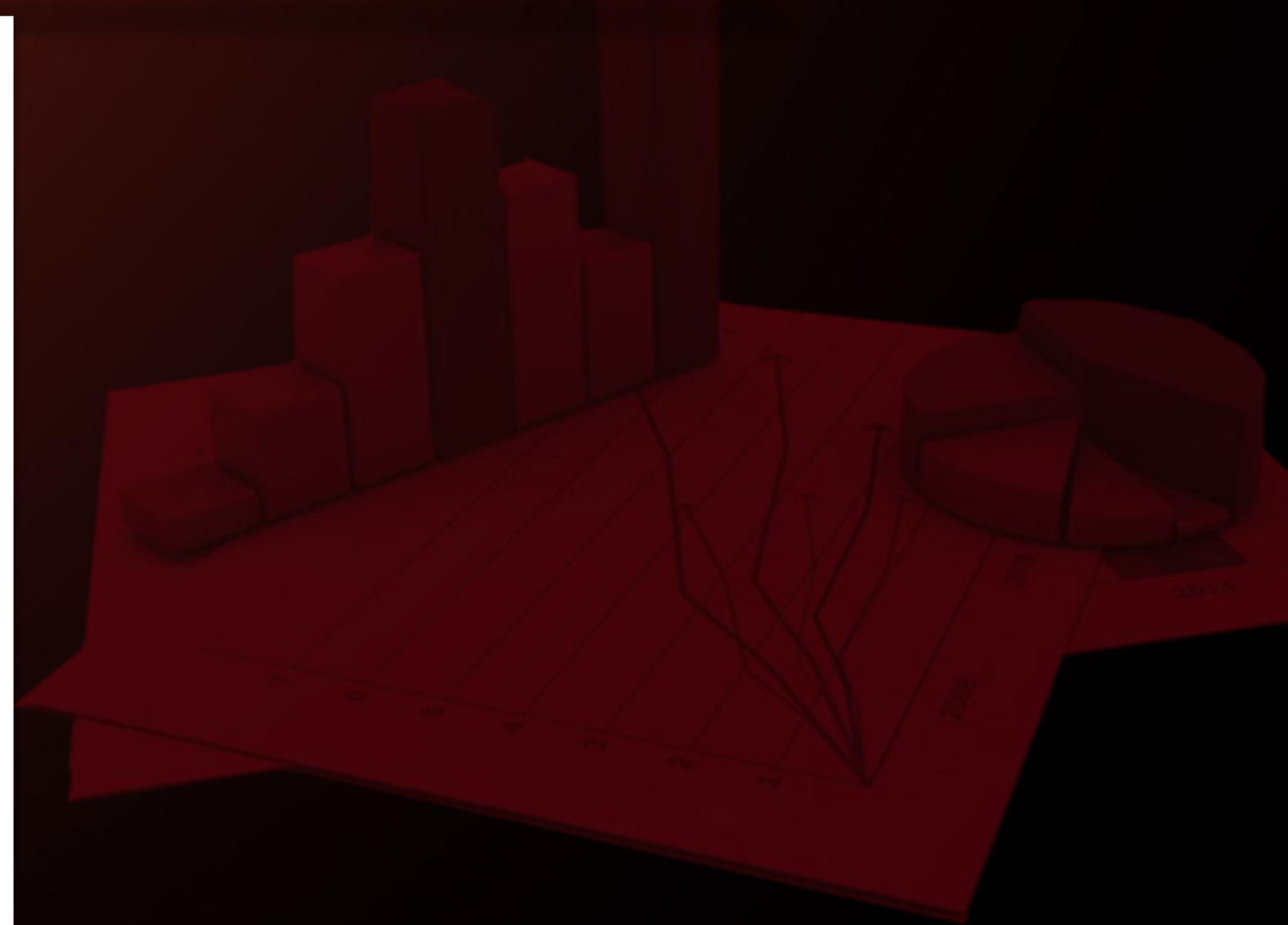
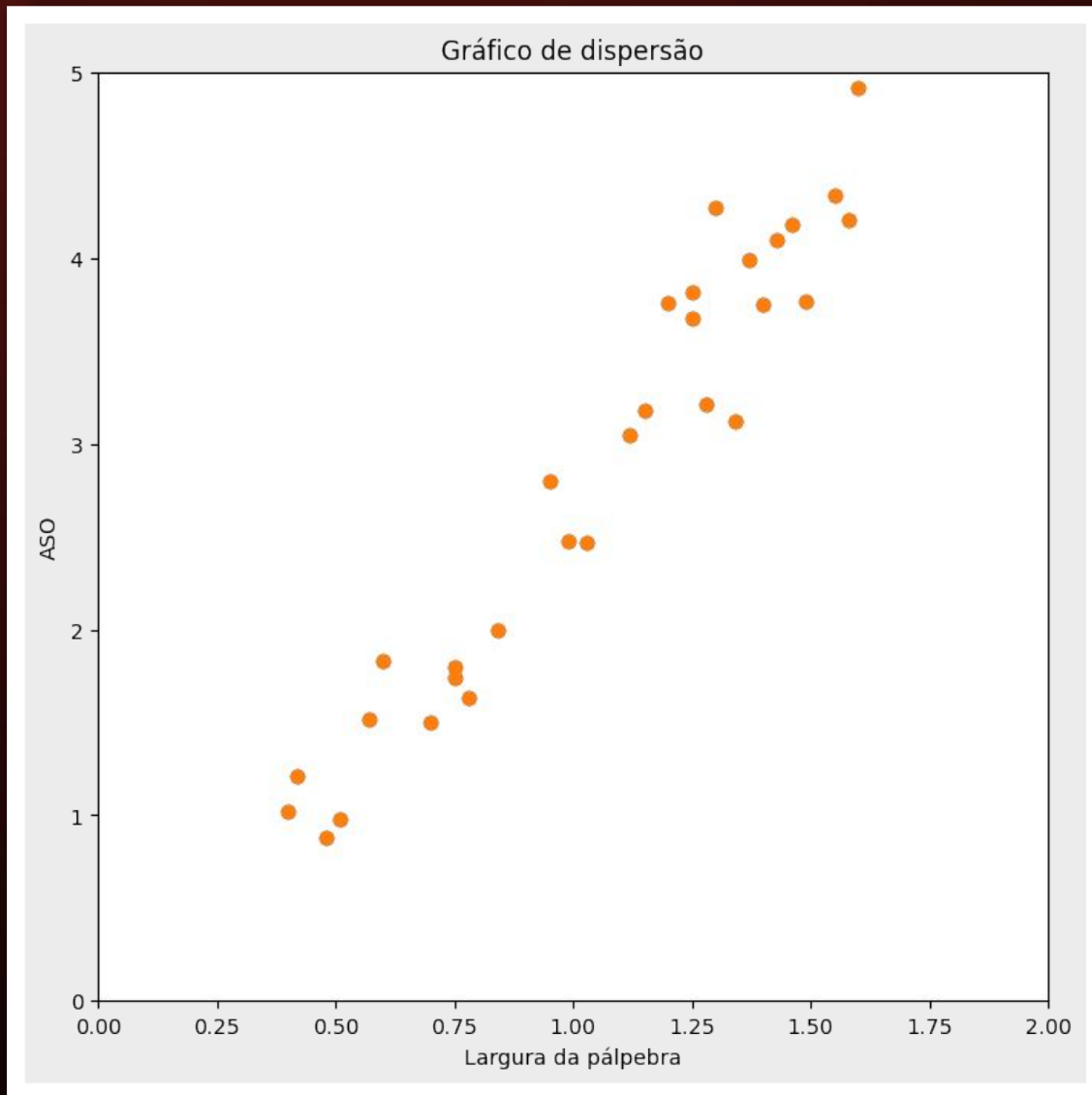
x	0,4	0,42	0,48	0,51	0,57	0,6	0,7	0,85	0,75	0,75	0,84	0,95	0,99	1,03	1,12
y	1,02	1,21	0,88	0,98	1,52	1,83	1,5	1,8	1,74	1,63	2	2,8	2,48	2,47	3,05
x	1,15	1,2	1,25	1,25	1,28	1,3	1,34	1,37	1,4	1,43	1,46	1,49	1,55	1,58	1,6
y	3,18	3,76	3,68	3,82	3,21	4,27	3,12	3,99	3,75	4,1	4,18	3,77	4,34	4,21	4,92

Dados:

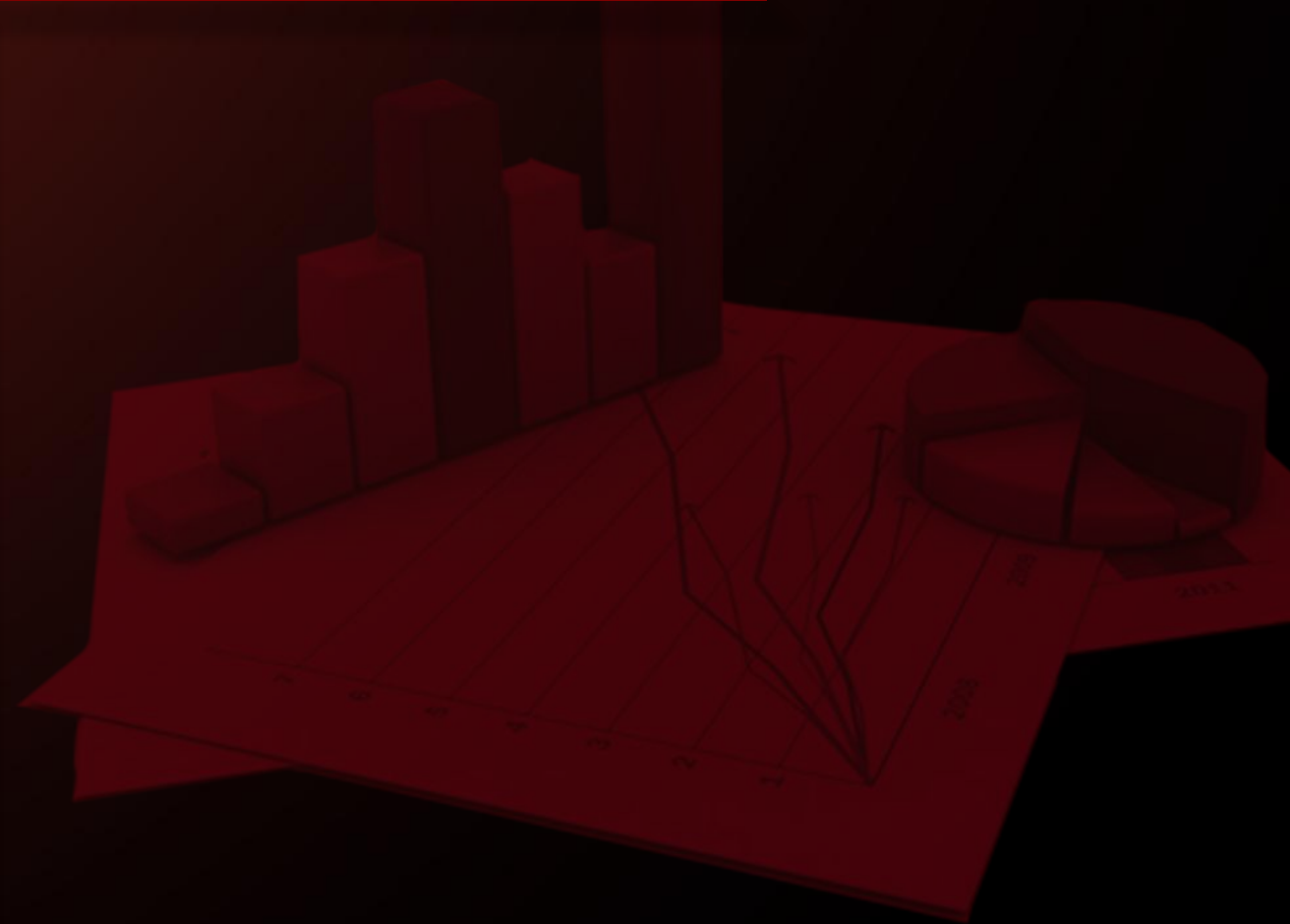
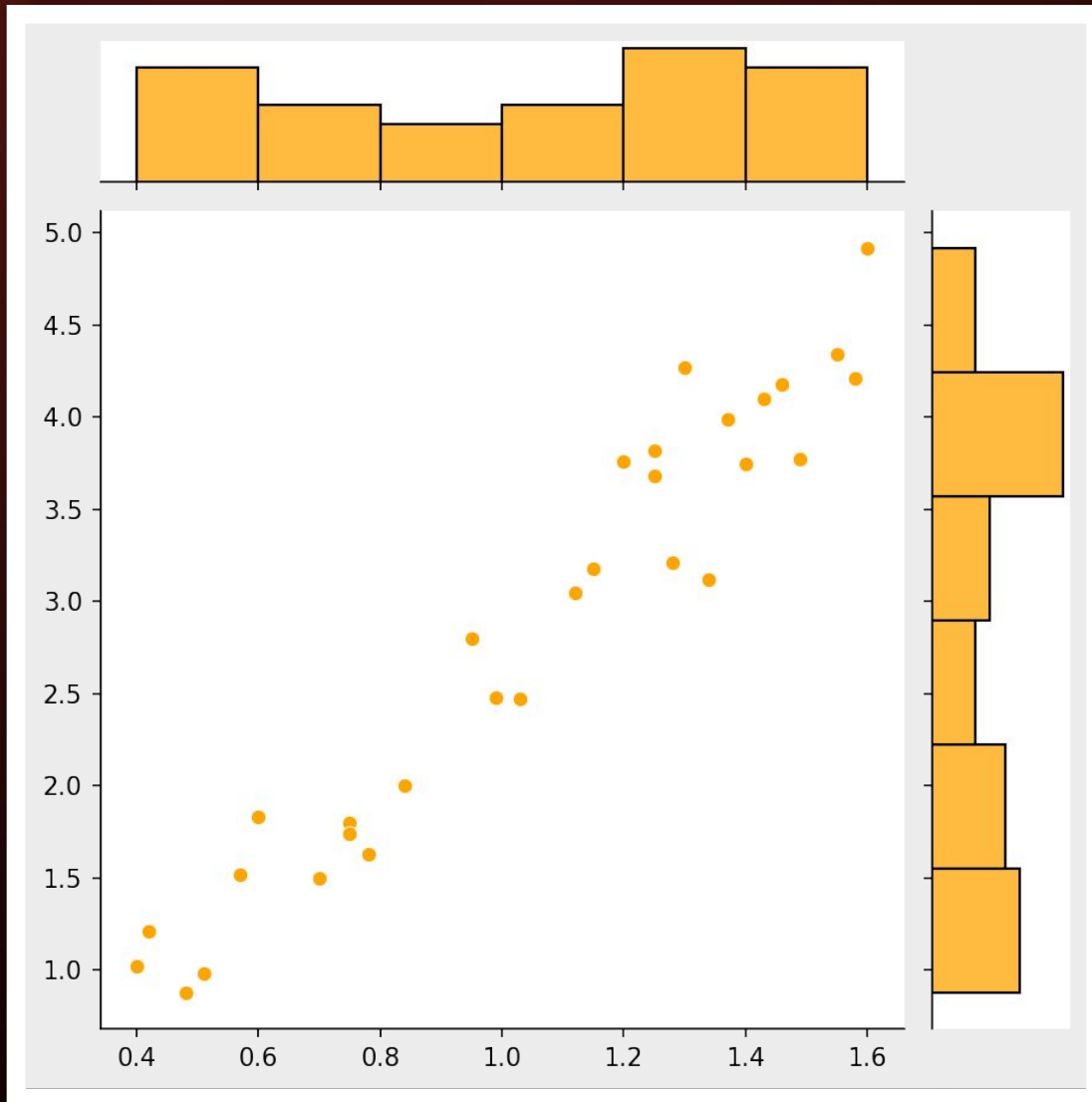
y = Área da Superfície Ocular, ASO (cm<sup>2</sup>)

x = Largura da fissura da pálpebra (isto é, a largura horizontal da abertura do olho, em cm)

# EXEMPLO 1: DIAGRAMA DE DISPERSÃO



# EXEMPLO 1: DIAGRAMA DE DISPERSÃO



## EXEMPLO 2 (DEVORE, 2018, P. 457)

O arsênio é encontrado em muitas águas subterrâneas e algumas águas de superfície. Uma pesquisa recente sobre os efeitos de saúde levou a Agência de Proteção Ambiental a reduzir os níveis de arsênio permitidos na água potável, de modo que muitos sistemas de água não são mais compatíveis com os padrões.

Isso estimulou o interesse no desenvolvimento dos métodos para remover o arsênio. Os dados a seguir ( $x$  = pH e  $y$  = arsênio removido (%) por um determinado processo) foram lidos em um gráfico de dispersão no artigo

“Optimizing arsenic removal during iron removal: theoretical and practical considerations”

(J. of Water Supply Res. and Tech., 2005: 545-560).

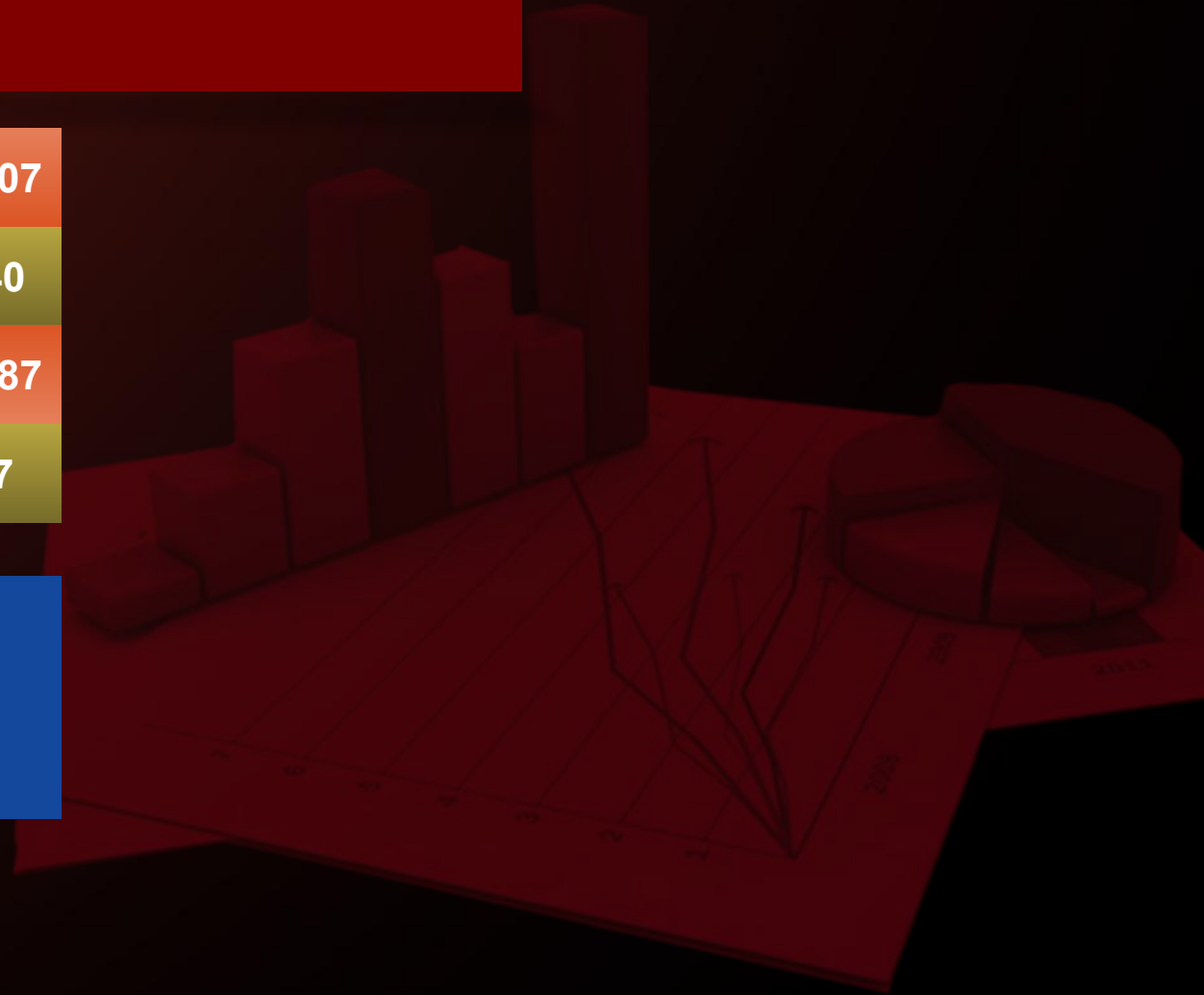
## EXEMPLO 2: DADOS

x	7,01	7,11	7,12	7,24	7,94	7,94	8,04	8,05	8,07
y	60	67	66	52	50	45	52	48	40
x	8,9	8,94	8,95	8,97	8,98	9,95	9,86	9,86	9,87
y	23	20	40	31	26	9	22	13	7

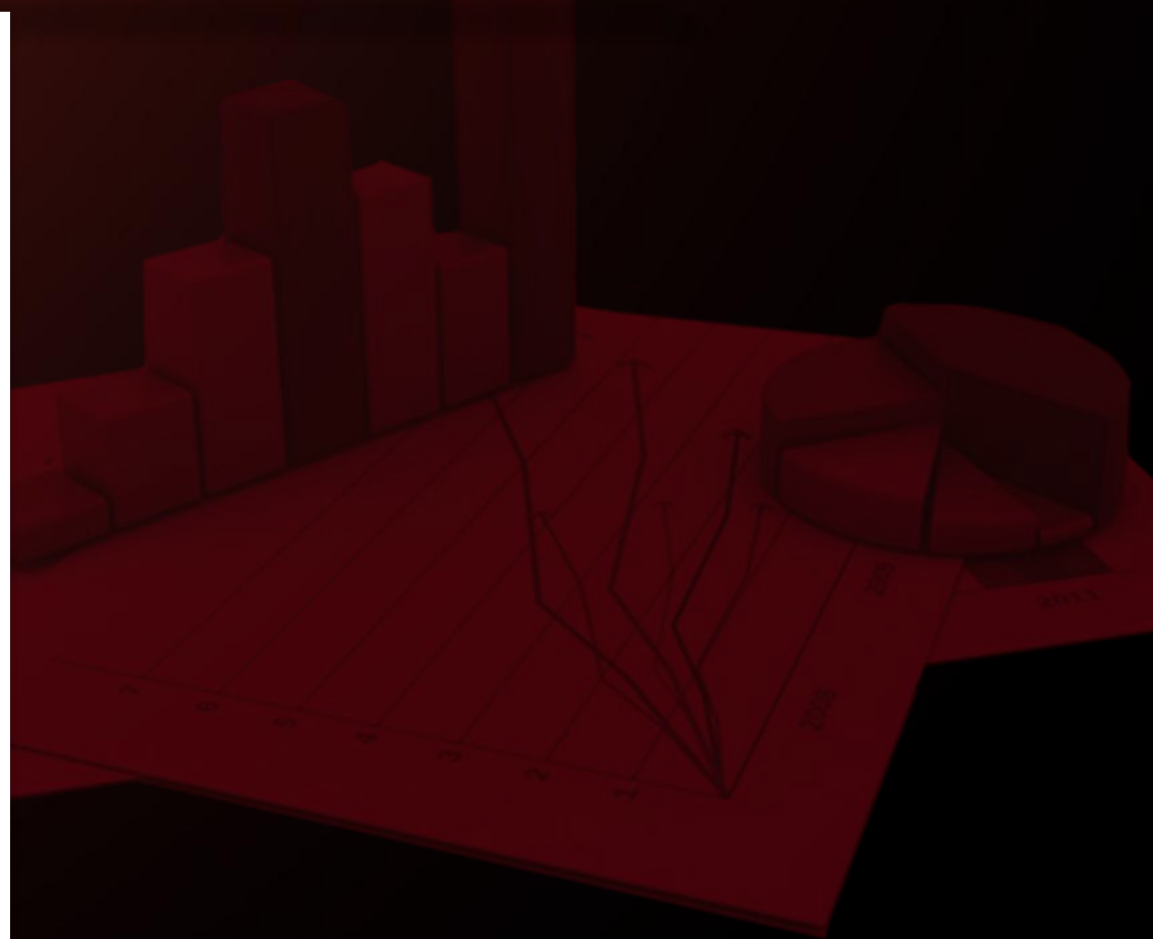
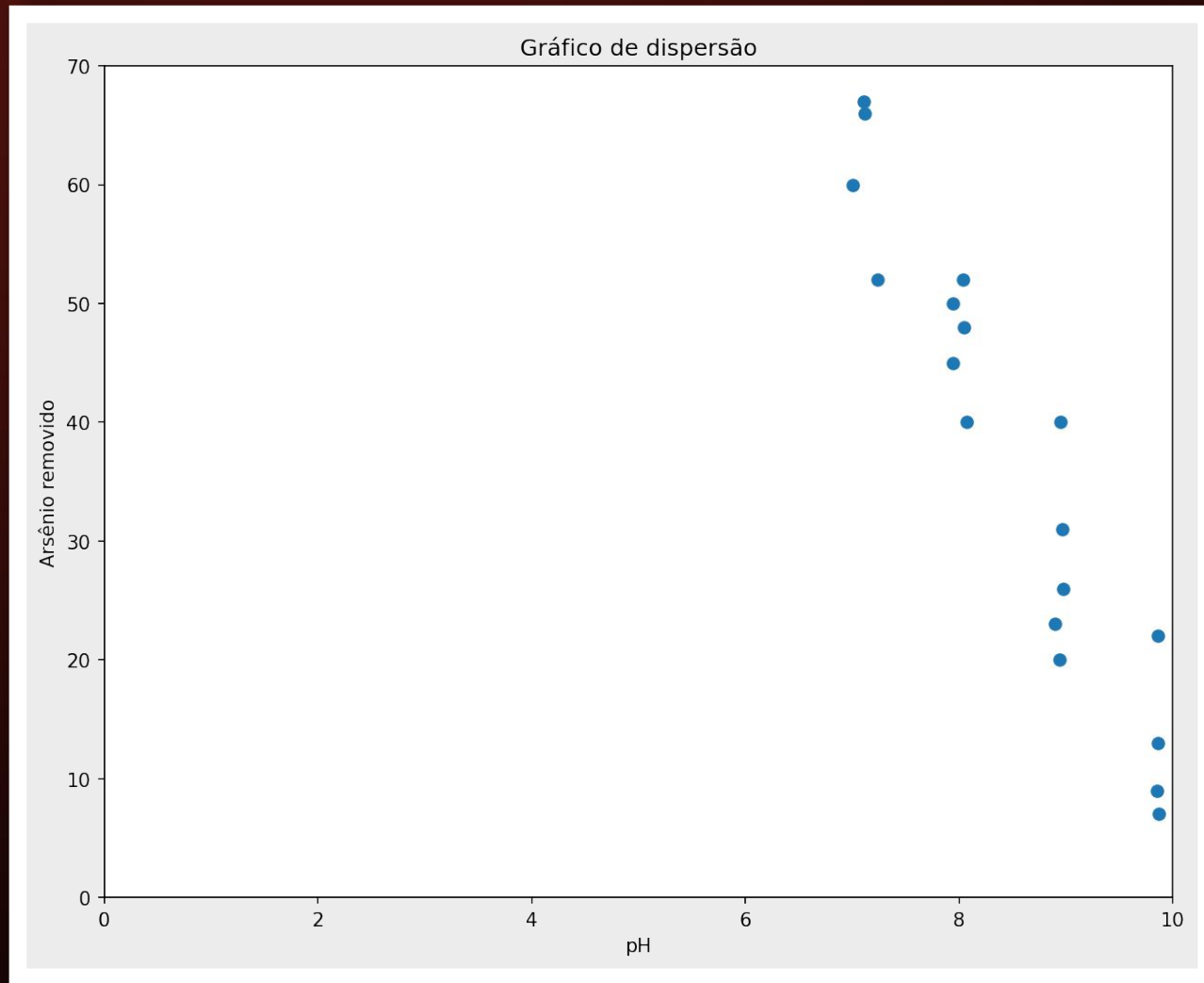
Dados:

x = pH

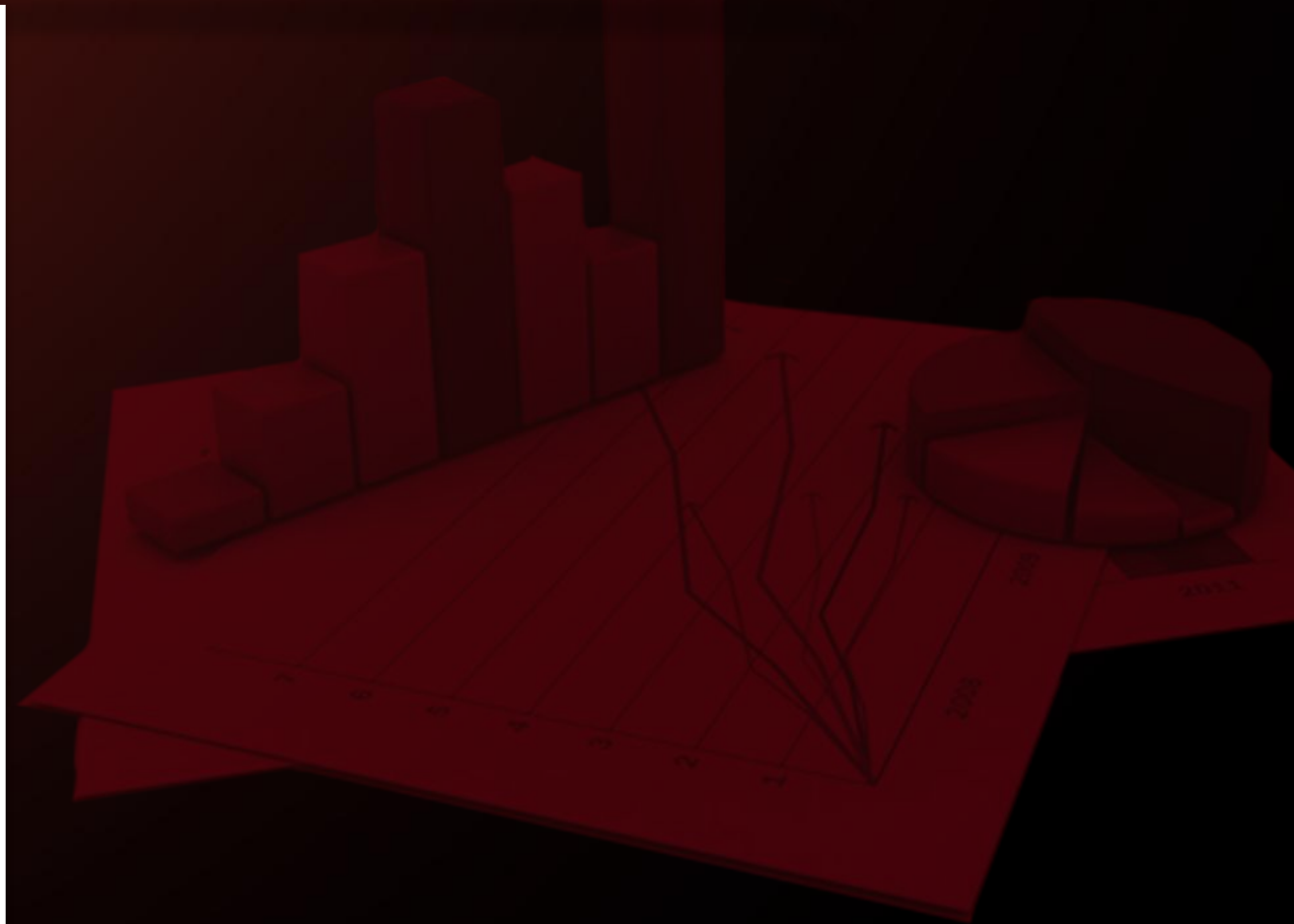
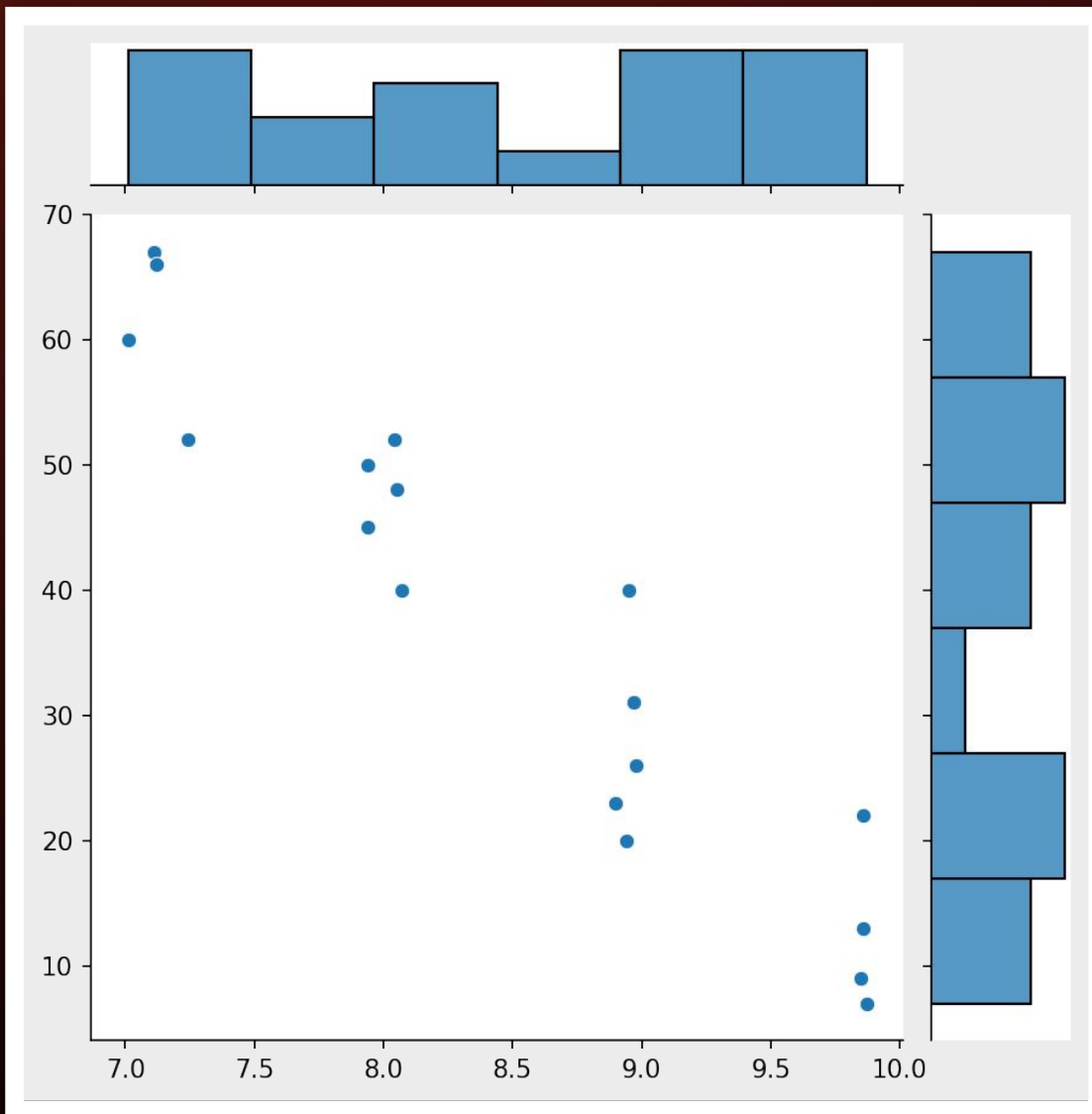
y = arsênio removido (%)



# EXEMPLO 2: DIAGRAMA DE DISPERSÃO



## EXEMPLO 2: DIAGRAMA DE DISPERSÃO



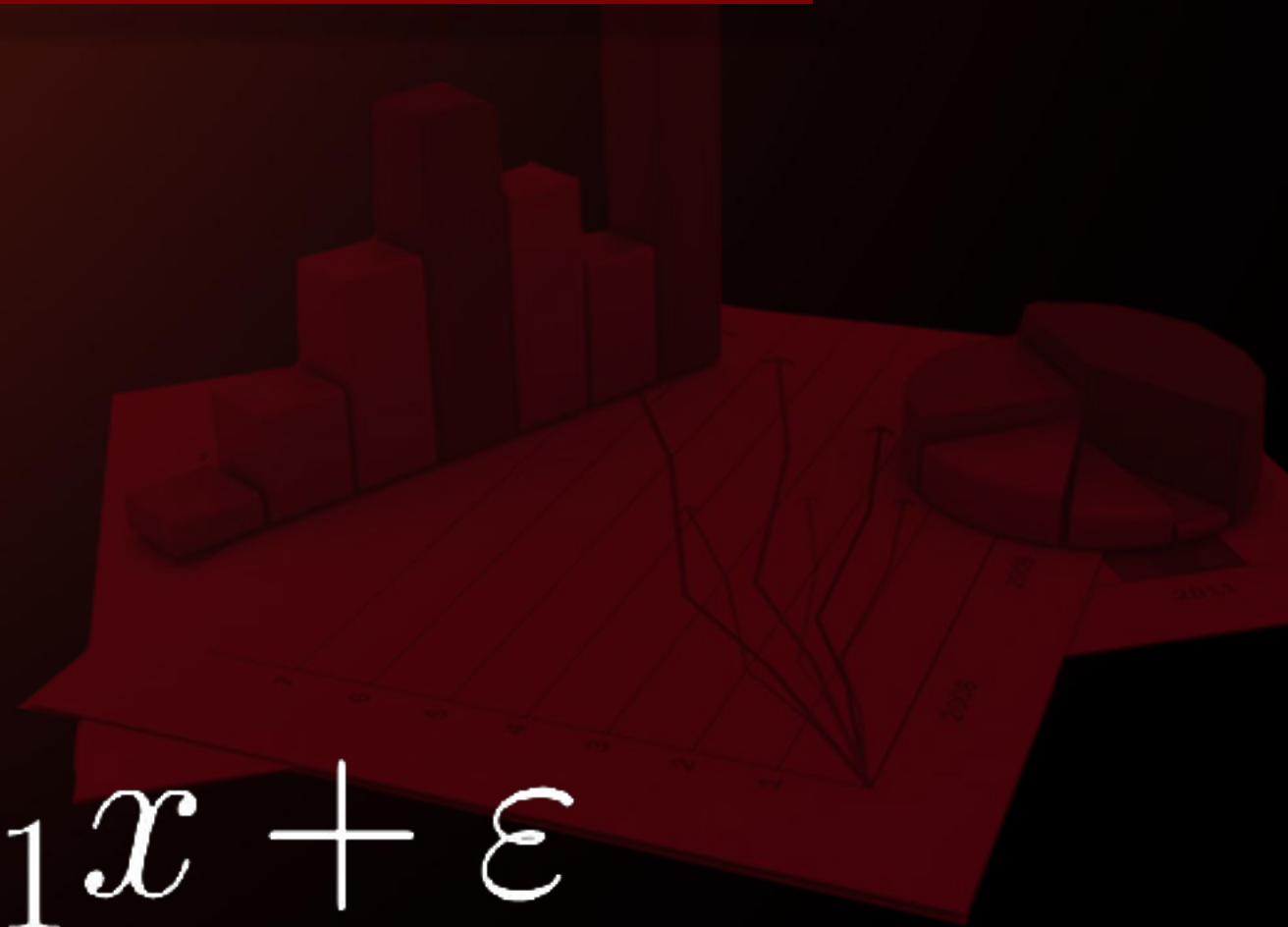
# MODELO PROBABILÍSTICO LINEAR

➤ Modelo de regressão simples

$$y = \beta_0 + \beta_1 x_1$$

➤ Modelo probabilístico

$$Y = \beta_0 + \beta_1 x + \varepsilon$$





# MODELO PROBABILÍSTICO LINEAR

➤ Modelo probabilístico

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Variável aleatória  
Variável dependente

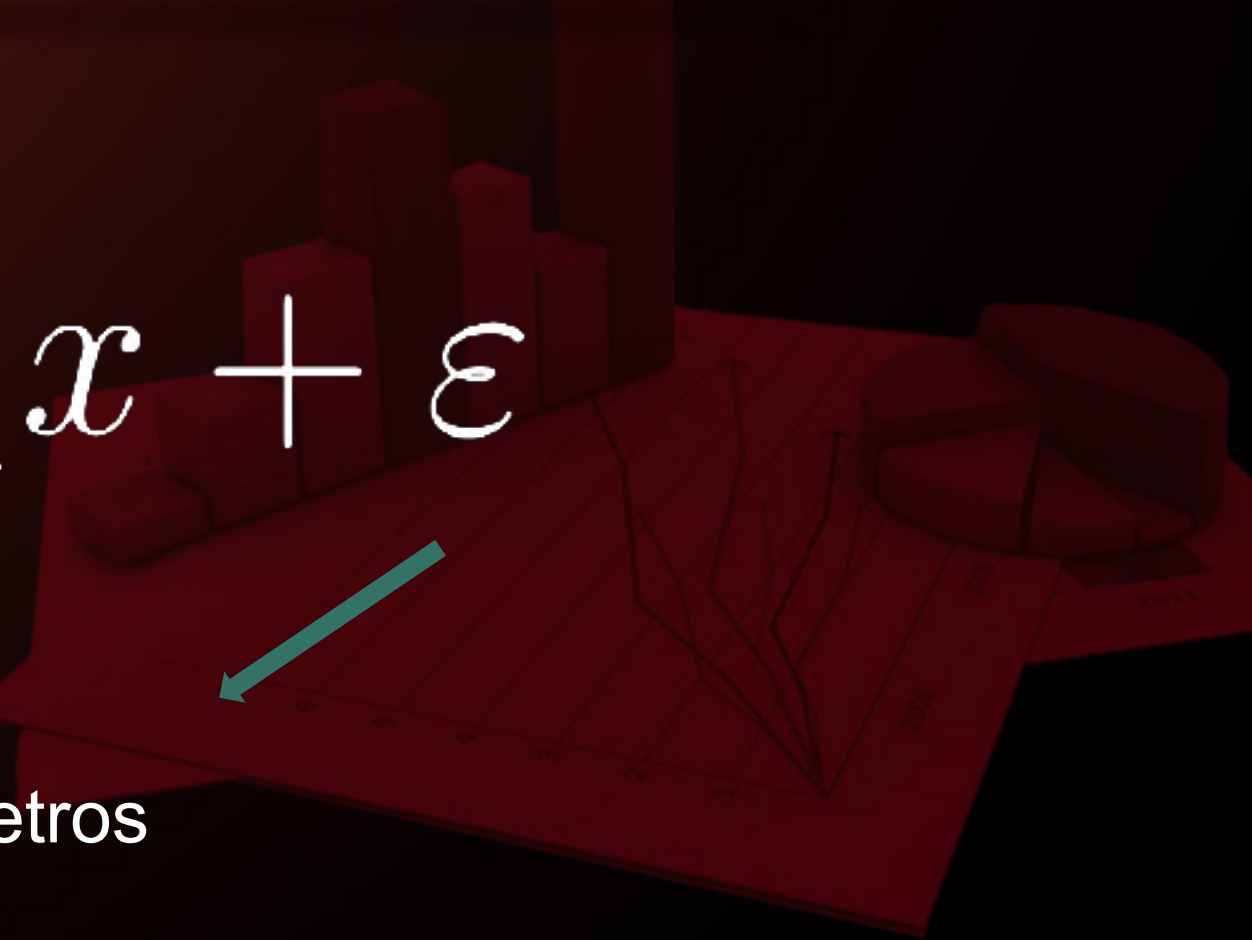
Variável  
independente

# MODELO PROBABILÍSTICO LINEAR

➤ Modelo probabilístico

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Parâmetros



# MODELO PROBABILÍSTICO LINEAR

➤ Modelo probabilístico

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

- ✓ Desvio aleatório ou termo do erro aleatório
- ✓ Variável aleatória com distribuição normal, com  $E(\varepsilon) = 0$  e  $V(\varepsilon) = \sigma^2$

# DISTRIBUIÇÕES DE PROBABILIDADE

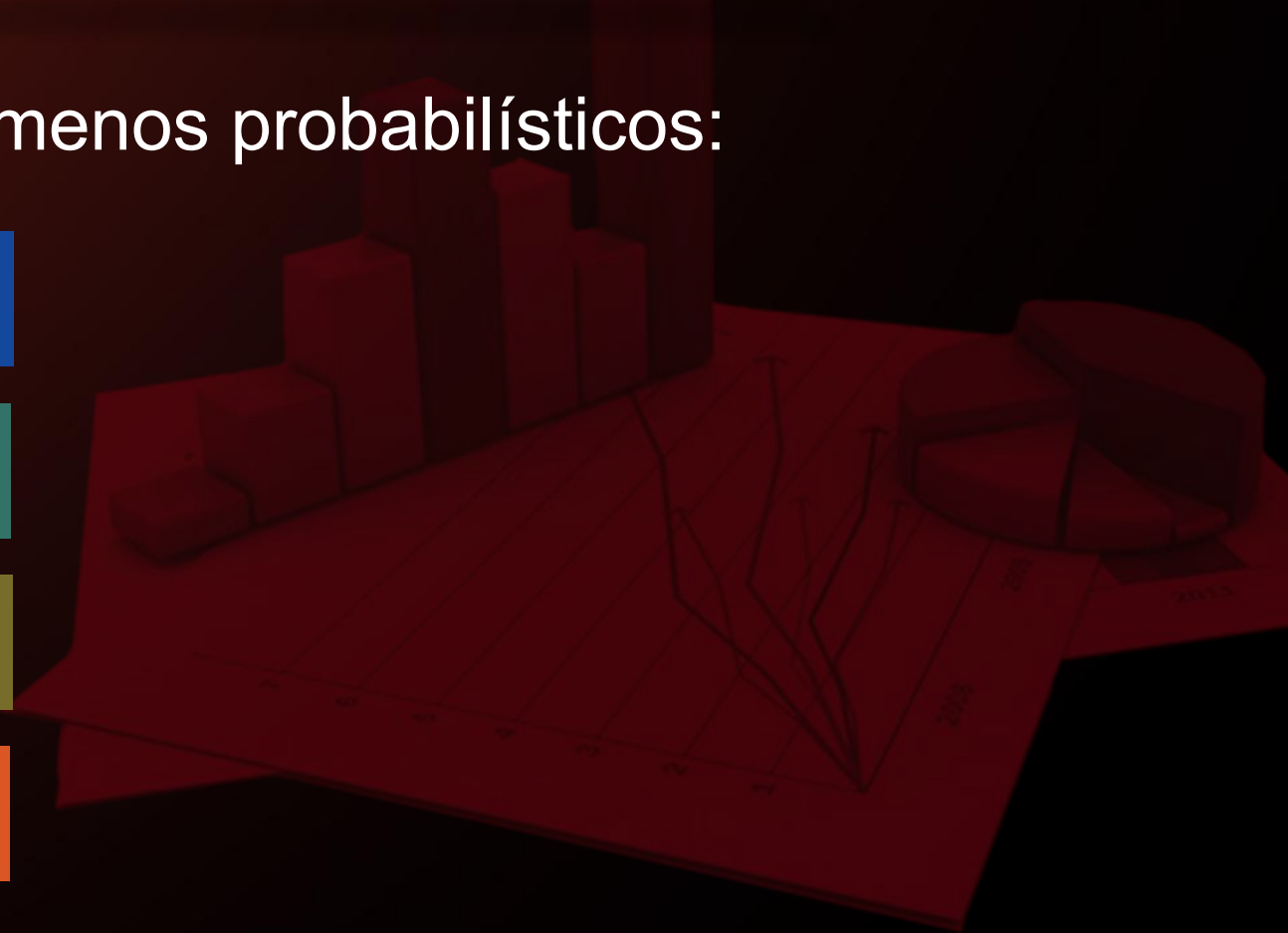
➤ Distribuições descrevem fenômenos probabilísticos:

Distribuição de Bernoulli.

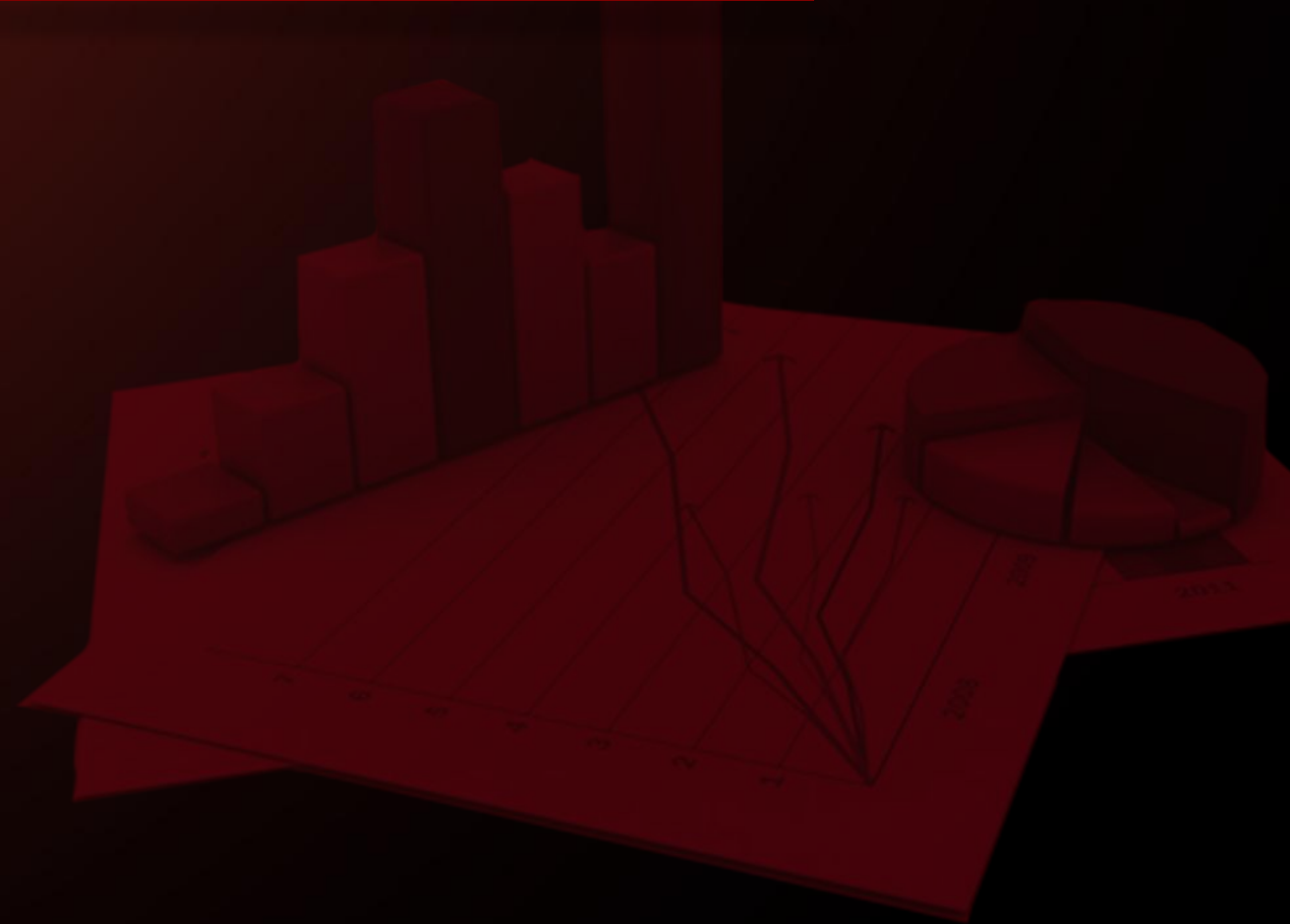
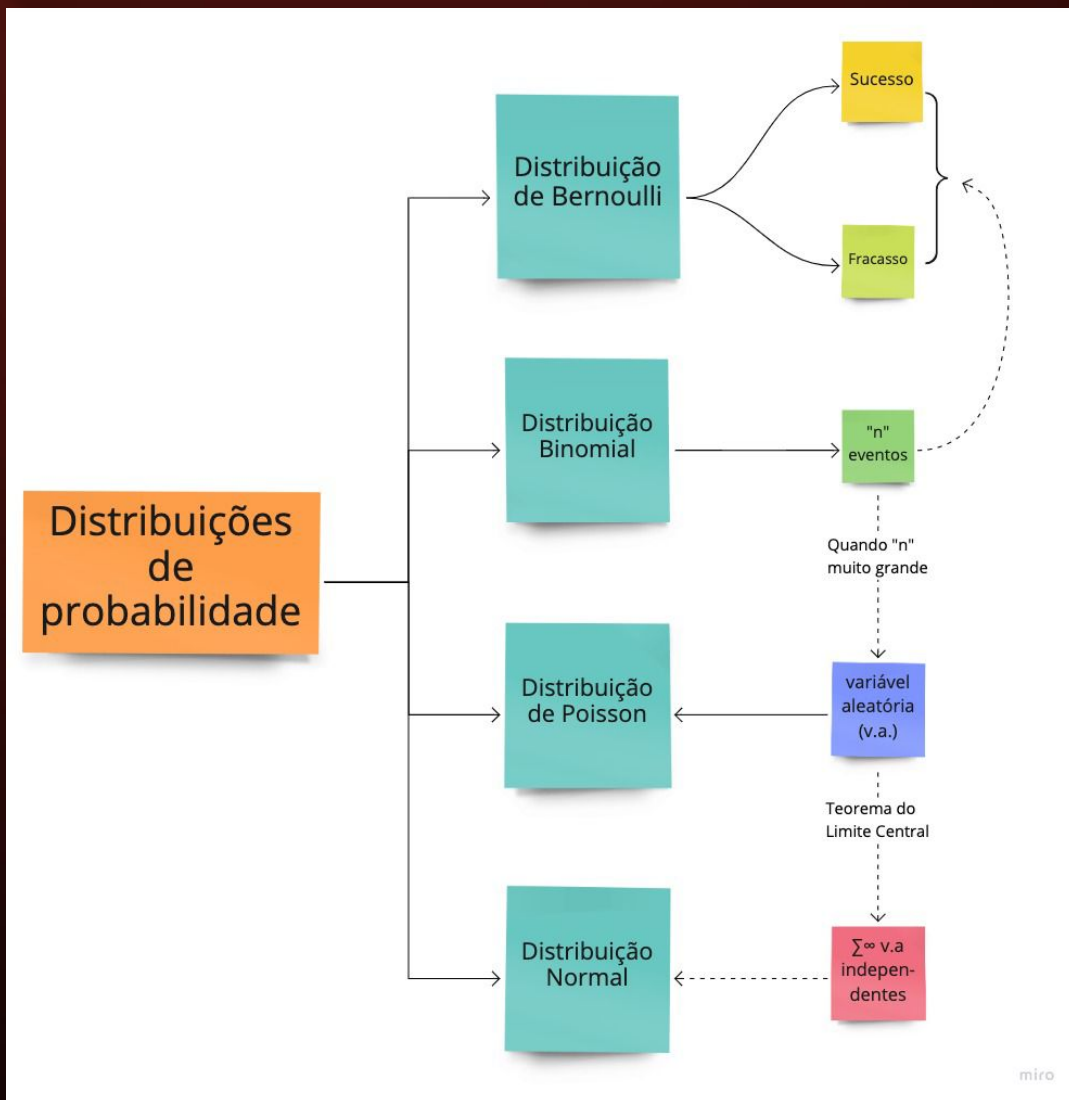
Distribuição Binomial.

Distribuição de Poisson.

Distribuição Normal.



# DISTRIBUIÇÕES DE PROBABILIDADE



# DISTRIBUIÇÃO NORMAL

➤ Onde se aplica:

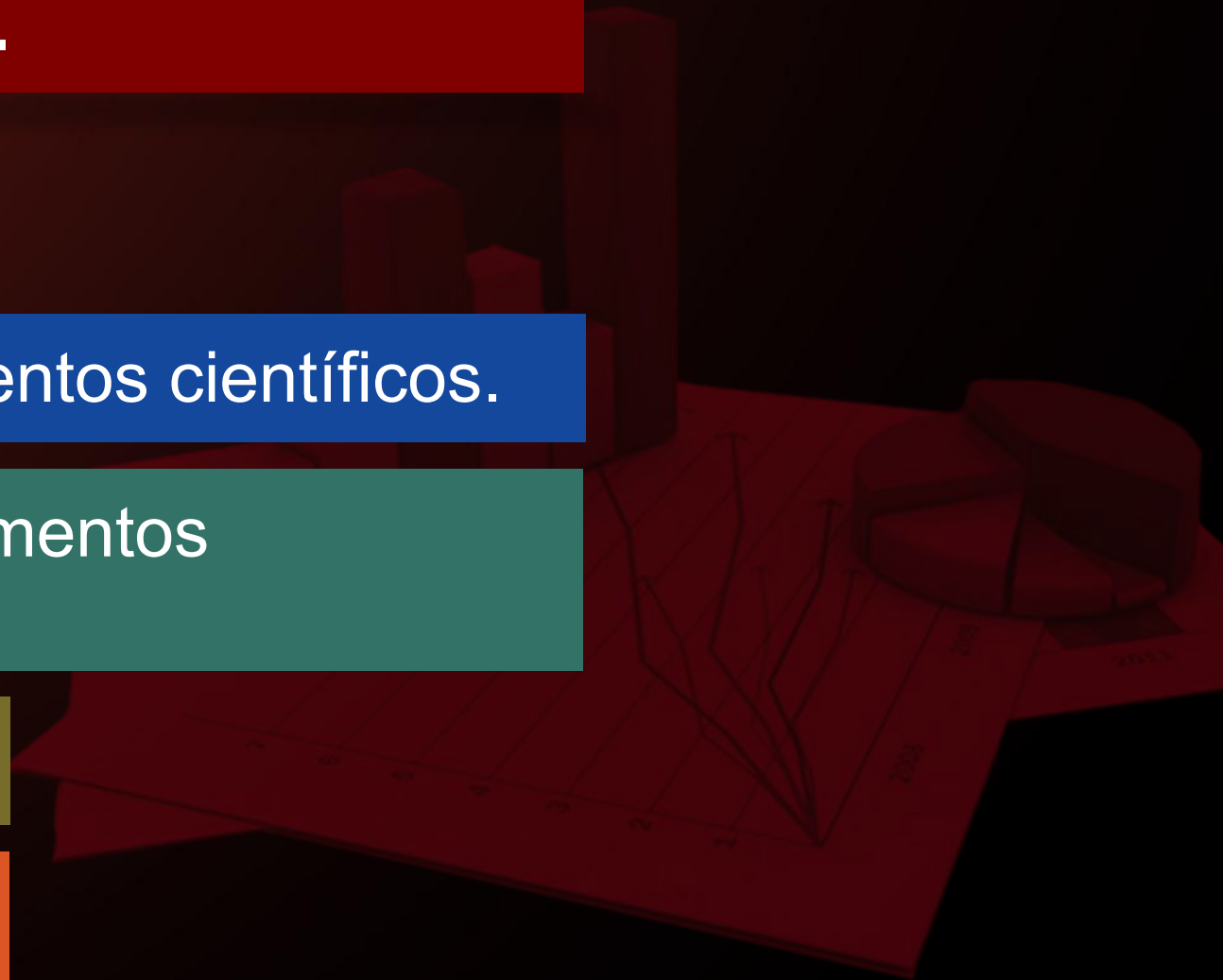
Erro de medida em experimentos científicos.

Tempo de reação em experimentos psicológicos.

Medidas de aptidão.

Indicadores econômicos.

Pontuações em testes.



# TERMOS IMPORTANTES

Função de distribuição (densidade) de probabilidade

População → amostra → variável

Variável aleatória (v.a.)

- Discreta ou Contínua

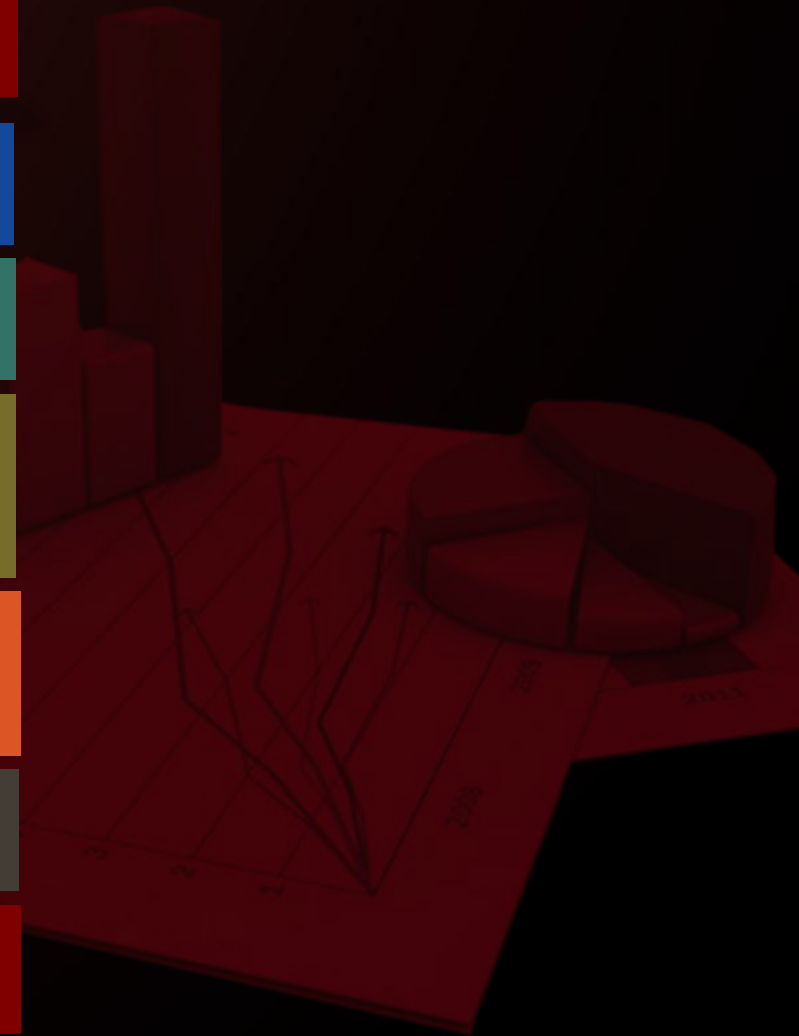
Dados ← Observação de uma variável

- Univariados, bivariados, multivariados

Variância

Desvio padrão

Valor médio ou valor esperado



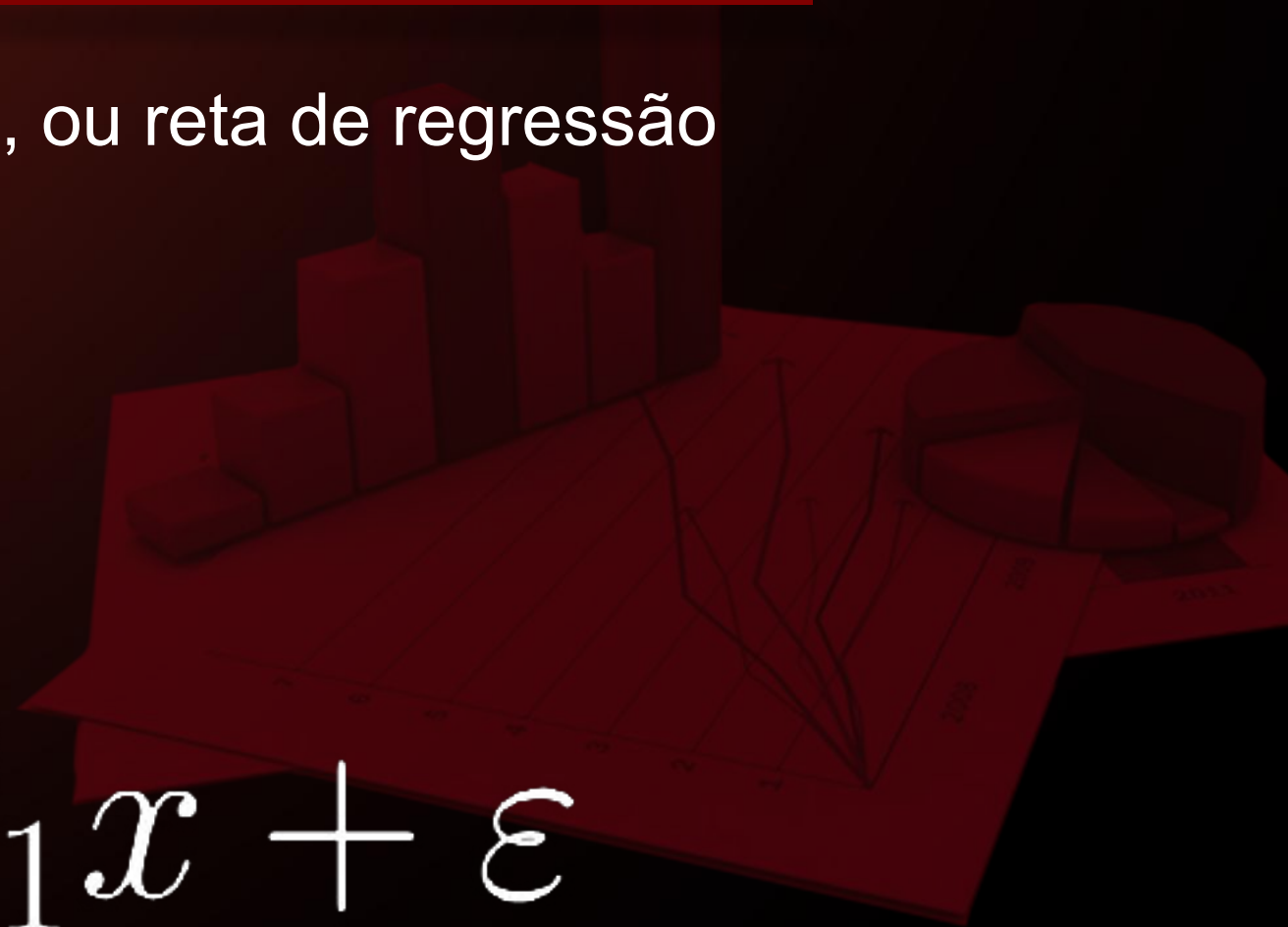
# MODELO PROBABILÍSTICO LINEAR

- Modelo de regressão simples, ou reta de regressão verdadeira (ou populacional)

$$y = \beta_0 + \beta_1 x_1$$

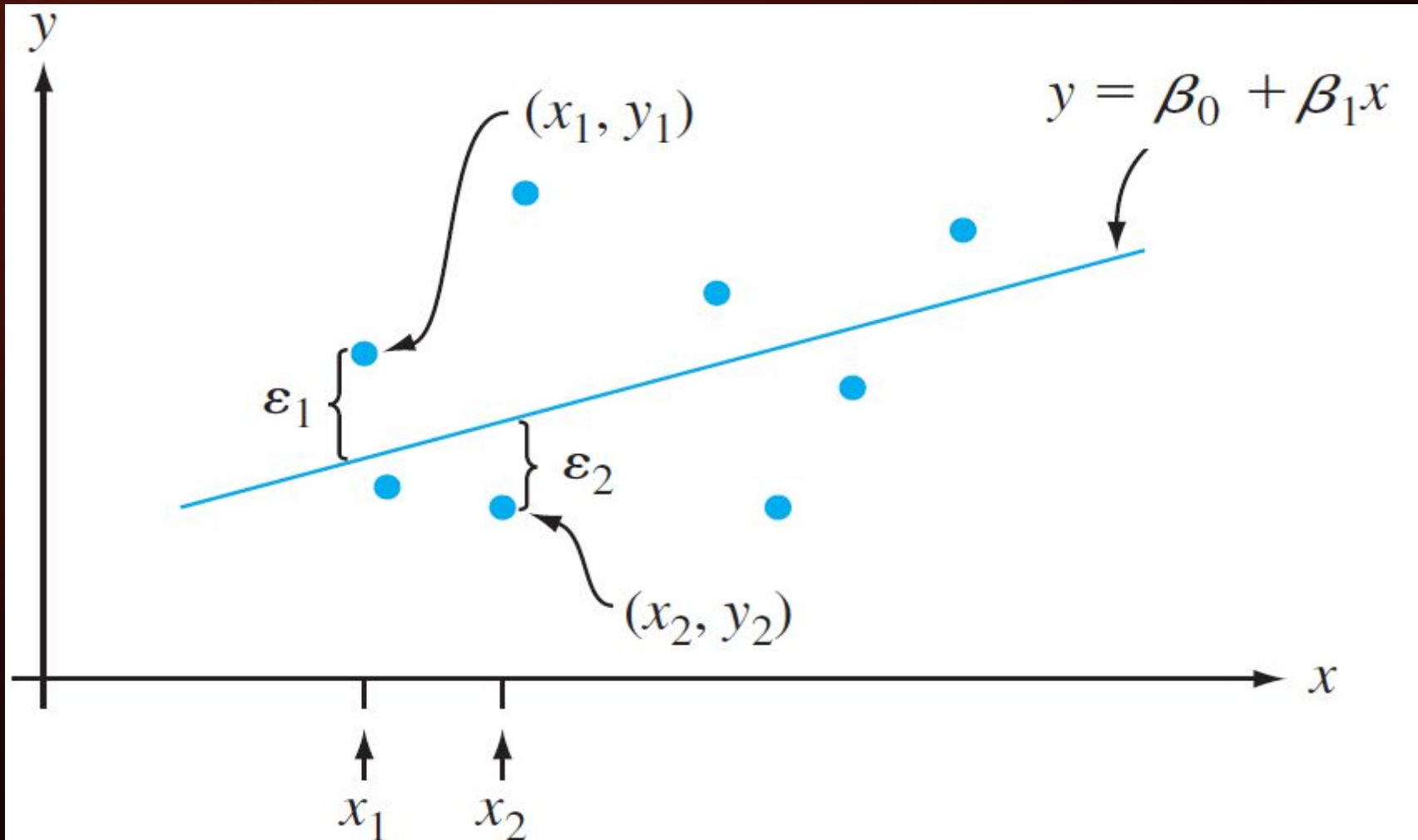
- Modelo probabilístico

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

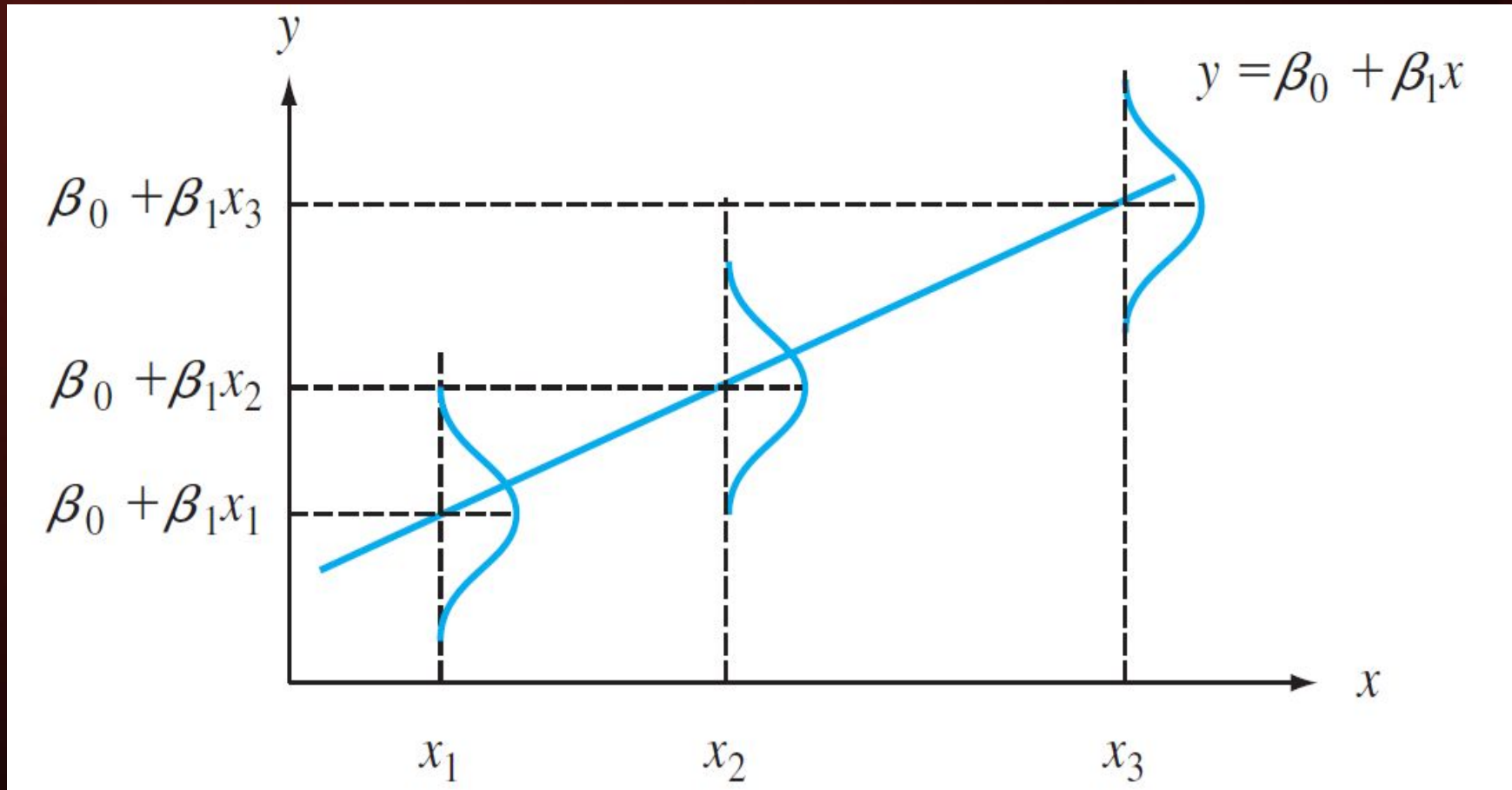




# MODELO PROBABILÍSTICO LINEAR



# MODELO PROBABILÍSTICO LINEAR



# MODELO PROBABILÍSTICO LINEAR

$\mu_{y,x^*} =$  Valor esperado de Y quando x adota o valor  $x^*$

$\sigma^2_{y,x^*} =$  Variância de Y quando x adota o valor  $x^*$

Ex.:  $\mu_{y,20} = 250$  e  $\sigma^2_{y,20} = 2,15$