

# Regressão Logística

Disciplina: Modelos Estatísticos  
Professora: Jéssica Assunção

# Comparação de modelos

## Critério de Informação de Akaike - AIC

$$AIC = -2 \log L(\hat{\theta}) + 2(p)$$

$$AIC = -n \ln(SQ_{Res}/n) + 2p.$$

## Critério de Informação Bayesiano - BIC

$$BIC = -2 \log f(x_n | \theta) + p \log n,$$

$$BIC = -n \ln(SQ_{Res}/n) + \ln(n)p.$$

# MAE & MSE

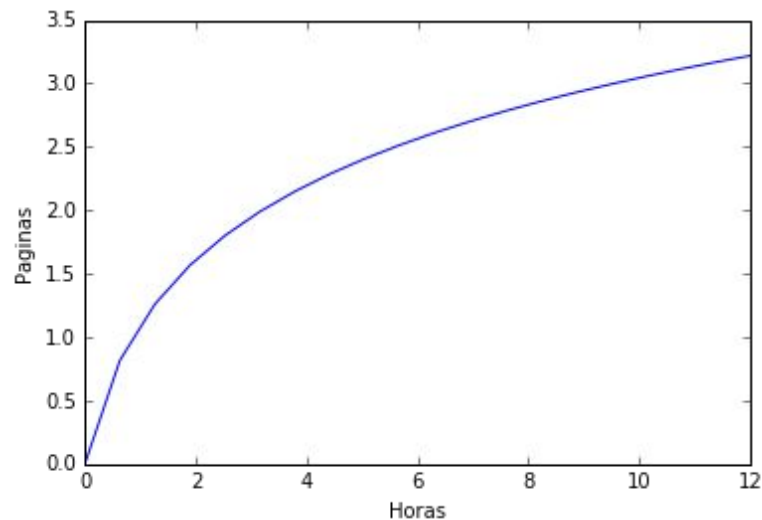
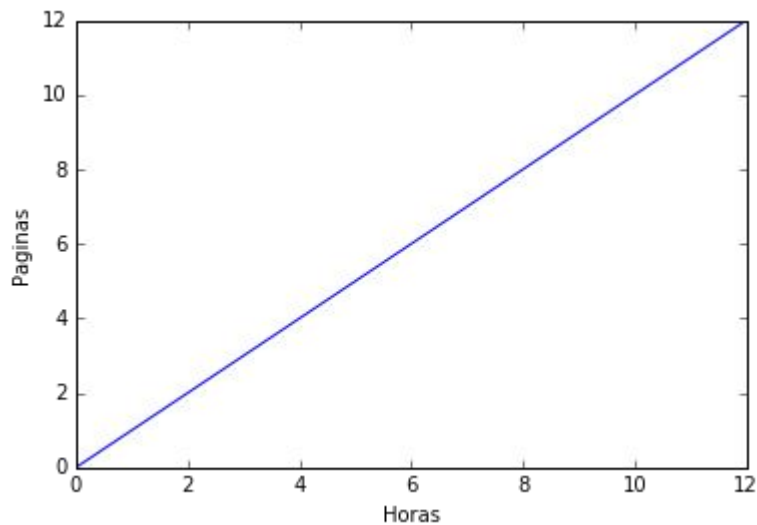
**MAE - Erro médio absoluto**

$$\frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j|$$

**MSE - Erro médio quadrático**

$$\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2$$

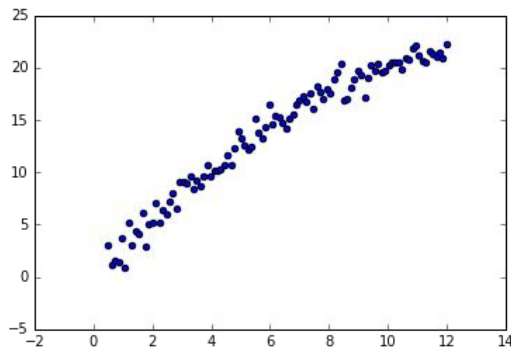
# Regressão Polinomial



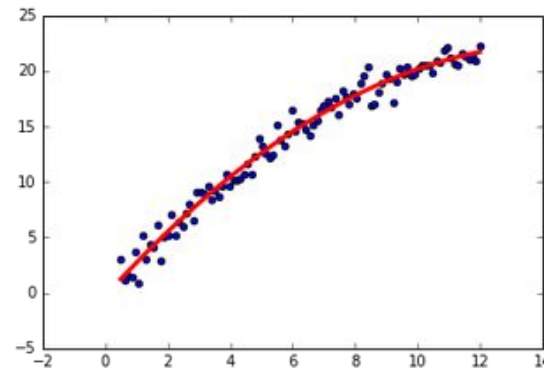
# Regressão Polinomial

Contornando o problema com modelos lineares!  
Vamos forjar novas variáveis...

$$y = 3x - 0,1x^2 + \epsilon$$



	x	x^2	y
0	0.500000	0.250000	3.099345
1	0.616162	0.379655	1.198763
2	0.732323	0.536297	1.615168
3	0.848485	0.719927	1.400493
4	0.964646	0.930543	3.666293



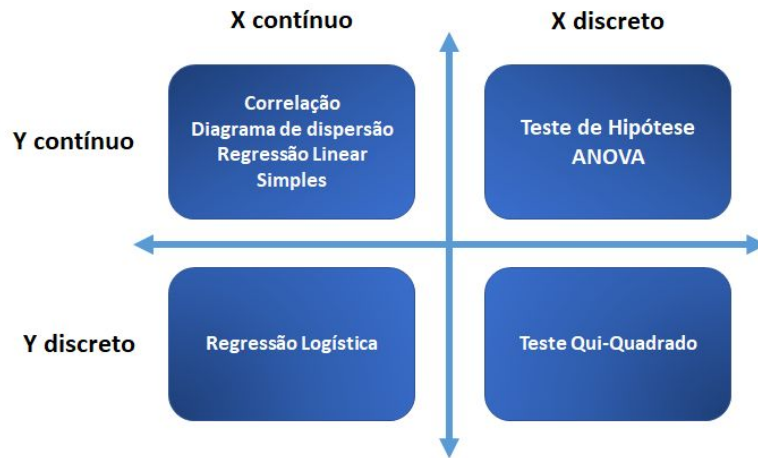
# Regressão Polinomial

## Algumas Desvantagens:

- O algoritmo não aprende sobre a forma da não linearidade
- Responsabilidade atribuída ao analista
- Não performa bem com muitas variáveis
- Não performa bem quando o grau do polinômio é muito alto

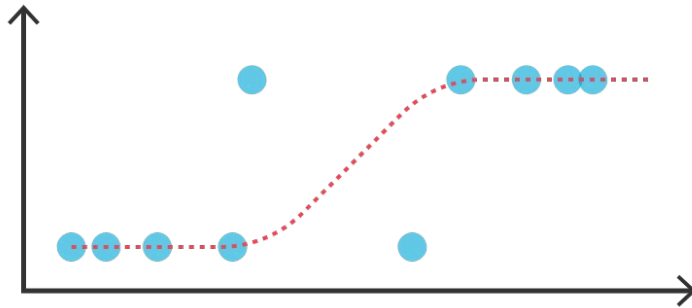
# Regressão Logística

# Regressão Logística





# Regressão Logística



## Diferenças

	Linear	Logística
Reta	Reta	Curva - S
Variável Dependente	Continua	Categórica
Interpretação	$\hat{y}$	$\ln\left(\frac{p}{1+p}\right)$

# Regressão Logística

$$\ln \left( \frac{p}{1+p} \right) = a + bX$$

$$\left( \frac{p}{1+p} \right) = e^{a+bX}$$

$$p = \left( \frac{e^{a+bX}}{1 + e^{a+bX}} \right)$$

# Regressão Logística

$$p = \left( \frac{1}{1 + e^{-a+bX}} \right)$$

$$p = \frac{1}{1 + e^{-(b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n)}}$$