

## Disciplina de Séries Temporais

### Índice

Capítulo 1:	Modelos de Regressão*	2
	1.1. Regressão Linear $Y = A + BX + e$	2
	1.2. Regressão Quadrática: $Y = A + BX + CX^2 + e$	13
	1.3. Regressão Múltipla	20
Capítulo 2	Séries Temporais	28
	2.1. Definição	29
	2.2. Objetivo	29
	2.3. Fundamento Probabilístico	30
	2.4. Previsão de Séries Temporais	36
	2.5. Métodos de Previsão de Séries Temporais	38
	2.6. Métodos de Decomposição de Séries Temporais	42
	2.7. Métodos Simples de Previsão de Séries Temporais	43
	2.7.1. Média Móvel	43
	2.7.2. Alisamento Exponencial Simples	50
	2.7.3. Alisamento Exponencial Linear	51
	2.7.4. Alisamento Exponencial Sazonal e Linear de Winter	51
Capítulo 3.	Modelos ARIMA	55
	3.1. Métodos Avançados de Previsão de Séries Temporais	55
	3.2. Modelos Autoregressivo e de Médias Móveis	56
	3.2.1 Modelo Autoregressivo	56
	3.2.2 Modelo de Médias Móveis	56
	3.2.3 Modelo Autoregressivo e de Médias Móveis	57
	3.2.4 Modelo de Box e Jenkins	57
Capítulo 4.	Sobre as previsões	69
	4.1 Considerações sobre os Métodos de Previsão de Séries Temporais	69
	4.2. Medidas de Acuidade dos Métodos de Previsão de Séries Temporais	98
Bibliografia.		99

### Observação

\* O conteúdo do capítulo 1 (Regressão) foi gentilmente doado pelo professor Dr Manuel Luiz Figueirôa

**Capítulo 1****1.1****Regressão Linear:  $Y = A + BX + e$** 

- Introdução**

Ao se adotar o modelo estatístico acima é necessário que se esclareça as pressuposições que são admitidas quando a questão é a estimativa dos parâmetros A e B.

As pressuposições são:

- a) Normalidade, i. é,  $e$  tem distribuição normal;
- b) Média zero, i. é,  $E(e) = 0$ ;
- c) Homocedasticidade, i. é,  $E(e^2) = \sigma^2$ ;
- d) Independência, i. é,  $E(e_i, e_j) = 0$ ;  $i \neq j$ ;
- e) A variável independente X é não-estocástica, i. é, os valores de X são determinados de tal forma que  $\frac{1}{n} \cdot \sum (X_i - \bar{X})^2$  é um número finito diferente de zero.

Se o modelo é  $Y_i = A + BX_i + e_i$ , qual a distribuição de probabilidade da variável dependente  $Y_i$ ?

Esperança matemática de  $Y_i$

$$E(Y_i) = E(A + BX_i + e_i)$$

$$E(Y_i) = A + BX_i + E(e_i)$$

$$E(Y_i) = A + BX_i$$

Variância de  $Y_i$

$$V(Y_i) = E[(A + BX_i + e_i) - E(A + BX_i + e_i)]^2$$

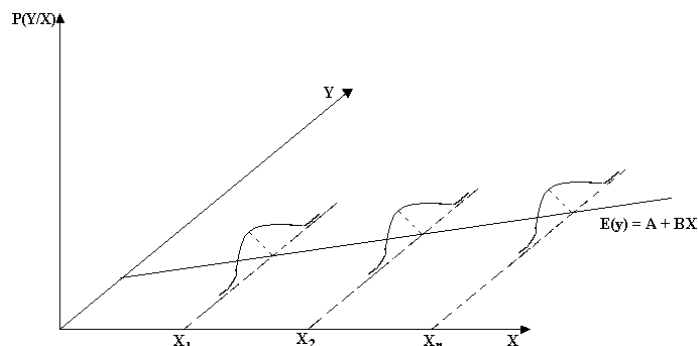
$$V(Y_i) = E[A + BX_i + e_i - A - BX_i]^2$$

$$V(Y_i) = E[e_i^2]$$

$$V(Y_i) = \sigma^2$$

Como  $Y_i$  é uma função linear de  $e_i$ , então a distribuição de  $Y_i$  é normal.

Graficamente pode-se observar que a reta de regressão  $E(Y) = A + BX$  tem os seus pontos passando pela média da distribuição.

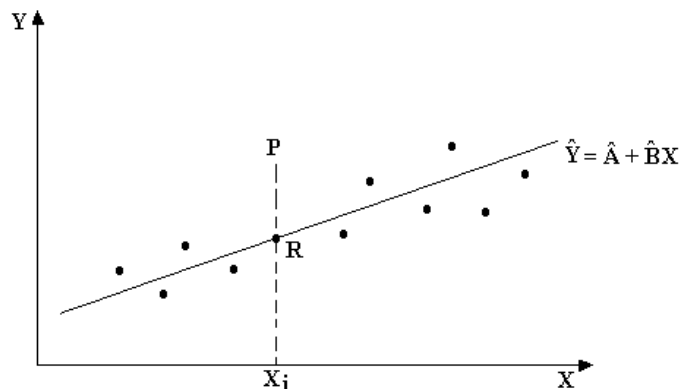


- Estimativa dos parâmetros**

Considere-se a tabela abaixo que registra os valores de  $Y_i$  para os respectivos valores de  $X_i$ .

$Y_i$	$Y_1$	$Y_2$	$Y_3$	...	$Y_n$
$X_i$	$X_1$	$X_2$	$X_3$	...	$X_n$

Aloque-se, esses pontos em um gráfico de dispersão e considere-se a reta  $Y = A + BX$  que se pretende estimar.



Seja um ponto  $P(X_i, Y_i)$  onde  $Y_i$  é o valor observado experimentalmente. Seja  $R(X_i, \hat{Y}_i)$  onde  $\hat{Y}_i$  é o valor obtido através da reta  $\hat{Y} = \hat{A} + \hat{B}X$ . Desta forma  $\overline{PR} = Y_i - \hat{Y}_i$  é o erro ou desvio que denota-se por  $e_i$ . Portanto,

$$e_i = Y_i - \hat{Y}_i$$

$$e_i = Y_i - (\hat{A} + \hat{B}X_i)$$

O que se pretende é encontrar estimadores dos parâmetros sob a condição da soma de quadrados do erro ser um mínimo.

$$\begin{aligned}
 \text{Seja } Z &= \sum e_i^2 = \sum [Y_i - (\hat{A} + \hat{B} X_i)]^2 \\
 \left\{ \begin{aligned} \partial Z / \partial \hat{A} &= -2 \sum (Y_i - \hat{A} - \hat{B} X_i) = 0 \\ \partial Z / \partial \hat{B} &= -2 \sum (Y_i - \hat{A} - \hat{B} X_i) \cdot X_i = 0 \end{aligned} \right. \\
 \left\{ \begin{aligned} \sum Y_i - N\hat{A} - \hat{B} \sum X_i &= 0 \\ \sum X_i Y_i - \hat{A} \sum X_i - \hat{B} \sum X_i^2 &= 0 \end{aligned} \right. \\
 \left\{ \begin{aligned} N\hat{A} + \hat{B} \sum X_i &= \sum Y_i \\ \hat{A} \sum X_i + \hat{B} \sum X_i^2 &= \sum X_i Y_i \end{aligned} \right. \quad (I)
 \end{aligned}$$

O sistema acima é conhecido como de equações normais e a sua solução fornece os valores de  $\hat{A}$  e  $\hat{B}$ , respectivamente.

Exemplo: A tabela abaixo indica o número de candidatos inscritos no vestibular da UFS de 1989 a 1998. Admita que a demanda por vagas na Universidade Federal é linear e estime a respectiva equação de mínimos quadrados.

ANO	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Demanda por vagas	7486	7306	11747	9127	8786	12063	11825	11868	11085	13343

Fonte: CCV/UFS

Solução:

É evidente que a variável independente  $Y$  será a demanda por vagas na UFS e a variável dependente  $X$  será o respectivo ano.

Quando a variável independente obedece uma lei aritmética (P.A.) esta pode ser substituída por outra nas mesmas condições e nesta situação indica-se uma nova série que apresenta a propriedade da sua soma ser nula.

No caso presente tem-se:

ANO	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
ANO	- 9	- 7	- 5	- 3	- 1	1	3	5	7	9

Calcula-se:

$$N = 10 ; \quad \sum X = 0 ; \quad \sum Y = 104636 ; \quad \sum X^2 = 330 ; \quad \sum X \cdot Y = 91142$$

e o sistema ( I ) assume a forma:

$$10 \hat{A} + \hat{B} \cdot 0 = 104636 \quad \Rightarrow \quad \hat{A} = 10463,6$$

$$\hat{A} \cdot 0 + \hat{B} \cdot 330 = 91142 \quad \Rightarrow \quad \hat{B} = 276,19$$

e a equação procurada é

$$\hat{Y} = 10463,6 + 276,19 X$$

### - Análise da Variância

Causas da Variância	G. L.	S. Q.	Q. M.	F	F(5%)
Regressão Linear	1 (K-1)	25172508,58	25172508,58	13,44	5,32
Desvio de regressão	8 (N-K)	14972899,42	1871612,43		
Total	9 (N-1)	40145408	-		

Onde K são o total de parâmetros a serem estimados

$$C = (\sum y)^2 / N = 1.094.869.250 \quad SQTot = \sum y^2 - C = 40.145.408$$

$$SQRgLinear = \beta' X' Y - C = [ 10463,6 \quad 276,19 ] \begin{bmatrix} 104636 \\ 91142 \end{bmatrix} - C = 25172508,58$$

Conclusão: A regressão é significativa para  $\alpha = 5\%$

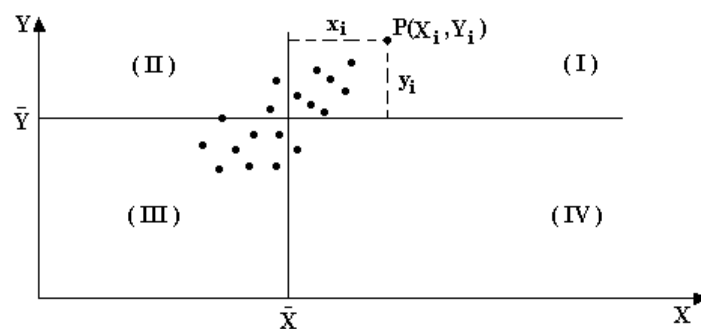
$$\text{Coef. de determinação} = \frac{SQRL}{SQTot} = 0,6270 = r^2$$

Cef. de correlação:  $r = 0,7918$

### • Coeficiente de Correlação Linear

Considere-se os pares  $(X_i, Y_i)$  em um gráfico de dispersão e construa as retas:  $X = \bar{X}$  e  $Y = \bar{Y}$  conforme figura abaixo.

Divida a dispersão em quatro partes I, II, III e IV conforme indicado acima.



Considere um ponto genérico  $P(X_i, Y_i)$  e defina:

$$x_i = X_i - \bar{X} \quad ; \quad y_i = Y_i - \bar{Y}$$

É fácil de observar que para pontos de ( I ) e ( III ) tem-se  $\sum x_i \cdot y_i > 0$  enquanto que para de ( II ) e ( IV ) tem-se  $\sum x_i \cdot y_i < 0$ .

Quando ocorrer  $\sum_{i=1}^n x_i \cdot y_i > 0$  é porque a maioria dos pontos situam-se em ( I ) e ( III ) e quando

$\sum_{i=1}^n x_i \cdot y_i < 0$  a maioria dos pontos encontram-se ( II ) e ( IV ). Entretanto se  $\sum_{i=1}^n x_i \cdot y_i = 0$  ou

próximo de zero os pontos estão dispersos nos quatro quadrantes e não existe relação de dependência entre X e Y.

Observe que  $\sum_{i=1}^n x_i \cdot y_i$  é uma medida de correlação absoluta sofrendo, portanto, influência das unidades.

Pearson eliminando a influencia das unidades definiu uma medida relativa que é conhecida por:

$$\text{Coeficiente de correlação de Pearson} = r = \frac{\sum x_i y_i}{n S_x S_y}$$

$$\text{onde } S_x = \sqrt{\frac{\sum x_i^2}{n}} \quad \text{e} \quad S_y = \sqrt{\frac{\sum y_i^2}{n}}$$

Realizando as substituições correspondentes resulta

$$r = r \cdot \frac{S_y}{S_x} \frac{\sum x_i y_i}{\sqrt{\sum x^2 \sum y^2}}$$

Ao considerarmos a reta de mínimos quadrados através do modelo  $y = a + bx$ , onde  $y = Y - \bar{Y}$  e  $x = X - \bar{X}$  obtem-se:  $a = 0$  e  $b = \frac{\sum x \cdot y}{\sum x^2}$ .

Substituindo  $\sum x_i y_i = r \sqrt{(\sum x^2) (\sum y^2)}$  em  $b = \frac{\sum x \cdot y}{\sum x^2}$  tem-se:

$$b = \frac{r \cdot \sqrt{\sum x^2 \cdot \sum y^2}}{\sum x^2} = \frac{r \cdot \sqrt{\sum x^2 \cdot \sum y^2}}{(\sum x^2)^2} = r \cdot \sqrt{\frac{\sum y^2}{\sum x^2}}$$

$$b = r \cdot \sqrt{\frac{\frac{\sum y^2}{n}}{\frac{\sum x^2}{n}}} = r \cdot \frac{S_y}{S_x}$$

A expressão  $b = r \cdot \frac{S_y}{S_x}$  permite concluir que o estimador  $b$  tem mesmo sinal do coeficiente de correlação  $r$ .

Por outro lado ao se definir a reta de mínimos quadrados adotou-se  $y_i = \hat{y}_i + e_i$ .

$$\text{Logo } \sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 + 2 \sum \hat{y}_i e_i$$

Ocorre que  $y = bx$  pois a reta de mínimos quadrados passa por  $(\bar{X}, \bar{Y})$  isto é,  $(Y - \bar{Y}) / (X - \bar{X}) = b$ .

$$\text{Daí, } \sum \hat{y}_i e_i = b \sum e_i x_i = b \sum x_i (y_i - \hat{y}_i) = b \sum x_i (y_i - \hat{b} x_i) = b [ \sum x_i y_i - \hat{b} \sum x_i^2 ]$$

Ocorre que  $\hat{b} \sum x_i^2 = \sum x_i y_i$  logo:

$$\sum y_i e_i = b [ \sum x_i y_i - \sum x_i y_i ] \quad ; \quad \sum y_i e_i = 0 \quad \text{e resulta}$$

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2$$

A expressão acima afirma que: A variação total dos valores de  $Y$  em torno de sua média amostral  $\bar{Y}$  pode ser decomposta em duas partes onde a primeira  $\sum y_i^2$  que é a variação dos valores de  $Y$  em torno da sua média  $\bar{Y}$ , que é a variação explicada pela regressão linear, e, a segunda  $\sum e_i^2$  é a variação não explicada devido as perturbações aleatórias.

A expressão acima também indica que a soma dos quadrados das variações totais é igual a soma de quadrados devido a regressão linear mais a soma de quadrados dos fatores aleatórios.

O quociente entre a soma de quadrados das variações explicadas pela regressão linear sobre a soma de quadrados das variações totais fornece:

$$(\sum \hat{y}_i^2) / (\sum y_i^2) = [\sum (bx_i)^2] / \sum y_i^2 = b^2 [(\sum x_i^2) / (\sum y_i^2)] = r^2$$

que é o quadrado do coeficiente de correlação de agora por diante denominado de coeficiente de determinação.

Dividindo-se,  $\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2$ , ambos os membros por  $\sum y_i^2$  tem-se:

$$1 = r^2 + \frac{\sum e_i^2}{\sum y_i^2}$$

$$r^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2}$$

Conclusões:

- O valor máximo de  $r^2$  é **um** e isto só ocorre quando  $\sum e_i^2 = 0$ , isto é, todos os pontos do diagrama de dispersão situam-se em uma reta;
- O valor mínimo de  $r^2$  é zero isto ocorrendo quando  $\sum e_i^2 = \sum y_i^2$ , isto é, não existe correlação entre as variáveis;
- Como  $0 \leq r^2 \leq 1$  então  $-1 < r < 1$ .

### Análise de Variância

$$\text{Ocorre que } \frac{B - \hat{B}}{\frac{\sigma}{\sqrt{\sum x_i^2}}} \cap N(0;1), \quad \text{logo } \frac{(B - \hat{B})^2}{\frac{\sigma^2}{\sum x_i^2}} \cap \chi_1^2.$$

Como  $(\sum e^2 / \sigma^2) \cap \chi_{n-2}^2$ , conclui-se que:

$$F = \frac{\frac{(B - \hat{B})^2}{\frac{\sigma^2}{\sum x_i^2}}}{\frac{\frac{\sum e^2}{\sigma^2}}{n-2}} = \frac{(B - \hat{B})^2 \cdot \sum x_i^2}{\frac{\sum e^2}{\sigma^2 \cdot (n-2)}} = \frac{(B - \hat{B})^2 \cdot \sum x_i^2}{\sigma^2} \cdot \frac{\sigma^2 \cdot (n-2)}{\sum e^2} = \frac{(B - \hat{B})^2 \cdot \sum x_i^2}{\sum e_i^2}$$

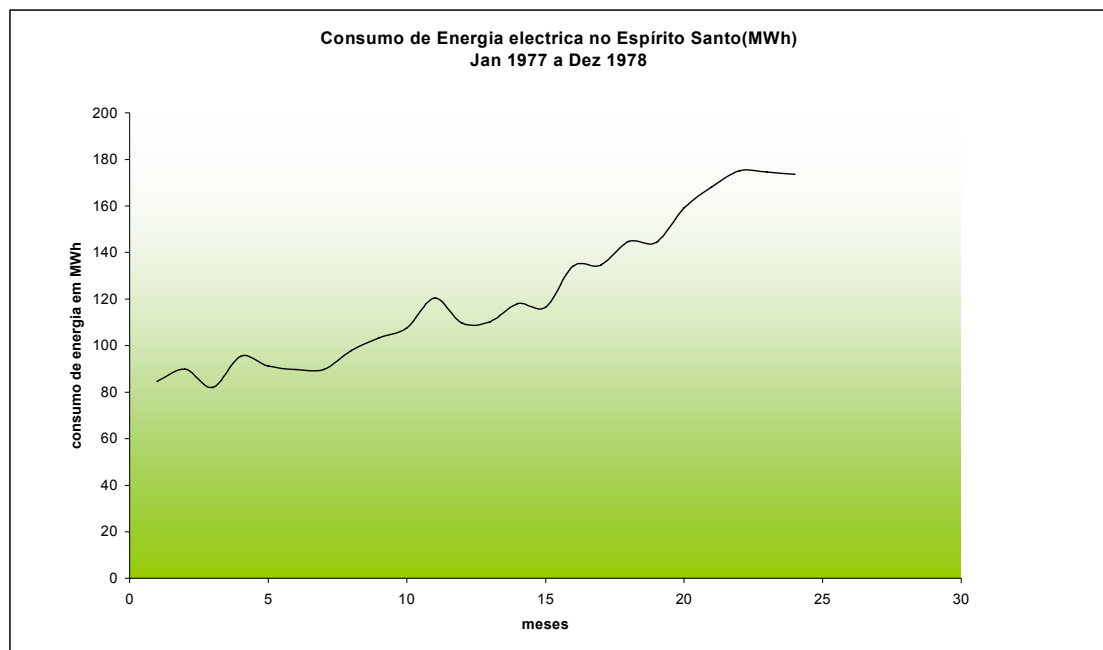
Se  $B = 0$  então  $F = (\hat{B}^2 \cdot \sum x_i^2) / [(\sum e_i^2) / (n-2)] = \sum y_i^2 / [\sum e_i^2 / (n-2)]$ , o que permite rejeitar a hipótese se  $F_c > F_{\text{tab}}$ .



**Consumo de Energia electrica no Espírito Santo(MWh)****Jan 1977 a Dez 1978**

Ajuste de mínimos quadrados

t	z(t)	t	z(t)
1	84,6	13	110,3
2	89,9	14	118,1
3	81,9	15	116,5
4	95,4	16	134,2
5	91,2	17	134,7
6	89,8	18	144,8
7	89,7	19	144,4
8	97,9	20	159,2
9	103,4	21	168,2
10	107,6	22	175,2
11	120,4	23	174,5
12	109,6	24	173,7



Saída de regressão em Excel

## RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0,962318
R-Quadrado	0,926055
R-quadrado ajustado	0,922694
Erro padrão	8,667021
Observações	24

## ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	20696,29	20696,29	275,5199	6,3E-14
Resíduo	22	1652,579	75,11725		
Total	23	22348,87			

	<i>Coefficients</i>	<i>Error standard</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>
Interseção	68,43841	3,651853	18,74073	5,16E-15	60,86492
Variável X 1	4,242261	0,255576	16,59879	6,3E-14	3,712227

## RESULTADOS DE RESÍDUOS

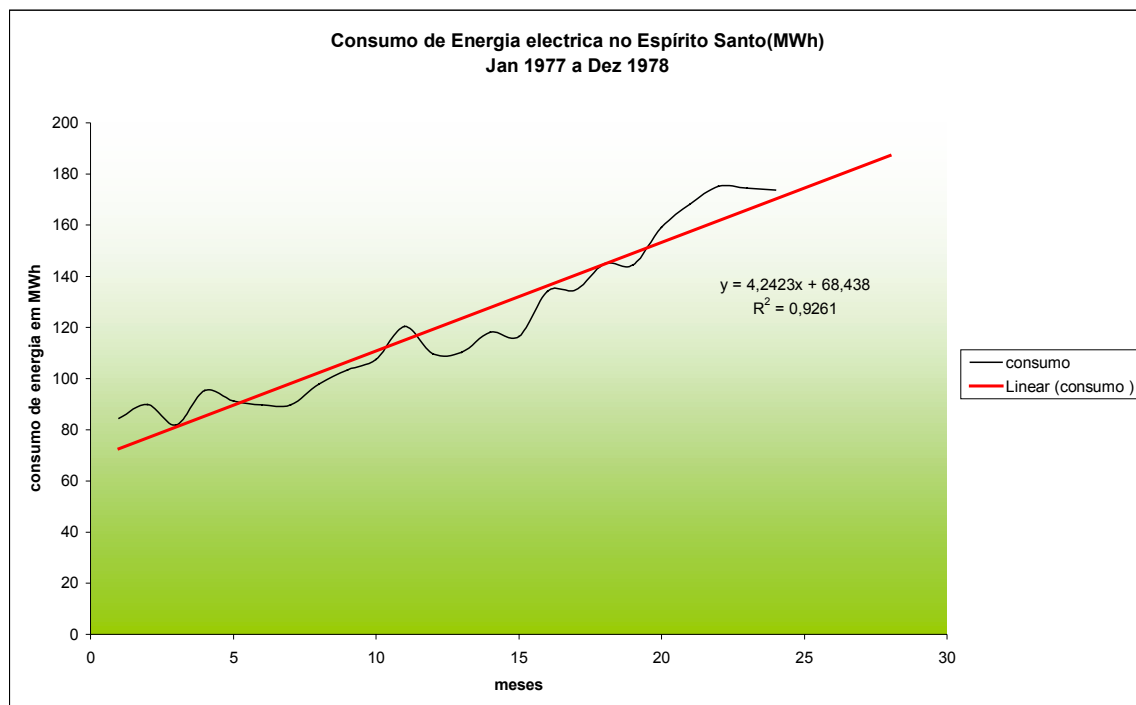
RESULTADOS DE  
PROBABILIDADE

<i>Observação</i>	<i>Y previsto</i>	<i>Resíduos</i>	<i>Percentil</i>	<i>Y</i>
1	72,68067	11,91933	2,083333	81,9
2	76,92293	12,97707	6,25	84,6
3	81,16519	0,734812	10,41667	89,7
4	85,40745	9,992551	14,58333	89,8
5	89,64971	1,55029	18,75	89,9
6	93,89197	-4,09197	22,91667	91,2
7	98,13423	-8,43423	27,08333	95,4
8	102,3765	-4,47649	31,25	97,9
9	106,6188	-3,21875	35,41667	103,4
10	110,861	-3,26101	39,58333	107,6
11	115,1033	5,296725	43,75	109,6
12	119,3455	-9,74554	47,91667	110,3
13	123,5878	-13,2878	52,08333	116,5
14	127,8301	-9,73006	56,25	118,1
15	132,0723	-15,5723	60,41667	120,4
16	136,3146	-2,11458	64,58333	134,2
17	140,5568	-5,85684	68,75	134,7
18	144,7991	0,000899	72,91667	144,4

19	149,0414	-4,64136	77,08333	144,8
20	153,2836	5,916377	81,25	159,2
21	157,5259	10,67412	85,41667	168,2
22	161,7681	13,43186	89,58333	173,7
23	166,0104	8,489594	93,75	174,5
24	170,2527	3,447333	97,91667	175,2

Valores observados após do ajuste

observação	consumo	Previsto	Erro
25	<b>179,8</b>	174,4955	-5,3045
26	<b>185,8</b>	178,7378	-7,0622
27	<b>270,3</b>	182,9801	-87,3199
28	<b>196,9</b>	187,2224	-9,6776



**1.2 Regressão Quadrática:  $Y = A + BX$** 

$$+ CX^2 + e$$

- Introdução**

As mesmas pressuposições da regressão linear são aqui admitidas.

A esperança matemática de  $Y_i$  é:

$$E(Y_i) = E(A + BX_i + CX_i^2 + e_i)$$

$$E(Y_i) = A + BX_i + CX_i^2$$

$$E(Y_i) = A + BX_i + CX_i^2 + E(e_i)$$

Variância de  $Y_i$

$$V(Y_i) = E[(A + BX_i + CX_i^2 + e_i) - E(A + BX_i + CX_i^2 + e_i)]^2$$

$$V(Y_i) = E[A + BX_i + CX_i^2 + e_i - A - BX_i - CX_i^2]^2$$

$$V(Y_i) = E(e_i^2)$$

$$V(Y_i) = \sigma^2$$

Como  $Y_i$  é uma função linear de  $e_i$ , então a distribuição de  $Y_i$  é normal.

- Estimativa dos parâmetros**

Considere-se  $e_i = Y_i - \hat{Y}_i$

$$e_i = Y_i - (\hat{A} + \hat{B}X_i + \hat{C}X_i^2).$$

O que se pretende é encontrar estimadores dos parâmetros sob a condição da soma de quadrados do erro ser um mínimo.

$$\text{Seja } Z = \sum e_i^2 = \sum [Y_i - (\hat{A} + \hat{B}X_i + \hat{C}X_i^2)]^2$$

$$\partial Z / \partial \hat{A} = -2 \sum [Y_i - \hat{A} - \hat{B}X_i - \hat{C}X_i^2] = 0$$

$$\partial Z / \partial \hat{B} = -2 \sum [(Y_i - \hat{A} - \hat{B}X_i - \hat{C}X_i^2) X_i] = 0$$

$$\partial Z / \partial \hat{C} = -2 \sum [(Y_i - \hat{A} - \hat{B}X_i - \hat{C}X_i^2) X_i^2] = 0$$

Resulta:

$$\begin{aligned} N\hat{A} + \hat{B} \sum X_i + \hat{C} \sum X_i^2 &= \sum Y_i \\ \hat{A} \sum X_i^2 + \hat{B} \sum X_i^3 + \hat{C} \sum X_i^4 &= \sum X_i Y_i \quad (\text{II}) \\ \hat{A} \sum X_i^2 + \hat{B} \sum X_i^3 + \hat{C} \sum X_i^4 &= \sum X_i^2 Y \end{aligned}$$

Que é o sistema de equações normais cuja solução fornece  $\hat{A}$ ,  $\hat{B}$  e  $\hat{C}$ .

Exemplo: Considere o número de candidatos inscritos no vestibular da UFS de 1989 a 1998 do exercício anterior. Admita que a demanda por vagas na UFS é quadrática e estime a respectiva equação de mínimos quadrados.

Solução:

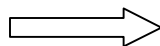
A série dos anos continua sendo substituída pela série ( -9, -7, -5, -3, -1, 1, 3, 5, 7, 9 ). Desta forma tem-se:

$$\begin{aligned} N &= 10; & \sum X &= 0; & \sum X^2 &= 330; & \sum X^3 &= 0; & \sum X^4 &= 19338 \\ \sum Y &= 104636; & \sum XY &= 91142; & \sum X^2 Y &= 3388100 \end{aligned}$$

e o sistema ( II ) assume a forma:

$$10 \hat{A} + 0 \hat{B} + 330 \hat{C} = 104636$$

$$0 + 330 \hat{B} + 0 \hat{C} = 91142$$



$$\hat{B} = 276,1879$$

$$330 \hat{A} + 0 \hat{B} + 19338 \hat{C} = 3388100$$

$$10 \hat{A} + 330 \hat{C} = 104636$$

$$330 \hat{A} + 19338 \hat{C} = 3388100$$

$$\begin{bmatrix} \hat{A} \\ \hat{B} \end{bmatrix} \begin{bmatrix} \hat{A} \\ \hat{C} \end{bmatrix} = (1 / 84480) \cdot \begin{bmatrix} 19338 & -330 \\ -330 & 10 \end{bmatrix} \begin{bmatrix} 104636 \\ 3388100 \end{bmatrix}$$

$$\begin{bmatrix} \hat{A} \\ \hat{C} \end{bmatrix} = \begin{bmatrix} 10717,0687 \\ -7,6809 \end{bmatrix}$$

e a equação procurada é

$$\hat{Y} = 10717,0687 + 276,1879 X - 7,6809 X^2$$

### - Análise da Variância

Causas da Variância	G. L.	S. Q.	Q. M.	F	F(5%)
Regressão Linear	2	25673987,5	12836993,75	6,21	4,74
Desvio de regressão	7	14471420,5	2067345,79		
Total	9	40145408			

$$C = (\sum y)^2 / N = 1.094.869.250$$

$$SQ_{Tot} = \sum y^2 - C = 40.145.408$$

$$SQR_{gLinear} = \beta' X' Y - C = \begin{bmatrix} 10717,07 & 276,19 & -7,68 \end{bmatrix} \begin{bmatrix} 104636 \\ 91142 \\ 3388100 \end{bmatrix} - C = 25673987,5$$

$$\text{Coef. de determinação} = r^2 = SQRQ / SQT = 0,6395 \text{ ou } 63,95\%$$

$$\text{Cef. de correlação: } r = 0,7997$$

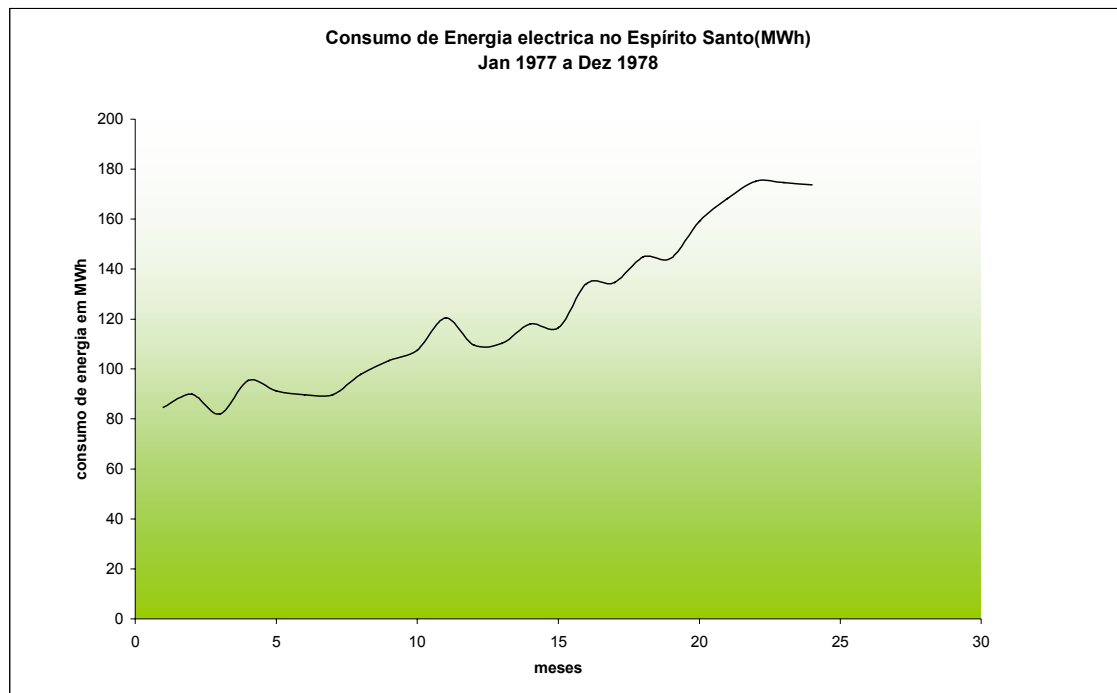
Conclusão: A regressão é significativa

**Consumo de Energia electrica no Espírito Santo(MWh)  
Jan 1977 a Dez 1978**

Ajuste de mínimos quadrados

t	z(t)	t <sup>2</sup>
1	84,6	1
2	89,9	4
3	81,9	9
4	95,4	16
5	91,2	25
6	89,8	36
7	89,7	49
8	97,9	64
9	103,4	81
10	107,6	100
11	120,4	121
12	109,6	144
13	110,3	169
14	118,1	196
15	116,5	225
16	134,2	256
17	134,7	289
18	144,8	324
19	144,4	361
20	159,2	400
21	168,2	441
22	175,2	484
23	174,5	529
24	173,7	576





Saída de regressão quadrática em Excel

### RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0,984858
R-Quadrado	0,969944
R-quadrado ajustado	0,967082
Erro padrão	5,655637
Observações	24

### ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	2	21677,16	10838,58	338,8515	1,04E-16
Resíduo	21	671,7108	31,98623		
Total	23	22348,87			

	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	84,64032	3,773451	22,43048	3,74E-16	76,79299	92,48764	76,79299	92,48764
Variável X 1	0,503358	0,695474	0,723764	0,477197	-0,94296	1,949675	-0,94296	1,949675
Variável X 2	0,149556	0,027007	5,53763	1,7E-05	0,093391	0,205721	0,093391	0,205721

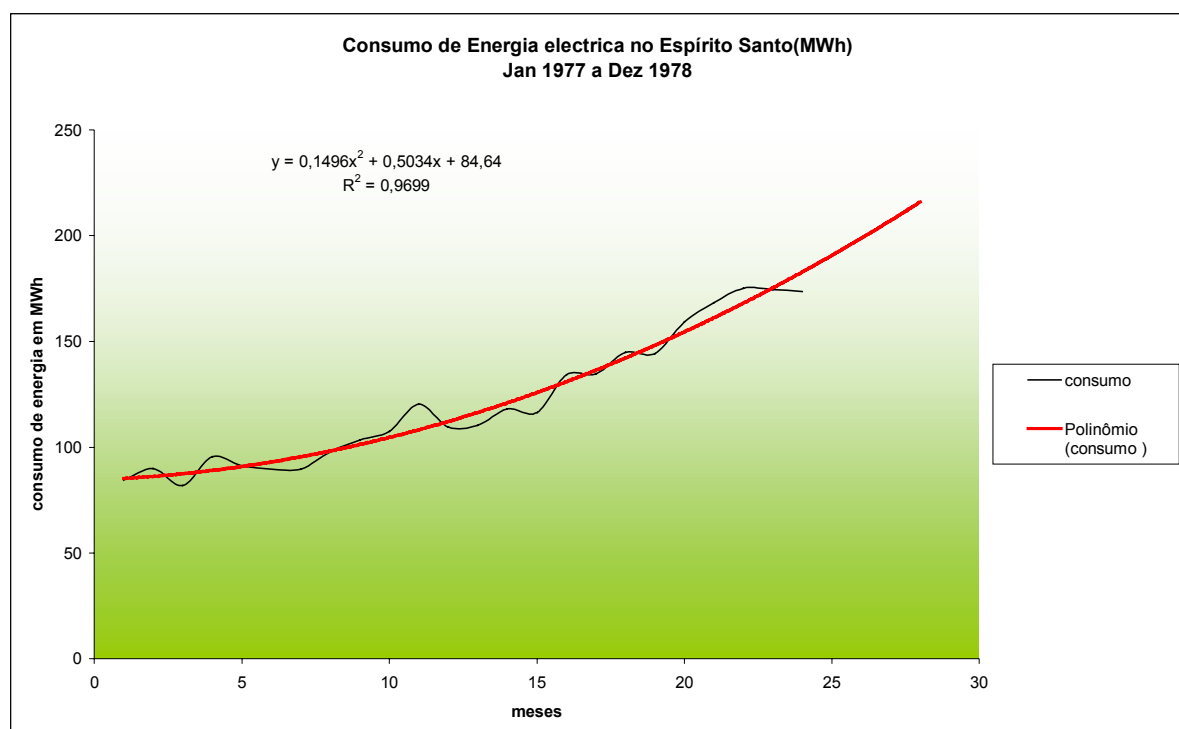
## RESULTADOS DE RESÍDUOS

<i>Observação</i>	<i>Y previsto</i>	<i>Resíduos</i>
1	85,29323	-0,69323
2	86,24526	3,654742
3	87,4964	-5,5964
4	89,04665	6,353352
5	90,89601	0,303989
6	93,04449	-3,24449
7	95,49207	-5,79207
8	98,23877	-0,33877
9	101,2846	2,115414
10	104,6295	2,97049
11	108,2735	12,12645
12	112,2167	-2,6167
13	116,459	-6,15896
14	121,0003	-2,90033
15	125,8408	-9,34081
16	130,9804	3,219588

17	136,4191	-1,71912
18	142,1569	2,643056
19	148,1939	-3,79388
20	154,5299	4,670076
21	161,1651	7,034918
22	168,0994	7,100647
23	175,3327	-0,83274
24	182,8652	-9,16523

Valores observados após do ajuste

observação	consumo	Previsto	Erro
25	<b>179,8</b>	190,725	10,925
26	<b>185,8</b>	198,858	13,058
27	<b>270,3</b>	207,2902	-63,0098
28	<b>196,9</b>	216,0216	19,1216



## 1.3

**Regressão Múltipla**• **Introdução**

Tem-se uma regressão linear múltipla quando uma variável aleatória  $Y$  depende de duas ou mais variáveis  $X_1, X_2, \dots, X_n$ .

Nestas condições o modelo estatístico assume a forma:

$$Y_i = A + B_1 X_{1i} + B_2 X_{2i} + \dots + B_p X_{pi} + e_i$$

onde:

$Y_i$  são os valores observados

$A, B_1, B_2, \dots, B_p$  são os parâmetros

$X_{ij}$  são os valores fixos, com  $i = 1, 2, \dots, n$  e  $j = 1, 2, \dots, p$ .

Com procedimento já conhecido conclui-se:

$$E(Y_i) = A + B_1 x_1 + B_2 x_2 + \dots + B_p x_p \quad \text{e} \quad V(Y_i) = \sigma^2$$

**Estimativa de parâmetros**

$$e_i = Y_i - (A + B_1 X_1 + B_2 X_2 + \dots + B_p X_p)$$

$$Z = \sum e_i^2 = \sum [Y_i - (A + B_1 X_1 + B_2 X_2 + \dots + B_p X_p)]^2$$

$$\partial Z / \partial A = -2 \sum [Y_i - A - B_1 X_1 - B_2 X_2 - \dots - B_p X_p] = 0$$

$$\partial Z / \partial B_1 = -2 \sum [(Y_i - A - B_1 X_1 - B_2 X_2 - \dots - B_p X_p) X_1] = 0$$

$$\partial Z / \partial B_2 = -2 \sum (Y_i - A - B_1 X_1 - B_2 X_2 - \dots - B_p X_p) X_2 = 0$$

.....

$$\partial Z / \partial B_p = -2 \sum (Y_i - A - B_1 X_1 - B_2 X_2 - \dots - B_p X_p) X_p = 0$$

$$\left\{ \begin{array}{l} N\hat{A} + \hat{B}_1 \sum X_{1i} + \hat{B}_2 \sum X_{2i} + \dots + \hat{B}_p \sum X_{pi} = \sum Y_i \\ \hat{A} \sum X_{1i} + \hat{B}_1 \sum X_{1i}^2 + \hat{B}_2 \sum X_{1i} X_{2i} + \dots + \hat{B}_p \sum X_{1i} X_{pi} = \sum X_{1i} Y_i \\ \hat{A} \sum X_{2i} + \hat{B}_1 \sum X_{1i} X_{2i} + \hat{B}_2 \sum X_{2i}^2 + \dots + \hat{B}_p \sum X_{2i} X_{pi} = \sum X_{2i} Y_i \end{array} \right.$$

O sistema acima conhecido como de equações normais fornecem os valores de  $\hat{A}$ ,  $\hat{B}_1$ ,  $\hat{B}_2$ , ...,  $\hat{B}_p$  que são os parâmetros estimados.

Exemplo: A tabela abaixo indica o peso ( kg ), a altura ( m ) e a idade ( anos ) de crianças em idade escolar.

Peso ( Y )	32	35,5	26,5	33,5	27,5	29	38,5	30,0	31,0	28,0	35,0	30,0
Altura ( X <sub>1</sub> )	1,14	1,18	0,98	1,24	1,02	1,00	1,10	0,96	1,04	0,84	1,22	1,14
Idade ( X <sub>2</sub> )	8	10	6	11	8	7	10	9	10	6	12	9

Estabeleça a equação de regressão  $Y = A + B_1 X_1 + B_2 X_2$ .

Solução:

Calcula-se:

$$N = 12; \quad \sum X_{1i} = 12,86; \quad \sum X_{2i} = 106; \quad \sum X_{1i}^2 = 13,94; \quad \sum X_{2i}^2 = 976$$

$$\sum X_{1i} X_{2i} = 115,58; \quad \sum Y_i = 376,5; \quad \sum X_{1i} Y_i = 406,74; \quad \sum X_{2i} Y_i = 3384,5$$

$$\left\{ \begin{array}{l} N\hat{A} + \hat{B}_1 \sum X_{1i} + \hat{B}_2 \sum X_{2i} = \sum Y_i \\ \hat{A} \sum X_{1i} + \hat{B}_1 \sum X_{1i}^2 + \hat{B}_2 \sum X_{1i} X_{2i} = \sum X_{1i} Y_i \\ \hat{A} \sum X_{2i} + \hat{B}_1 \sum X_{1i} X_{2i} + \hat{B}_2 \sum X_{2i}^2 = \sum X_{2i} Y_i \end{array} \right.$$

$$\begin{cases} 12 \hat{A} + 12,86 \hat{B}_1 + 106 \hat{B}_2 = 376,5 \\ 12,86 \hat{A} + 13,94 \hat{B}_1 + 115,58 \hat{B}_2 = 406,74 \\ 106 \hat{A} + 115,58 \hat{B}_1 + 976 \hat{B}_2 = 3384,5 \end{cases}$$

$$\begin{bmatrix} 12 & 12,86 & 106 \\ 12,86 & 13,94 & 115,58 \\ 106 & 115,58 & 976 \end{bmatrix} \begin{bmatrix} \hat{A} \\ \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} = \begin{bmatrix} 376,5 \\ 406,74 \\ 3384,5 \end{bmatrix}$$

$$\begin{bmatrix} \hat{A} \\ \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} = (1 / 28,1792) \begin{bmatrix} 246,7036 & -299,88 & 8,7188 \\ -299,88 & 476 & -23,8 \\ 8,7188 & -23,8 & 1,9004 \end{bmatrix} \begin{bmatrix} 376,5 \\ 406,74 \\ 3384,5 \end{bmatrix}$$

$$\begin{bmatrix} \hat{A} \\ \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} = \begin{bmatrix} 14,8866 \\ 5,4054 \\ 1,2108 \end{bmatrix}$$

e a equação procurada é

$$\hat{Y} = 14,8866 + 5,4054 X_{1i} + 1,2108 X_{2i}$$

#### - Análise de variância

Causas de Variação	G. L.	S. Q.	Q. M.	F	F(5%)
Regressão múltipla	2	89,10	44,55	7,10	4,26
Desvio da regressão	9	56,46	6,27		
Total	11	145,56			

$$C = \frac{(\sum y)^2}{N} = 11812,69$$

$$(\sum Y)^2 / N = 11812,69$$

$$SQ_{Tot} = \sum Y^2 - C = 145,56$$

$$SQ_{Reg} \text{ múltipla} = \beta' X' Y - C = [14,89 \quad 5,41 \quad 1,21] \begin{bmatrix} 376,5 \\ 406,74 \\ 3384,5 \end{bmatrix} - C = 89,10$$

Conclusão: A regressão múltipla é significativa.

$$\text{Coef. det} = r^2 = SQR.\text{múltiplo} / SQ_{tot} = 0,6121 \text{ ou } 61,21\%$$

$$r = 0,7823$$

### Exemplo

Preço(R\$)	Demanda(Kg)	Oferta(Kg)
8	0	10,1
7	3,1	9,6
6	4,8	8,8
5	5,8	8
4	6,6	6,8
3	7,5	5
2	8,3	2,3
1	8,9	0
0	9,5	0

Saída de regressão cubica em Excel

### RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0,995124
R-Quadrado	0,990272
R-quadrado ajustado	0,987029
Erro padrão	0,311905
Observações	9

## ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	2	59,41629	29,70815	305,3735	9,21E-07
Resíduo	6	0,583708	0,097285		
Total	8	60			

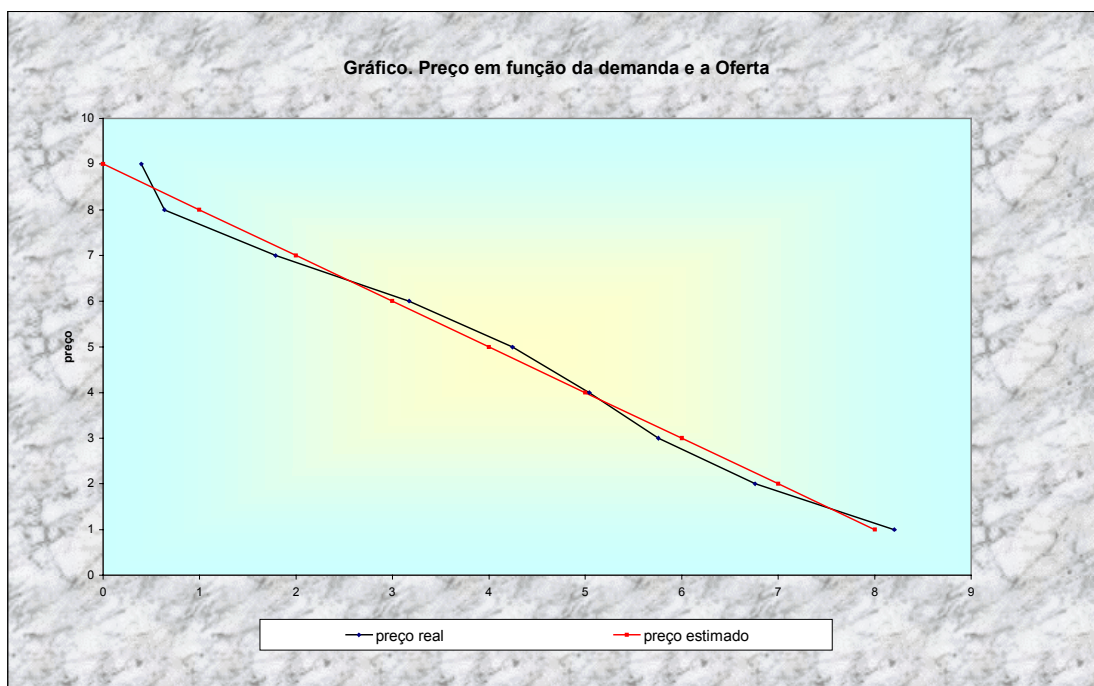
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	4,218947	0,79757	5,289751	0,001848	2,267362	6,170532	2,267362	6,170532
Variável X 1	-0,40249	0,078583	-5,12177	0,002174	-0,59477	-0,2102	-0,59477	-0,2102
Variável X 2	0,394565	0,060002	6,575912	0,000593	0,247746	0,541383	0,247746	0,541383



## RESULTADOS DE RESÍDUOS

<i>Observação</i>	<i>Y previsto</i>	<i>Resíduos</i>
1	8,204051	-0,20405
2	6,759062	0,240938
3	5,759183	0,240817
4	5,041045	-0,04105
5	4,245578	-0,24558
6	3,173124	-0,17312
7	1,78581	0,21419
8	0,636819	0,363181
9	0,395327	-0,39533





## Capítulo 2

## SÉRIES TEMPORAIS

Este curso é dedicado ao estudo de séries temporais e sua aplicabilidade na previsão de valores futuros dessas séries. São introduzidos os conceitos de série temporal e previsão, além da classificação e descrição de alguns métodos de previsão. Também são feitas determinadas considerações a respeito desses métodos e sobre as medidas de acuidade usualmente adotadas na avaliação de tais procedimentos.

Uma série temporal é qualquer conjunto de observações ordenadas no tempo. Alguns exemplos:

- i) Estimativas Trimestrais do PNB;
- ii) Valores diários de temperatura da cidade de São Paulo;
- iii) Índice diários da Bolsa de Valores do Rio de Janeiro;
- iv) Quantidade anual de chuva na cidade de Fortaleza;
- v) Valores mensais de vendas de automóveis no Brasil;
- vi) Um registro de marés no porto de Santos.

Nos exemplos i), v), temos séries temporais discretas, enquanto vi), é um exemplo de uma série contínua. Muitas vezes uma série temporal discreta é obtida através da amostragem de uma série temporal contínua em intervalos de tempos iguais,  $\Delta t$ . Assim para analisar a série vi) será necessário amostrá-la ( em intervalos de tempo de uma hora, por exemplo), convertendo a série contínua, observada no intervalo  $[0, T]$ , digamos em uma série discreta com  $N$  pontos, onde  $N = (T / \Delta t)$ . Em outros casos, como para as séries iv) ou v), temos que o valor da série num instante é obtido acumulando-se ( ou agregando-se ) valores em intervalos de tempos iguais.

## 2.1. Definição

A classe de fenômenos cujo processo observacional e conseqüente quantificação numérica gera uma seqüência de dados distribuídos no tempo é denominada *série temporal* [SOUZA1996].

A natureza de uma série temporal e a estrutura de seu mecanismo gerador estão relacionadas com o intervalo de ocorrência das observações no tempo. Caso o levantamento das observações da série possa ser feito a qualquer momento do tempo, a série temporal é dita *contínua*, sendo denotada por  $x(t)$  (1.1)

Entretanto, de acordo com e , na maioria das séries, as observações são tomadas em intervalos de tempo discretos e equidistantes.

Uma série temporal *discreta* pode ser representada por  $X^T = \{x_1, x_2, \dots, x_T\}$ , sendo que cada observação discreta  $x_t$  está associada a um instante de tempo distinto, existindo uma relação de dependência serial entre essas observações [SOUZA1989].

## 2.2. Objetivo

O objetivo inicial da análise de séries temporais a realização de inferências sobre as propriedades ou características básicas do mecanismo gerador do processo estocástico das observações da série. Assim, através da abstração de regularidades contidas nos fenômenos observáveis de uma série temporal existe a possibilidade de se construir um modelo matemático como uma representação simplificada da realidade.

Após a formulação do modelo matemático, obtido pela seleção entre as alternativas de classes de modelos identificadas como apropriadas para essa representação e subsequente estimação de seus parâmetros, é possível utilizá-lo para testar alguma hipótese ou teoria a respeito do mecanismo gerador do processo estocástico e realizar a previsão de valores futuros da série temporal.

## 2.3 Fundamento Probabilístico

**Definição.** Seja  $T$  um conjunto arbitrário. Um processo estocástico é uma família  $\{Y(t), t \in T\}$  tal que,  $\forall t \in T$ ,  $Y(t)$  é uma variável aleatória (v.a.)

Série temporal é um processo aleatório.

Uma Forma alternativa de definição de processo estocástico é uma família de variáveis aleatórias  $\{Y(t), t \in T\}$  é um processo estocástico se as v.a  $Y(t_1), Y(t_2), \dots, Y(t_n)$  tem função finito dimensionais,  $F(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n) = P(Y(t_1) \leq y_1, Y(t_2) \leq y_2, \dots, Y(t_n) \leq y_n)$ , conhecidas para todo  $n \geq 1$  satisfazendo as condições de:

- simetria: para qualquer permutação  $j_1, j_2, \dots, j_n$ , dos índices  $1, 2, \dots, n$  temos:

$$F(y_{j_1}, y_{j_2}, \dots, y_{j_n}; t_{j_1}, t_{j_2}, \dots, t_{j_n}) = F(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n)$$

exemplo ( $n=3$ )  $F(y_2, y_1, y_3; t_2, t_1, t_3) = F(y_1, y_2, y_3; t_1, t_2, t_3)$

- consistência:  $\lim_{y_n \rightarrow \infty} F(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n) = F(y_1, y_2, \dots, y_{n-1}; t_1, t_2, \dots, t_{n-1})$

Essa definição não é muito útil na prática pois é muito difícil a especificação de todas as distribuições finito-dimensionais.

Normalmente o que se faz é concentrar nos primeiro momentos. Estes são:

i) Função média:  $\mu(t) = E\{Y(t)\}$

ii) Função de auto- covariância (facv):

$$\gamma(t_1, t_2) = E\{[Y(t_1) - \mu(t_1)][Y(t_2) - \mu(t_2)]\} = E\{Y(t_1)Y(t_2)\} - \mu(t_1)\mu(t_2)$$

Em particular se  $t_1 = t_2 = t$ ,

$\gamma(t, t) = V\{Y(t)\}$  é a variância de  $Y(t)$  denotado por  $V(t)$  ou  $\sigma^2(t)$ .

A facv fornece a forma de dependência temporal do processo  $Y(t)$ . Ela não traduz a força dessa dependência pois depende da unidade de medição de  $Y$ .

Para contornar esse problema, a facv é comumente substituída pela:

$$\text{iii) Função de auto correlação: } \rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sigma(t_1)\sigma(t_2)}$$

A importância da função média e da facv deve-se ao fato de que se as distribuições finito dimensionais de  $Y(t)$  são normais então basta conhecer  $\mu$  e  $\gamma$  para conhecer todo o processo.

De agora em diante consideraremos apenas processos estacionários ou possíveis de estacionarização por transformações.

Após observar a série temporal discreta  $Y_1, \dots, Y_n$ ,  $\gamma_k$  será estimado por:

$$C_k = \frac{1}{n} \sum_{i=1}^{n-k} (Y_i - \bar{Y})(Y_{i+k} - \bar{Y}) \text{ onde, } \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, k=0,1,\dots,n-1$$

$C_k$  é a facv amostral

$\rho_k$  pode então ser estimado por  $r_k = \frac{C_k}{C_0}$ , função de autocorrelação amostral.

Modelos de séries temporais pode ser dividido em duas classes:

- i) Paramétricos – número finito de parâmetros, análise é feita no domínio do tempo.
- ii) Não paramétricos – número infinito de parâmetros, análise é feita no domínio da frequência.

Curso será basicamente sobre modelos paramétricos

Os modelos dessa classe podem ser escritos genericamente como:  $Y_t = S_t + \varepsilon_t$  ou seja:

Observação = sinal + ruído.

Assim temos:

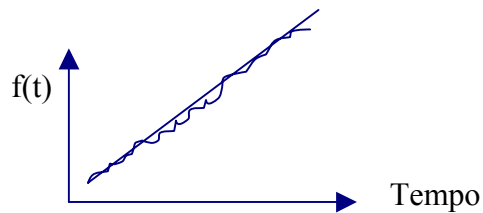
Modelos de regressão.

- Ruídos são não correlacionados
- $S_t = \underline{X}_t \underline{\beta}$

Exemplos

### 1. Modelos de tendência linear

$$S_t = (1 \quad t) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha + \beta t$$



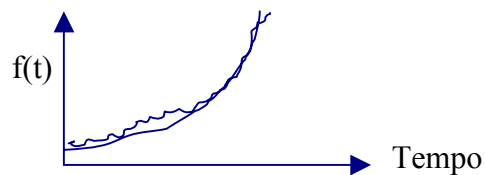
### 2. Modelo de curva de crescimento

$$Y_t = \alpha e^{\beta t + \varepsilon_t}$$

ou seja

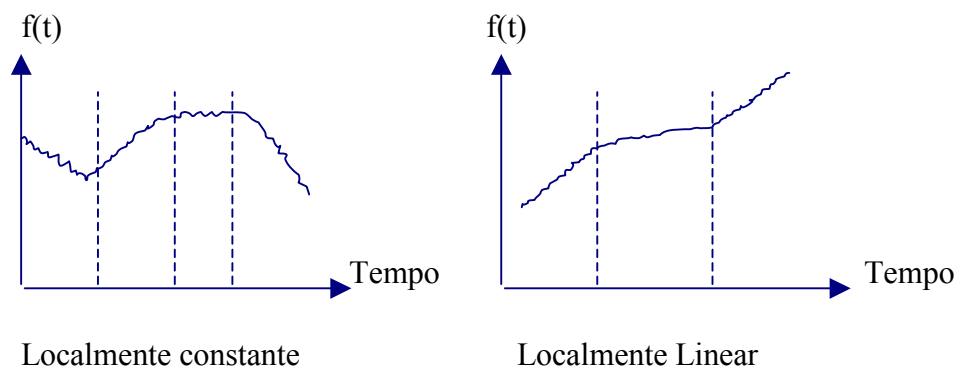
$$\log Y_t = \log \alpha + \beta t + \varepsilon_t$$

onde  $S_t = \log \alpha + \beta t$





Parâmetros sujeitos a pequenas variações temporais descreve a série da forma:



### Modelos Autorregressivos (AR)

Como em (a) com

$$S_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p}$$

Nível (Sinal) atual depende dos níveis passados

### Modelos Lineares estacionários

$$S_t = \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots + \psi_p \varepsilon_{t-p} = \sum_{i=1}^{\infty} \psi_i \varepsilon_{t-i}$$

Este processo é estacionário se  $\sum_{i=1}^{\infty} \psi_i^2 < \infty$

Inclui os modelos ARMA, que inclui os modelos AR.

Os modelos ARIMA são uma generalização dos modelos ARMA, que visam basicamente tornar o processo estacionário através de operações de diferença.

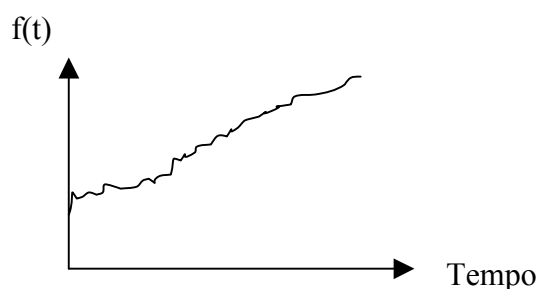
### - Fenômenos típicos de Séries Temporais.

Boa parte das séries tem características típicas

As principais são;

#### i) Tendência

É o efeito de longo prazo na média. Especificação de longo prazo é difícil.



A série acima apresenta tendência de crescimento linear .

#### ii) Sazonalidade.

Efeitos ligados a variações periódicas (semanal, mensal, anual, ...) exemplo: Medidas de temperatura: inverno, verão.

#### iii) Ciclos

Variações que apesar de periódicas não são associáveis automaticamente a nenhuma medida temporal.

Exemplos: Ciclos econômicos (5 a 7 anos), ciclos de epidemias.

Uma das tarefas mais importantes em séries temporais é identificar essas componentes visando a decomposição da série estudada.

**Modelos de regressão.**

Conforme visto nos capítulos anteriores podemos pensar a série temporal como uma coleção de observações determinadas por um sinal dependendo de uma forma determinista do tempo às quais são super impostas erros não correlacionados.

Procura-se nesse caso, usar modelos de regressão para caracterizar o sinal que controla a série.

Exemplos:

1) Modelos de tendência linear:  $S_t = \alpha + \beta t$

2) Modelos de crescimento exponencial:  $Y_t = \alpha e^{\beta t + \varepsilon_t}$

3) Modelos de regressão linear simples :  $S_t = \alpha + \beta X_t$

4) Modelos de regressão não-linear simples :  $S_t = 1/(\alpha + \beta X_t)$

Os modelos 2 e 4 não são lineares nos parâmetros embora possa ser linearizado através das transformações logarítmica dando

$$\log Y_t = \log \alpha + \beta t + \varepsilon_t$$

Os modelos a serem considerados serão lineares nos parâmetros.

Exemplo: Um modelo que pode ser usado para séries temporais descrevendo uma curva S é  $\log S_t = \alpha + \beta t$  que embora não seja linear em t é linear em  $\alpha$  e  $\beta$  após transformações logarítmicas.

Revisão de Modelos Lineares

Variável independente: Y

Variável explicativa:  $X_1, X_2, \dots, X_p$

Relacionadas através de  $y_t = \beta_0 + \sum_{i=1}^p \beta_i x_{it} + \varepsilon_t \quad t = 1, \dots, n$

Onde:  $S_t = \beta_0 + \sum_{i=1}^p \beta_i x_{it}$ ,

podemos escrever,  $S_t = \underline{x}_t^T \underline{\beta}$ ,  $\underline{x}_t^T = (1, x_{1t}, x_{2t}, \dots, x_{pt})$ ,  $\underline{\beta}^T = (\beta_0, \beta_1, \dots, \beta_p)$

os erros  $\varepsilon_t$  são não correlacionados com  $E\{\varepsilon_t\} = 0$ ,  $V\{\varepsilon_t\} = \sigma^2$ , comumente se assume também que  $\varepsilon_t$  são normais.

Usando notação matricial pode-se escrever  $\underline{y} = \underline{x}\underline{\beta} + \underline{\varepsilon}$ , onde

$\underline{y}^T = (y_1, y_2, \dots, y_n)$ ,  $\underline{x}^T = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$  e  $E\{\underline{\varepsilon}\} = 0$ ,  $V\{\underline{\varepsilon}\} = \sigma^2 I_n$ .

Critério de estimação para  $\underline{\beta}$  é a minimização de  $S(\underline{\beta}) = \sum_{t=1}^n (y_t - \underline{x}_t^T \underline{\beta})^2 = (\underline{y} - \underline{x}\underline{\beta})^T (\underline{y} - \underline{x}\underline{\beta})$

Para  $\hat{\underline{\beta}}$  minimizar  $S(\underline{\beta})$  é preciso que  $\underline{x}^T \underline{x} \hat{\underline{\beta}} = \underline{x}^T \underline{y}$  se  $\underline{x}$  tem posto máximo. (variáveis explicativas são linearmente independentes)

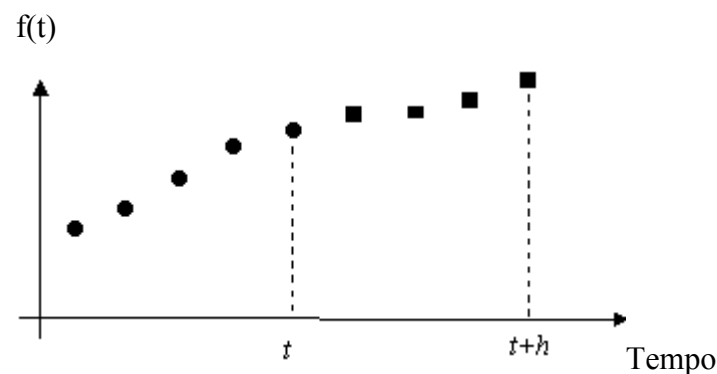
Então  $\underline{x}^T \underline{x}$  pode ser invertido fornecendo  $\hat{\underline{\beta}} = (\underline{x}^T \underline{x})^{-1} \underline{x}^T \underline{y}$

## 2.4. Previsão de Séries Temporais

Uma previsão é uma manifestação relativa a sucessos desconhecidos em um futuro determinado. A previsão não constitui um fim em si, mas um meio de fornecer informações e subsídios para uma conseqüente tomada de decisão, visando atingir determinados objetivos.

Considerando um conjunto de observações de uma série temporal coletadas até o instante  $t$  e de um modelo que represente esses fenômenos, a previsão do valor da série no tempo  $t+h$  (figura 1) pode ser obtida.

Figura 1. Observações de uma série temporal com previsões de origem  $t$  e horizonte  $h$ .



Dependendo do valor assumido pelo horizonte de previsão, [SOUZA1989] classifica as previsões dos valores futuros de uma série temporal como de curto, médio ou longo prazo. Desse modo, diante da possibilidade de existência de diferentes horizontes de previsão, se especifica técnicas distintas para prognosticar os valores futuros de uma série temporal :

- *previsão múltiplos passos* : esta abordagem, adotada para longos horizontes de previsão, procura identificar as tendências gerais e os pontos de inflexão mais relevantes na série temporal. Na *previsão múltiplos passos*, o conjunto de valores correntes é empregado na realização da previsão para determinado instante; esta previsão é, então, introduzida entre as observações passadas, compondo, desta forma, um novo conjunto de dados, sobre o qual será obtida a previsão do tempo subsequente.
- *previsão simples passo* : nesta técnica não há incorporação de previsões aos dados utilizados para encontrar a próxima previsão, sendo esta, independente dos valores anteriormente previstos. A previsão é feita apenas para o período de tempo imediatamente posterior ao atual, a partir das observações da série temporal.

Para [SOUZA1989], a garantia da otimalidade das previsões de uma série temporal somente é alcançada adotando como horizonte de previsão o instante de tempo imediatamente

subseqüente à origem  $t$ .

Naturalmente, a investigação do poder preditivo do modelo especificado como o mais adequado para explicar o mecanismo gerador das observações de uma série temporal é um processo empírico de verificação, visto que são feitas comparações entre as previsões e as observações, a fim de confirmar a habilidade do modelo matemático em descrever a estrutura definida pelos dados da série temporal analisada. É importante salientar que o caráter previsional de um determinado modelo não pode ser sustentado quando forem identificadas mudanças estruturais nas observações obtidas a partir da origem, em relação àquelas utilizadas para a elaboração do modelo.

## 2.5 Métodos de Previsão de Séries Temporais

Um método de previsão considera-se como um conjunto de procedimentos usados no desenvolvimento de uma determinada previsão.

Os métodos de previsão de séries temporais, classificados como métodos quantitativos, baseiam suas previsões na extrapolação de características de observações passadas e no inter-relacionamento entre essas observações, fornecendo previsões acuradas se o futuro apresentar comportamento similar ao passado .

A maioria dos métodos de previsão de séries temporais se baseia na suposição de que observações passadas contém todas as informações sobre o padrão de comportamento da série temporal e esse padrão é recorrente no tempo. O propósito dos métodos de previsão consiste em distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse padrão para prever os valores futuros da série temporal. Assim, pela identificação desse componente, a previsão para períodos de tempo subseqüentes ao observado pode ser desenvolvida.

Apesar de quase a totalidade dos métodos de previsão de séries temporais estar fundamentada apenas na análise das observações da série de interesse para a especificação de algum modelo que descreva essas observações, alguns procedimentos de previsão tentam explicar o comportamento de uma série temporal pela evolução dos fenômenos observacionais de outras

séries. Desta forma, dependendo do número de séries temporais envolvidas na modelagem, [SOUZA1989] classifica os métodos de previsão em *univariados*, *funções de transferência* e *multivariados*.

Os *métodos univariados*, que compreendem a maior parte dos métodos de previsão de séries temporais, consideram somente uma única série para a realização dos prognósticos. As previsões decorrentes da aplicação de métodos univariados podem estar relacionadas apenas com as informações contidas na série histórica de interesse (métodos baseados na estatística clássica) ou também, além de incorporarem essas informações, consideram outras supostamente relevantes e que não estão contidas na série analisada (métodos baseados na estatística bayesiana).

Aquelas metodologias nas quais a série de interesse é explicada não só pelo seu passado histórico, como também por outras séries temporais não-correlatas entre si, são conhecidas como *funções de transferência*. Esta classe de métodos de previsão envolve, portanto, mais de uma série temporal, com a ressalva de que a relação de causalidade entre estas séries é perfeitamente conhecida.

Os *métodos multivariados* abrangem os procedimentos de previsão que associam mais de uma série temporal na efetivação de prognósticos sem, no entanto, qualquer imposição com relação à causalidade entre essas séries.

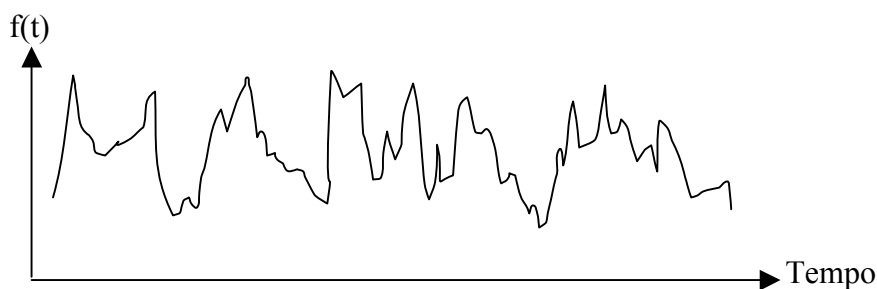
Diante disso, há uma variedade enorme de métodos de previsão de séries temporais, cada qual com suas capacidades e limitações. Qualquer que seja a classificação desses métodos, é possível utilizar um número muito grande de métodos diferentes para descrever o comportamento de uma série particular. A seleção do método de previsão adequado depende de vários fatores, tais como o comportamento do fenômeno observável ou o conhecimento a priori que se tenha sobre a sua natureza e do objetivo da análise.

[SOUZA1989] relaciona os métodos de previsão baseados exclusivamente em uma única série histórica de dados em métodos de *decomposição*, métodos *simples* de previsão e métodos *avançados* de previsão de séries temporais.

### **Processos Estacionários**

Como a quantidade de parâmetros é usualmente maior que o número de observações são necessárias hipóteses simplificadas.

A mais comum em séries temporais é a de estacionariedade. Basicamente isto significa que o comportamento da série não se altera com o passar do tempo, ou seja, média e facv não mudam se caminhamos no tempo.  $\Delta^n y_t = \Delta(\Delta^{n-1} y_t)$



Como normalmente as séries não apresentam esse comportamento estável, recorre-se a transformações dos dados.

As transformações mais comuns são

#### i) **Transformação Box Cox**

Considere a função da forma

$$g(y) = \frac{y^\lambda - 1}{\lambda}$$

se  $\lambda = 1$ ,  $g$  é a identidade

se  $\lambda = -1$ ,  $g$  é a transformação inversa

se  $\lambda = 1/2$ ,  $g$  transforma  $y$  em  $\sqrt{y}$

$$\text{e } \lim_{\lambda \rightarrow 0} g(y) = \log y \quad \lim_{\lambda \rightarrow 0} g(y) = \log y .$$

Essas transformações são usadas principalmente para estabilizar a variância



**Operação diferença ( $\Delta$ )**

Define-se o operador diferença através de  $\Delta y_t = y_t - y_{t-1}$

Aplicando-se o operador novamente obtém-se a 2ª diferença  $\Delta^2 y_t = \Delta(\Delta y_t) = \Delta(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$

$$\Delta^2 y_t = \Delta(\Delta y_t) = \Delta(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$$

A n-ésima diferença de  $y_t$  é obtida recursivamente por

$$\Delta^n y_t = \Delta(\Delta^{n-1} y_t)$$

Normalmente uma ou duas diferenças são suficientes para tornar a série estacionária.

- Estrita (ou Forte)
- Fraca (ou ampla ou de 2ª ordem).

Um processo estocástico  $Y(t)$  é estritamente estacionário sse suas distribuições finito dimensionais são invariantes por translações no tempo, isto é;

$$F(y_1, y_2, \dots, y_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau) = F(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n), \forall t_1, \dots, t_n, \tau$$

O processo deslocado  $\tau$  unidades no tempo permanece com as mesmas características.

Em particular com  $n=1$  e  $t_2 = t_1 + \tau$  temos que,

$$F(y_1; t_2) = F(y_2; t_1) \text{ e por tanto } \mu = \mu(t) \text{ e } \sigma^2 = \sigma^2(t) \text{ são constantes}$$

$$\text{Alem disso, } F(y_1, y_2; t_1 + \tau, t_2 + \tau) = F(y_1, y_2; t_1, t_2)$$

Logo

$$\text{Cov}\{Y(t_1), Y(t_2)\} = \gamma(t_1, t_2) = \gamma(t_1 - t_2) \quad \gamma(t_1, t_2) = \gamma(t_1 + \tau, t_2 + \tau), \quad \text{fazendo } \tau = -t_2 \text{ e } t = t_1 - t_2, \text{ temos que: } \gamma(t_1, t_2) = \gamma(t_1 - t_2, 0) = \gamma(t, 0) = \gamma(t)$$

A facv depende apenas da distancia dos pontos considerados.

Um processo estocástico  $Y(t)$  é fracamente estacionário sse:

- i)  $\mu(t) = E\{Y(t)\} = \mu$
- ii)  $V\{Y(t)\} = \sigma^2(t) = \sigma$
- iii)  $\text{Cov}\{Y(t_1), Y(t_2)\} = \gamma(t_1, t_2) = \gamma(t_1 - t_2)$

Se os momentos existem então: Estacionariedade Forte  $\longrightarrow$  Estacionariedade Fraca

A volta só vale se:

Distribuição finito dimensionais de  $Y(t)$  são normais (Processo Gaussiano).

Observe que a função de autocorrelação de processo estacionário é

$$\rho(t_1, t_2) = \frac{\sigma(t_1, t_2)}{\sigma(t_1)\sigma(t_2)} = \frac{\gamma(t_1 - t_2)}{\sigma^2(t_1)} = \frac{\gamma(t)}{\gamma(0)} = \rho(t) \quad \rho(t_1, t_2) = \frac{\sigma(t_1, t_2)}{\sigma(t_1)\sigma(t_2)} = \frac{\gamma(t_1 - t_2)}{\sigma^2(t_1)} = \frac{\gamma(t)}{\gamma(0)} = \rho(t)$$

## 2.6. Métodos de Decomposição de Séries Temporais

Os métodos de decomposição assumem que uma série temporal é constituída por um conjunto de componentes não-observáveis [SOUZA1989]. Dessa forma, pela identificação das componentes individuais presentes no padrão básico da série histórica de dados (tendência, ciclo, sazonalidade e aleatoriedade), a extrapolação para o futuro pode ser realizada .

Expressa-se o relacionamento entre as componentes não-observáveis da série temporal através da equação (1.1).

$$x_t = f( S_t, T_t, C_t, E_t ) \quad (1.1)$$

onde  $S_t$  corresponde à componente sazonal para o período  $t$ ;

$T_t$  é a componente de tendência no período  $t$ ;

$C_t$  é a componente de ciclo no período  $t$  e

$E_t$  é a componente aleatória no período  $t$ .

A componente sazonal representa as flutuações da série de acordo com algum fator de sazonalidade. O ciclo apresenta um comportamento similar à componente sazonal, embora tenha normalmente comprimento maior que aquela. Justamente pelo fato de não apresentar duração uniforme, a identificação da componente ciclo é mais problemática. A tendência representa o aumento ou declínio gradual nos valores das observações de uma série temporal. Com a remoção das componentes de sazonalidade, ciclo e tendência, a componente aleatória

fica determinada.

Vários procedimentos para a decomposição de séries temporais foram desenvolvidos, cada qual tentando isolar as componentes não-observáveis da série o mais acuradamente possível. O objetivo desses procedimentos consiste em remover cada uma das componentes, permitindo que o comportamento da série temporal seja melhor compreendido e, conseqüentemente, prognosticar valores futuros mais apropriados.

## 2.7. Métodos Simples de Previsão de Séries Temporais

Os métodos assim classificados efetuam a previsão do valor futuro da série temporal pelo alisamento das observações passadas da série de interesse. Assumindo que os valores extremos da série representam flutuações aleatórias, o propósito desses métodos consiste em identificar o padrão básico presente nos dados históricos e, então, usar esse padrão para prever valores futuros.

Associa-se a grande popularidade desses métodos à simplicidade, à eficiência computacional e à razoável previsão obtida. Entre os métodos simples de previsão destacam-se o da *Média Móvel*, o *Alisamento Exponencial Simples*, o *Alisamento Exponencial Linear* e o *Alisamento Exponencial Sazonal e Linear de Winter*, os quais são apresentados sucintamente na seqüência.

### 2.7.1. Média Móvel

Esse método considera como previsão para o período futuro a média das observações passadas recentes.

A média móvel para o período de tempo  $t$  é definida por

$$x_t = \frac{x_{t-1} + x_{t-2} + \dots + x_{t-n}}{n} \quad (1.2)$$

onde  $n$  representa o número de observações incluídas na média  $x_t$ .

O termo *média móvel* é utilizado porque à medida que a próxima observação se torna disponível, a média das observações é recalculada, incluindo essa observação no conjunto de observações e desprezando a observação mais antiga.

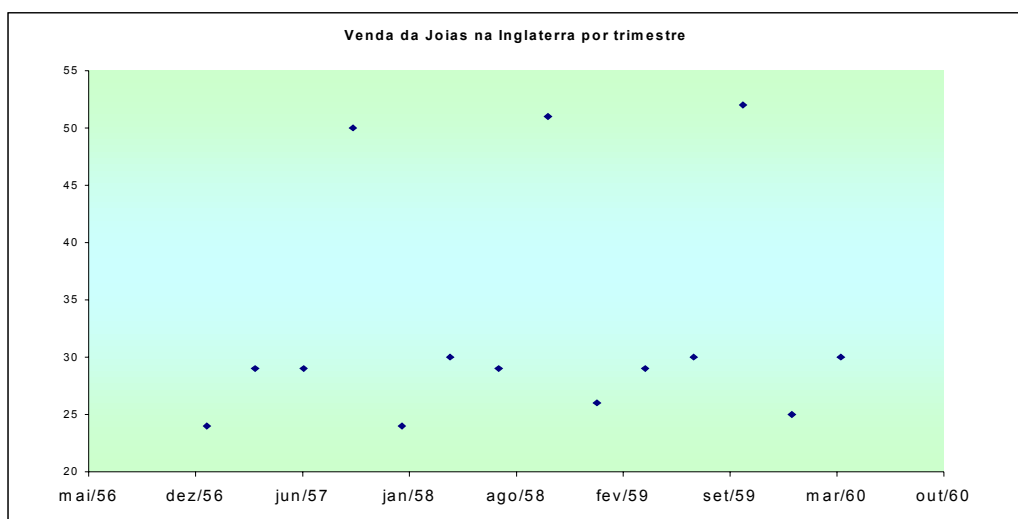
coloca que quanto maior o número de observações incluídas na média móvel, maior o efeito de alisamento na previsão. Assim, caso a série temporal apresente muita aleatoriedade ou pequenas mudanças nos padrões dessa série, um número maior de valores podem ser utilizados no cálculo da média móvel, obtendo-se uma previsão mais alisada. Entretanto, se houver pouca flutuação aleatória nos dados ou mudança significativa no padrão da série, um número menor de observações deve ser incluído no conjunto de valores empregado na determinação da média para que se possa reagir a essas alterações mais rapidamente.

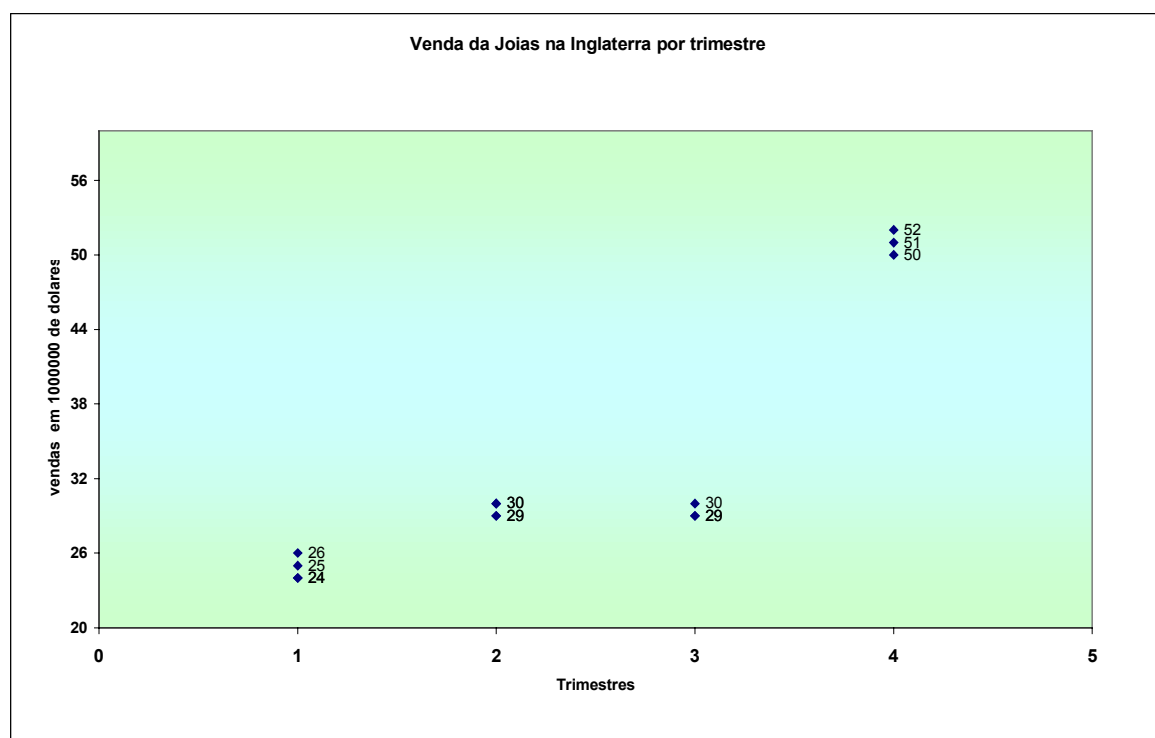
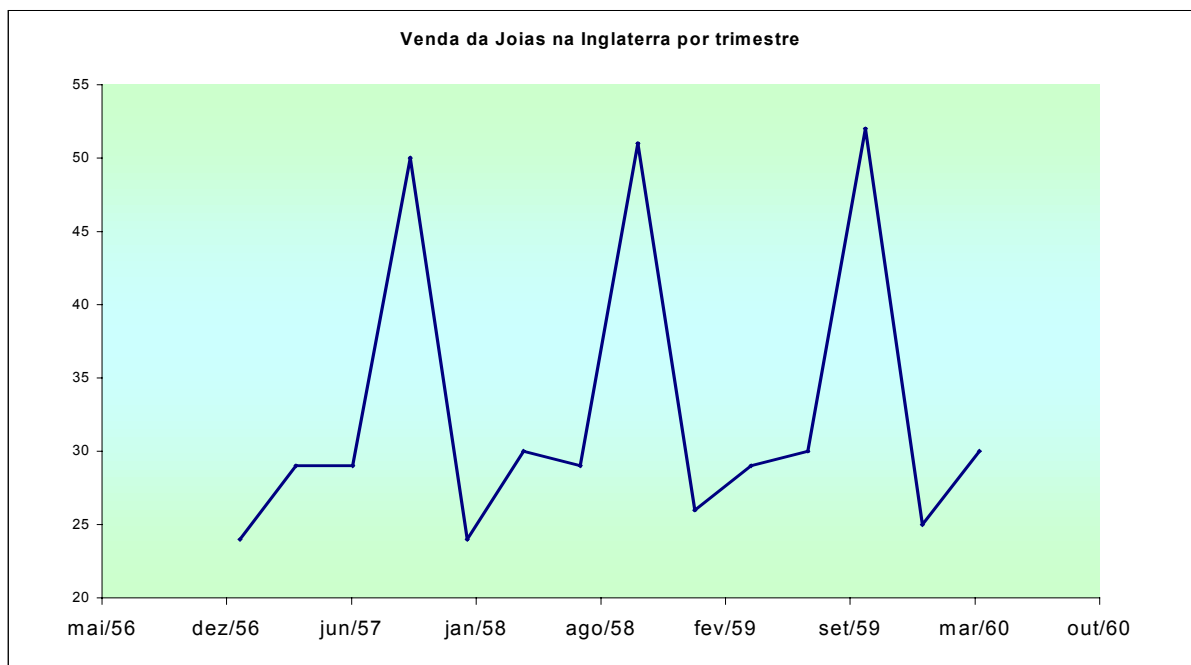
**Exemplo1** : Venda da Jóias na Inglaterra por ano( em milhões de dólares)

Data	em milhões
jan/57	24
abr/57	29
jul/57	29
out/57	50
jan/58	24
abr/58	30
jul/58	29
out/58	51
jan/59	26
abr/59	29
jul/59	30
out/59	52
jan/60	25

Considerando os dados acima pode-se dividir cada ano em trimestres da seguinte forma

Data	em milhões	Trimestrais	Média móvel
jan/57	24	1	
abr/57	29	2	33
jul/57	29	3	33
out/57	50	4	33,25
jan/58	24	1	33,25
abr/58	30	2	33,5
jul/58	29	3	34
out/58	51	4	33,75
jan/59	26	1	34
abr/59	29	2	34,25
jul/59	30	3	34
out/59	52	4	34,25
jan/60	25		





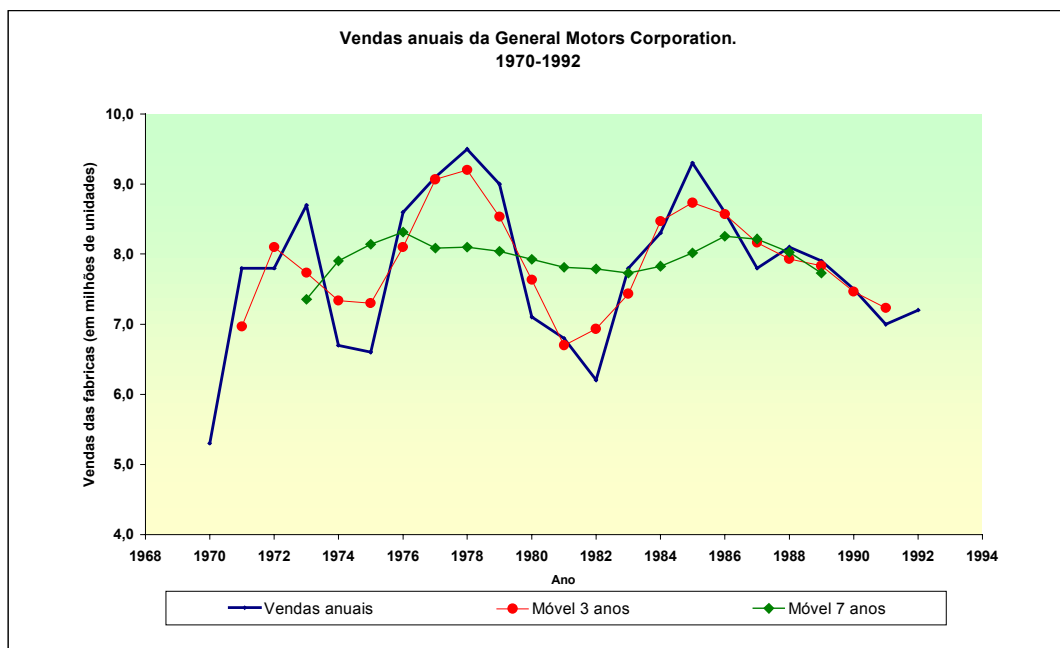
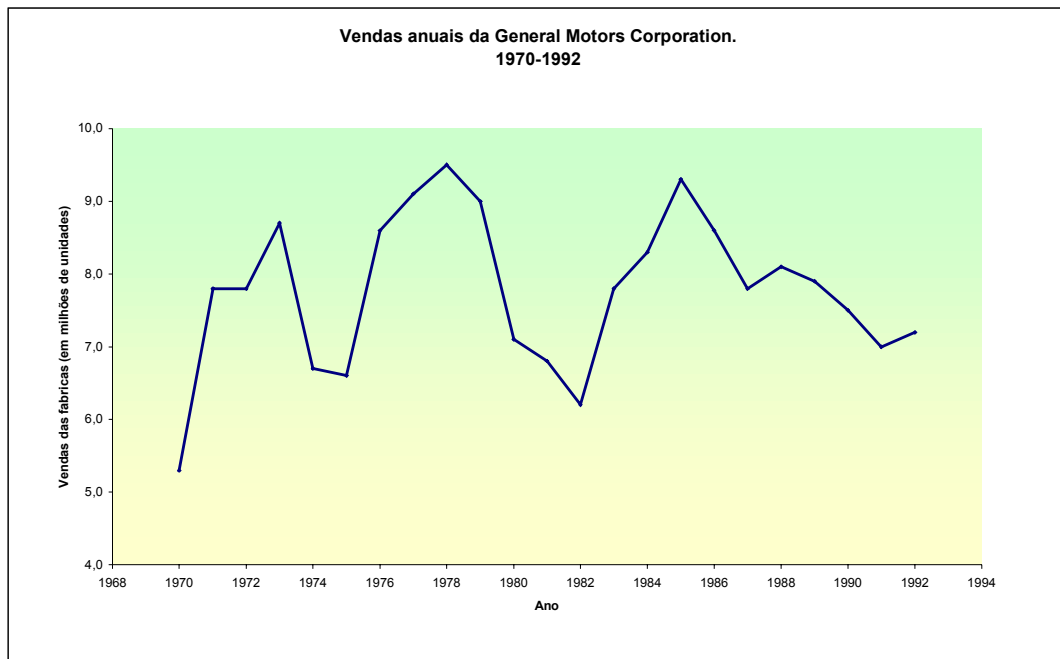
Vendas anuais das fabricas da General Motors Corporation (em milhões de unidades) 1970 – 1992

Ano	Vendas das fábricas	Ano	Vendas das fábricas
1970	5,3	1982	6,2
1971	7,8	1983	7,8
1972	7,8	1984	8,3
1973	8,7	1985	9,3
1974	6,7	1986	8,6
1975	6,6	1987	7,8
1976	8,6	1988	8,1
1977	9,1	1989	7,9
1978	9,5	1990	7,5
1979	9,0	1991	7,0
1980	7,1	1992	7,2
1981	6,8		

Fonte: Moody's Handbook of Common Stocks

Ano	Vendas das fábricas	Total movel de três anos	Média movel de três anos	Total movel de sete anos	Média movel de sete anos
1970	5,3	--	--	--	--
1971	7,8	20,9	6,97	--	--
1972	7,8	24,3	8,10	--	--
1973	8,7	23,2	7,73	51,5	7,36
1974	6,7	22,0	7,33	55,3	7,90
1975	6,6	21,9	7,30	57,0	8,14
1976	8,6	24,3	8,10	58,2	8,31
1977	9,1	27,2	9,07	56,6	8,09
1978	9,5	27,6	9,20	56,7	8,10
1979	9,0	25,6	8,53	56,3	8,04
1980	7,1	22,9	7,63	55,5	7,93
1981	6,8	20,1	6,70	54,7	7,81
1982	6,2	20,8	6,93	54,5	7,79
1983	7,8	22,3	7,43	54,1	7,73
1984	8,3	25,4	8,47	54,8	7,83
1985	9,3	26,2	8,73	56,1	8,01
1986	8,6	25,7	8,57	57,8	8,26
1987	7,8	24,5	8,17	57,5	8,21
1988	8,1	23,8	7,93	56,2	8,03
1989	7,9	23,5	7,83	54,1	7,73
1990	7,5	22,4	7,47	--	--
1991	7,0	21,7	7,23	--	--
1992	7,2	--	--	--	--





### 2.7.2. Alisamento Exponencial Simples

A princípio, o método conhecido como Alisamento Exponencial Simples se assemelha ao da Média Móvel por extrair das observações da série temporal o comportamento aleatório pelo alisamento dos dados históricos. Entretanto, a inovação introduzida pelo Alisamento Exponencial Simples advém do fato de este método atribuir pesos diferentes a cada observação da série. Enquanto que na Média Móvel as observações usadas para encontrar a previsão do valor futuro contribuem em igual proporção para o cálculo dessa previsão, no Alisamento Exponencial Simples as informações mais recentes são evidenciadas pela aplicação de um fator que determina essa importância.

Segundo Wheelwright, o argumento para o tratamento diferenciado das observações da série temporal é fundamentado na suposição de que as últimas observações contém mais informações sobre o futuro e, portanto, são mais relevantes para a previsão.

especifica o método Alisamento Exponencial Simples através da equação (1.3).

$$F_{t+1} = \alpha x_t + (1 - \alpha) F_t \quad (1.3)$$

onde  $F_{t+1}$  representa a previsão no tempo  $t + 1$  e

$\alpha$  é o peso atribuído à observação  $x_t$ ,  $0 < \alpha < 1$ .

O valor assumido determina o ajuste aplicado aos dados. Quanto menor o valor da constante, mais estáveis serão as previsões, visto que a utilização de baixo valor de  $\alpha$  implica na atribuição de peso maior às observações passadas e, conseqüentemente, qualquer flutuação aleatória no presente contribui com menor importância para a obtenção da previsão. Contudo, não há metodologia que oriente quanto à seleção de um valor apropriado para  $\alpha$ , sendo normalmente encontrado por tentativa e erro. Um procedimento mais objetivo seria a seleção do valor de  $\alpha$  que forneça a "melhor previsão das observações contidas na série temporal".

### 2.7.3. Alisamento Exponencial Linear

Quando o método Alisamento Exponencial Simples é aplicado na previsão de séries temporais que apresentam tendência entre as observações passadas, os valores prognosticados superestimam (ou subestimam) os valores reais. Desta forma, a acuidade das previsões fica prejudicada.

Para evitar esse erro sistemático, o método Alisamento Exponencial Linear foi desenvolvido procurando reconhecer a presença de tendência na série de dados. O valor da previsão obtido através deste método é alcançado pela aplicação da equação (1.4).

$$F_{t+m} = S_t + mT_t \quad (1.4)$$

onde  $S_t$  corresponde à previsão no tempo  $t$ , conforme equação (1.5);

$T_t$  representa a componente de tendência, obtida pela equação (1.6) e

$m$  é o horizonte de previsão.

$$S_t = \alpha x_t + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (1.5)$$

Onde  $\alpha$  é o peso atribuído à observação  $x_t$ ,  $0 < \alpha < 1$ .

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \quad (1.6)$$

onde  $\beta$  é o coeficiente de alisamento, análogo a  $\alpha$ .

### 2.7.4. Alisamento Exponencial Sazonal e Linear de Winter

Este método produz resultados similares ao Alisamento Exponencial Linear, sendo, no

entanto, capaz de manipular séries temporais que além de apresentarem tendência nos dados, apresentam também sazonalidade.

As equações (1.7), (1.8), (1.9) e (1.10) definem o referido método de previsão.

$$S_t = \alpha \frac{x_t}{I_{t-l}} + (1 - \alpha)(S_{t-1} + T_{t-1}) \quad (1.7)$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{t-1} \quad (1.8)$$

$$I_t = \gamma \frac{x_t}{S_t} + (1 - \gamma)I_{t-l} \quad (1.9)$$

$$F_{t+m} = (S_t + mT_t)I_{t-l+m} \quad (1.10)$$

onde  $I_t$  corresponde ao alisamento do fator de sazonalidade  $\frac{x_t}{S_t}$ ;

$l$  é o intervalo da sazonalidade e  $\gamma$  corresponde ao peso atribuído ao fator de sazonalidade.

### Ajuste exponencial

Ajuste exponencial é uma outra técnica que podemos utilizar para ajustar uma série temporal e, conseqüentemente, ter uma impressão sobre os movimentos gerais de longo prazo que ocorrem nos dados. Além disso, podemos empregar o método de ajuste exponencial para obter previsões de curto prazo. Que já é uma vantagem em relação ao método de médias móveis.

$$E_i = WY_i + (1-W)E_{i-1}$$

Onde:

$E_i$  : Valor da série exponencialmente ajustada, sendo calculada no período de tempo  $i$

$E_{i-1}$ : Valor da série exponencialmente ajustada, já calculada no período de tempo  $i-1$

$Y_i$  : Valor observado da série temporal no período  $i$

$W_i$  : Peso designado subjetivamente ou coeficiente de ajuste (onde  $0 < W < 1$ )

**Previsão:**  $\hat{Y}_{i+1} = E_i$

Vendas anuais das fabricas da General Motors Corporation  
(em milhões de unidades) 1970 - 1992

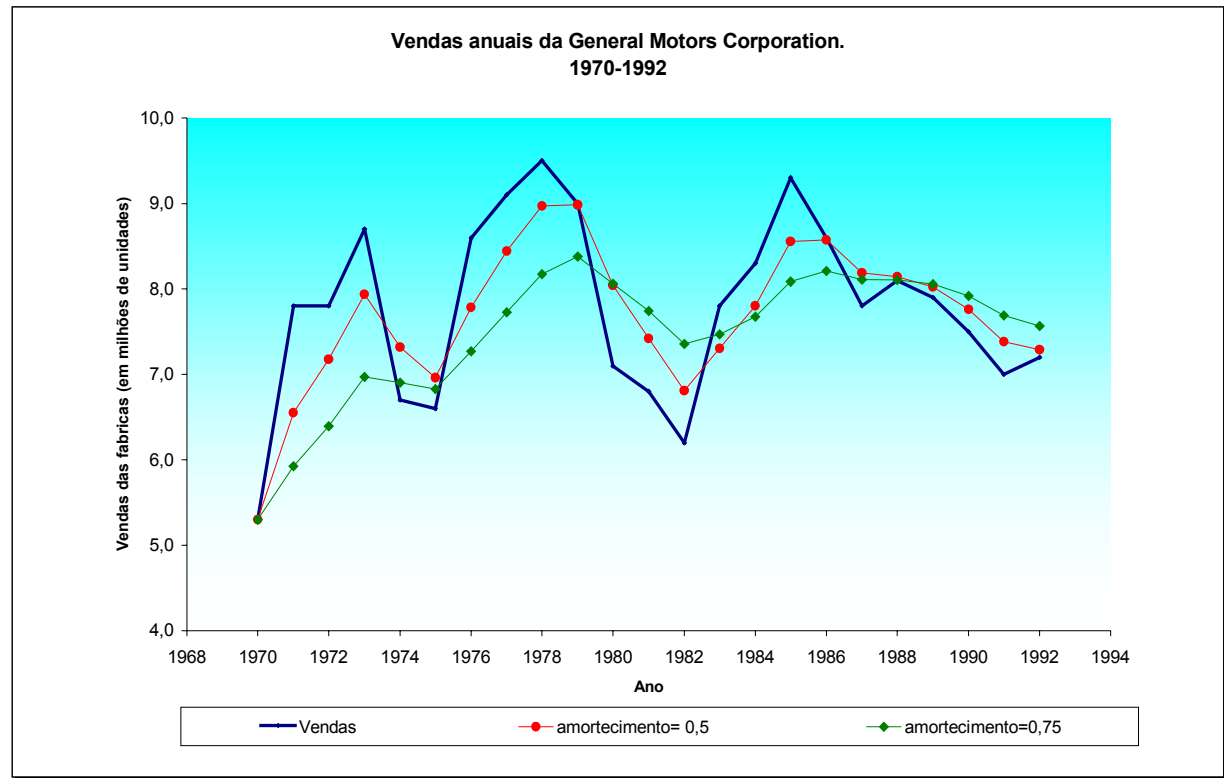
Ano	Vendas das fábricas
1970	5,3
1971	7,8
1972	7,8
1973	8,7
1974	6,7
1975	6,6
1976	8,6
1977	9,1
1978	9,5
1979	9,0
1980	7,1
1981	6,8
1982	6,2
1983	7,8
1984	8,3
1985	9,3
1986	8,6
1987	7,8
1988	8,1
1989	7,9
1990	7,5
1991	7,0
1992	7,2

Fonte: Moody's Handbook of Common Stocks

Vendas anuais das fabricas da General Motors Corporation (em milhões de unidades) 1970-1992

Ano	Vendas das fábricas	Amortecimento (1-W) = 0,5	Amortecimento (1-W) = 0,75
1970	5,3	5,30	5,30
1971	7,8	6,55	5,93
1972	7,8	7,18	6,39
1973	8,7	7,94	6,97
1974	6,7	7,32	6,90
1975	6,6	6,96	6,83
1976	8,6	7,78	7,27
1977	9,1	8,44	7,73
1978	9,5	8,97	8,17
1979	9,0	8,98	8,38
1980	7,1	8,04	8,06
1981	6,8	7,42	7,74
1982	6,2	6,81	7,36
1983	7,8	7,31	7,47
1984	8,3	7,80	7,68
1985	9,3	8,55	8,08
1986	8,6	8,58	8,21
1987	7,8	8,19	8,11
1988	8,1	8,14	8,11
1989	7,9	8,02	8,05
1990	7,5	7,76	7,92
1991	7,0	7,38	7,69
1992	7,2	7,29	7,57

Fonte: Moody's Handbook of Common Stocks



Amortecimento =  $1-W$

## Capítulo 3. Modelos ARIMA

### 3.1. Métodos Avançados de Previsão de Séries Temporais

No universo dos métodos de previsão de séries temporais mais complexos encontram-se os modelos Autoregressivo e Médias Móveis (AR, MA e ARMA), modelos Autoregressivo Integrado de Médias Móveis (ARIMA), Filtros de Kalman e AEP, modelos ARARMA de Parzen, modelos ARMA Multivariáveis (MARMA), entre outros. Os métodos assim classificados obtêm a previsão de algum valor futuro da série temporal pela combinação dos valores reais passados e/ou dos erros ocorridos.

Os modelos AR, MA, ARMA e ARIMA são descritos a seguir.

### 3.2. Modelos Autoregressivo e de Médias Móveis

descreve três procedimentos capazes de representar as observações de uma série temporal estacionária : modelo *Autoregressivo* (AR), modelo de *Médias Móveis* (MA) e o modelo *Autoregressivo e de Médias Móveis* (ARMA).

#### 3.2.1 Modelo Autoregressivo

A especificação de um modelo Autoregressivo (AR) é dada pela equação (1.11).

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + e_t \quad (1.11)$$

onde  $x_t$  corresponde à observação da série temporal no tempo  $t$ ;

$\phi_p$  corresponde ao parâmetro do modelo AR de ordem  $p$  e

$e_t$  representa o erro de eventos aleatórios que não podem ser explicados pelo modelo.

Caso as observações da série temporal possam ser representadas pela equação (1.11), a ordem do modelo puder ser determinada e os parâmetros estimados, é possível prever o valor futuro da série em análise.

#### 3.2.2 Modelo de Médias Móveis

Um modelo de Médias Móveis (MA) fica definido conforme equação (1.12) .

$$x_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (1.12)$$

onde  $e_t$  representa o erro de eventos aleatórios que não podem ser explicados pelo modelo e



$\theta_q$  corresponde ao parâmetro do modelo MA de ordem  $q$ .

A equação (1.12) é similar à equação (1.11), exceto pelo fato de que o valor previsto para a observação  $\alpha$  depende dos valores dos erros observados em cada período passado, ao invés das observações propriamente ditas.

### 3.2.3 Modelo Autoregressivo e de Médias Móveis

Wheelwright e Makridakis especificam o modelo misto Autoregressivo e de Médias Móveis (ARMA) através da equação (1.13), como sendo a combinação dos modelos AR e MA.

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (1.13)$$

Analisando a equação (1.13) é possível verificar que os modelos ARMA relacionam os valores futuros com as observações passadas, assim como também com os erros passados apurados entre os valores reais e os previstos.

### 3.2.4 Modelo de Box e Jenkins

O destaque atribuído ao modelo de George Box e Gwilyn Jenkins, que também pode ser incluído nesta classificação, é devido principalmente a sua fundamentação teórica, sendo a princípio capaz de manipular séries temporais de qualquer natureza.

O método de Box e Jenkins consiste na busca de um modelo ARIMA (AutoRegressive Integrate Moving Average) que represente o processo estocástico gerador da série temporal, a partir de um modelo ARMA aplicável na descrição de séries temporais estacionárias, estendendo esse conceito para séries temporais não-estacionárias.

Genericamente, um processo ARIMA( $p, d, q$ ) pode ser representado pela equação (1.14).

$$(1.14)$$

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

sendo  $w_t = x_t - x_{t-d}$ ,

onde  $\phi_p$  e  $\theta_q$  são os parâmetros dos processos Autoregressivo e de Média Móvel de ordem  $p$  e  $q$  (ARMA( $p,q$ ));

$e_t$  corresponde ao erro de eventos aleatórios que não podem ser explicados pelo modelo e  $d$  equivale ao grau de homogeneidade não-estacionária.

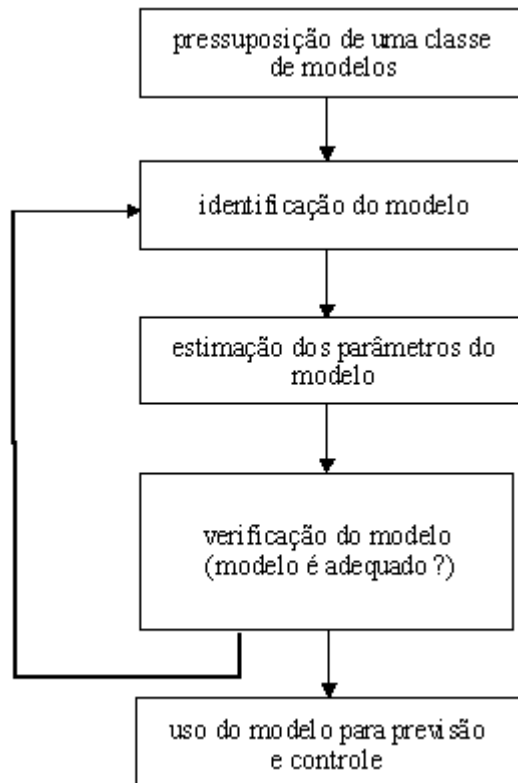
A estratégia para construção de um modelo ARIMA envolve uma abordagem iterativa que pode ser sumarizada conforme ilustra a figura 2.

O objetivo da *identificação* é determinar os valores de  $p$ ,  $d$  e  $q$  do modelo ARIMA( $p,d,q$ ). Inicialmente, a série temporal  $X^T$  é diferenciada para se obter uma série estacionária. Com isso, o processo fica reduzido a um modelo ARMA( $p,q$ ). Em seguida, a ordem do processo ARMA é identificada pela análise dos coeficientes de autocorrelação e autocorrelação parcial. Ainda nesta etapa são efetuadas estimativas preliminares dos parâmetros do modelo identificado.

Após a identificação do modelo que seja uma representação adequada do mecanismo gerador da série, a *estimação* dos parâmetros desse modelo é efetuada. Os parâmetros do processo AR são estimados através de métodos de regressão; caso o processo MA esteja envolvido, a estimação dos parâmetros deste modelo é obtida pela aplicação de algum algoritmo de otimização não-linear.

**Figura 2.** Estágios da construção de um modelo ARIMA.

Fonte : [BOX1976]



Estimado o modelo, a *verificação* de sua habilidade em representar os fenômenos observáveis da série temporal é confirmada pela análise dos erros do modelo proposto. Caso a inadequação fique evidenciada, o ciclo de identificação, estimação e verificação é novamente aplicado, até que a representação apropriada seja encontrada.

Após a validação do modelo, a *previsão* dos valores futuros da série temporal modelada pode, enfim, ser obtida.

Análise de séries temporais – é uma técnica diferente, útil para previsão de curto prazo.

A técnica de séries temporais relaciona os valores correntes de uma variável econômica com seus valores passados, eventualmente, com valores passados de outras variáveis e valores de erros aleatórios passados e presentes.

Os modelos que relacionam os valores correntes de uma variável com seus valores passados e com erros passados e correntes são chamados de modelos de séries temporais univariados.

Se os modelos incluem várias variáveis cujos seus valores correntes são explicados pelo conjunto dessas variáveis defasadas e de erros correntes e defasados são chamados de modelos de séries temporais multivariados.

Exemplo de um processo de séries temporais:

**Processo autoregressivo de 1ª ordem AR(1):**

$$Y_t = \alpha + \theta Y_{t-1} + e_t \text{ onde } -1 < \theta < 1$$

**Processo autoregressivo de 2ª ordem AR(2):**

$$Y_t = \alpha + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + e_t$$

**Processo de autoregressivo de ordem p AR(p):**

$$Y_t = \alpha + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + e_t$$

Os parâmetros deste tipo de modelo podem ser estimados por mínimos quadrados.

Exemplos

**AR**

Dados de desemprego no Brasil

		Período	Taxa %
		(IBGE) Brasil	
DESEMPREGO	1	JAN/96	5,26
	2	FEV/96	5,7
	3	MAR/96	6,38
	4	ABR/96	6,03
	5	MAI/96	5,91
	6	JUN/96	5,92
	7	JUL/96	5,58
	8	AGO/96	5,56
	9	SET/96	5,23

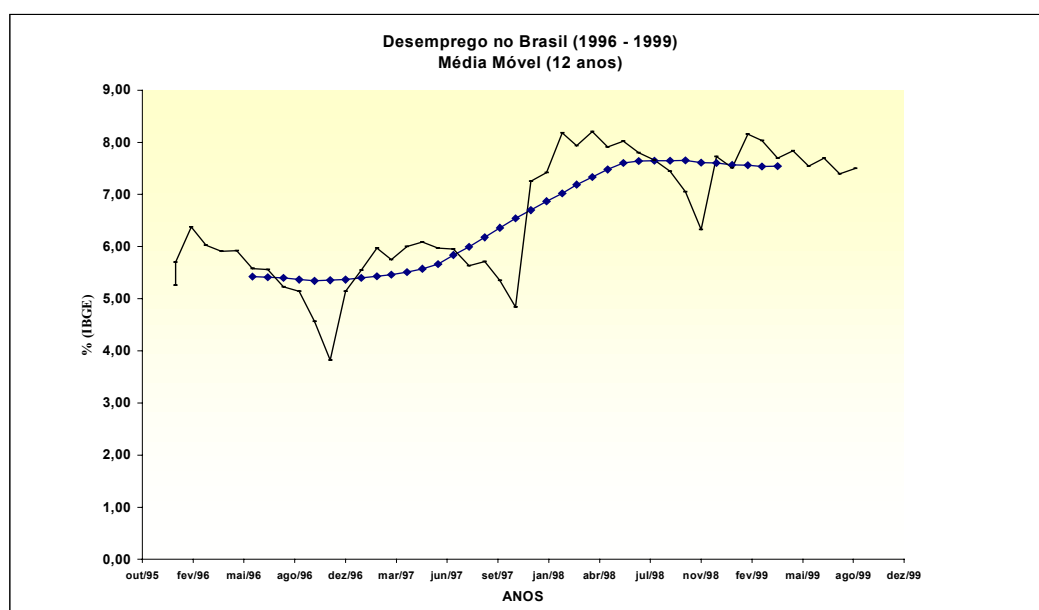
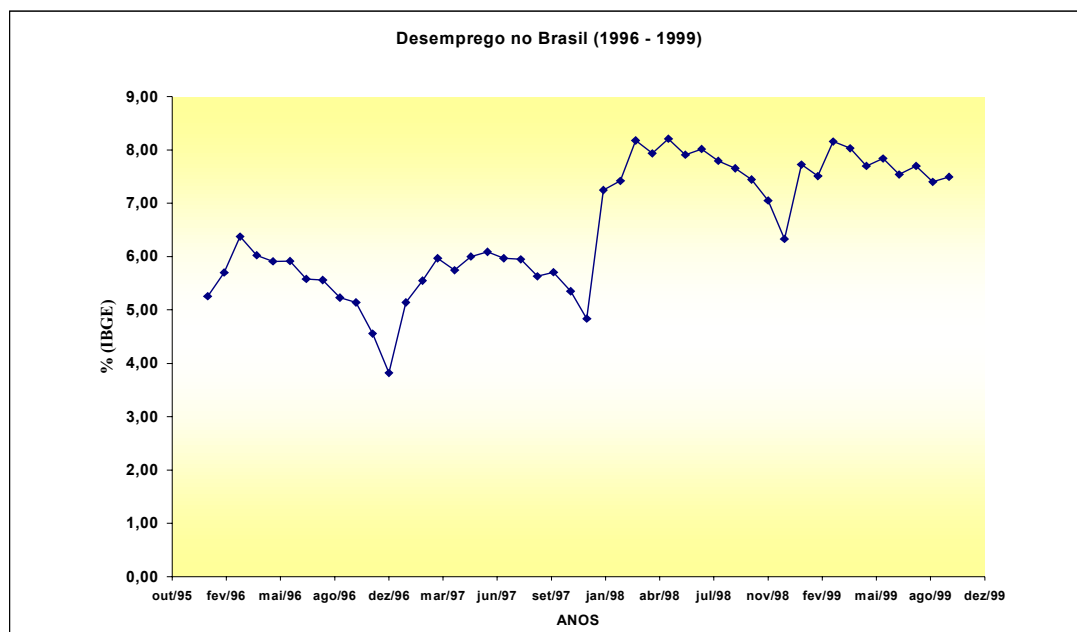
	10	OUT/96	5,14
	11	NOV/96	4,56
	12	DEZ/96	3,82
DESEMPREGO	13	JAN/97	5,14
	14	FEV/97	5,55
1997	15	MAR/97	5,97
	16	ABR/97	5,75
	17	MAI/97	6
	18	JUN/97	6,09
	19	JUL/97	5,97
	20	AGO/97	5,95
	21	SET/97	5,63
	22	OUT/97	5,71
	23	NOV/97	5,35
	24	DEZ/97	4,84
DESEMPREGO	25	JAN/98	7,25
	26	FEV/98	7,42
1998	27	MAR/98	8,18
	28	ABR/98	7,94
	29	MAI/98	8,21
	30	JUN/98	7,91
	31	JUL/98	8,02
	32	AGO/98	7,8
	33	SET/98	7,66
	34	OUT/98	7,45
	35	NOV/98	7,05
	36	DEZ/98	6,33
DESEMPREGO	37	JAN/99	7,73
	38	FEV/99	7,51
1999	39	MAR/99	8,16

40	ABR/99	8,03
41	MAI/99	7,7
42	JUN/99	7,84
43	JUL/99	7,54
44	AGO/99	7,7
45	SET/99	7,4
46	OUT/99	7,5

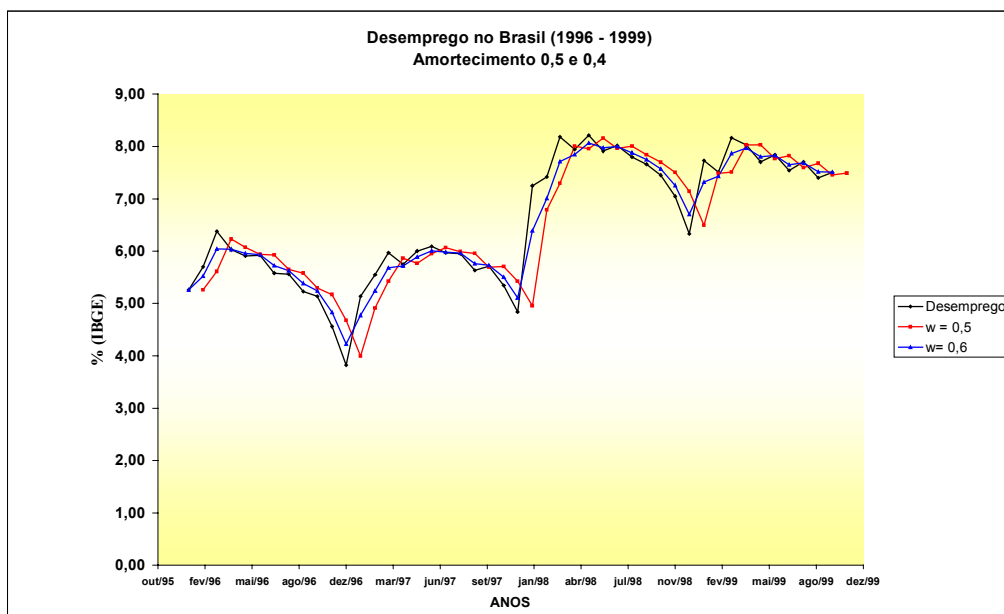
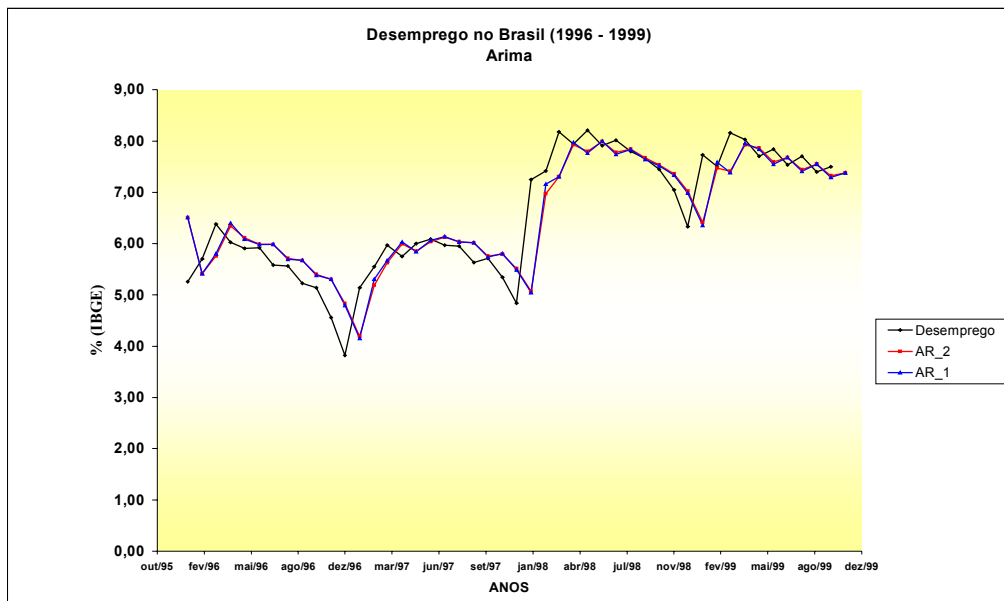
Período	Taxa %		AJUSTE		Modelos	Autorregressivos
	(IBGE) Brasil	MOVEL	EXPONENCIAL		AR 2	AR 1
			AMR=0.2	AMR=0.4		
JAN/96	5,26				6,51	6,51
FEV/96	5,70		5,26	5,26	5,41	5,41
MAR/96	6,38		5,61	5,52	5,76	5,80
ABR/96	6,03		6,23	6,04	6,34	6,40
MAI/96	5,91		6,07	6,03	6,11	6,09
JUN/96	5,92	5,42	5,94	5,96	5,99	5,98
JUL/96	5,58	5,41	5,92	5,94	5,99	5,99
AGO/96	5,56	5,40	5,65	5,72	5,71	5,69
SET/96	5,23	5,37	5,58	5,62	5,67	5,68
OUT/96	5,14	5,34	5,30	5,39	5,40	5,39
NOV/96	4,56	5,35	5,17	5,24	5,30	5,31
DEZ/96	3,82	5,37	4,68	4,83	4,83	4,80
JAN/97	5,14	5,40	3,99	4,22	4,19	4,15
FEV/97	5,55	5,43	4,91	4,77	5,19	5,31
MAR/97	5,97	5,46	5,42	5,24	5,63	5,67
ABR/97	5,75	5,51	5,86	5,68	6,00	6,04
MAI/97	6,00	5,58	5,77	5,72	5,85	5,84
JUN/97	6,09	5,66	5,95	5,89	6,04	6,06
JUL/97	5,97	5,84	6,06	6,01	6,13	6,14

AGO/97	5,95	5,99	5,99	5,99	6,04	6,04
SET/97	5,63	6,18	5,96	5,96	6,02	6,02
OUT/97	5,71	6,36	5,70	5,76	5,76	5,74
NOV/97	5,35	6,55	5,71	5,73	5,79	5,81
DEZ/97	4,84	6,70	5,42	5,50	5,51	5,49
JAN/98	7,25	6,87	4,96	5,11	5,07	5,05
FEV/98	7,42	7,02	6,79	6,39	6,97	7,16
MAR/98	8,18	7,19	7,29	7,01	7,30	7,31
ABR/98	7,94	7,34	8,00	7,71	7,93	7,97
MAI/98	8,21	7,48	7,95	7,85	7,80	7,76
JUN/98	7,91	7,60	8,16	8,07	8,00	8,00
JUL/98	8,02	7,64	7,96	7,97	7,78	7,74
AGO/98	7,80	7,65	8,01	8,00	7,84	7,83
SET/98	7,66	7,65	7,84	7,88	7,67	7,64
OUT/98	7,45	7,66	7,70	7,75	7,54	7,52
NOV/98	7,05	7,61	7,50	7,57	7,36	7,33
DEZ/98	6,33	7,61	7,14	7,26	7,02	6,98
JAN/99	7,73	7,57	6,49	6,70	6,41	6,35
FEV/99	7,51	7,56	7,48	7,32	7,48	7,58
MAR/99	8,16	7,54	7,50	7,43	7,41	7,39
ABR/99	8,03	7,54	8,03	7,87	7,92	7,96
MAI/99	7,70		8,03	7,97	7,87	7,84
JUN/99	7,84		7,77	7,81	7,59	7,55
JUL/99	7,54		7,83	7,83	7,68	7,68
AGO/99	7,70		7,60	7,65	7,45	7,41
SET/99	7,40		7,68	7,68	7,55	7,55
OUT/99	7,50		7,46	7,51	7,32	7,29
Previsão			7,49	7,51	7,38	7,38

---







**Desemprego No Brasil**

		(IBGE) Brasil	AJUSTE EXPONENCIAL				ARIMA			
			AMR=0.2		AMR=0.4		AR 2	AR 1		
			$(y - \hat{y})^2$		$(y - \hat{y})^2$			$(y - \hat{y})^2$	$(y - \hat{y})^2$	
1996	1 jan/96	5,26	27,668		27,668		6,509	1,560	6,511	1,565
	2 fev/96	5,70	5,26	0,194	5,260	0,194	5,414	0,082	5,414	0,082
	3 mar/96	6,38	5,61	0,590	5,524	0,733	5,756	0,389	5,800	0,337
	4 abr/96	6,03	6,23	0,039	6,038	0,000	6,339	0,096	6,396	0,134
	5 mai/96	5,91	6,07	0,025	6,033	0,015	6,113	0,041	6,089	0,032
	6 jun/96	5,92	5,94	0,000	5,959	0,002	5,988	0,005	5,984	0,004
	7 jul/96	5,58	5,92	0,119	5,936	0,127	5,986	0,165	5,993	0,170
	8 ago/96	5,56	5,65	0,008	5,722	0,026	5,713	0,023	5,694	0,018
	9 set/96	5,23	5,58	0,121	5,625	0,156	5,669	0,193	5,677	0,200
	10 out/96	5,14	5,30	0,025	5,388	0,061	5,402	0,069	5,387	0,061
	11 nov/96	4,56	5,17	0,374	5,239	0,461	5,303	0,551	5,308	0,560
	12 dez/96	3,82	4,68	0,744	4,832	1,023	4,828	1,016	4,800	0,960
1997	13 jan/97	5,14	3,99	1,317	4,225	0,838	4,185	0,912	4,151	0,979
	14 fev/97	5,55	4,91	0,409	4,774	0,602	5,188	0,131	5,308	0,058
	15 mar/97	5,97	5,42	0,300	5,240	0,534	5,626	0,119	5,668	0,091
	16 abr/97	5,75	5,86	0,012	5,678	0,005	5,997	0,061	6,036	0,082
	17 mai/97	6,00	5,77	0,052	5,721	0,078	5,854	0,021	5,844	0,024
	18 jun/97	6,09	5,95	0,018	5,888	0,041	6,037	0,003	6,063	0,001
	19 jul/97	5,97	6,06	0,009	6,009	0,002	6,130	0,026	6,142	0,029
	20 ago/97	5,95	5,99	0,001	5,986	0,001	6,041	0,008	6,036	0,007
	21 set/97	5,63	5,96	0,107	5,964	0,112	6,015	0,148	6,019	0,151
	22 out/97	5,71	5,70	0,000	5,764	0,003	5,756	0,002	5,738	0,001
	23 nov/97	5,35	5,71	0,128	5,731	0,146	5,794	0,197	5,808	0,210
	24 dez/97	4,84	5,42	0,338	5,503	0,439	5,511	0,450	5,493	0,426

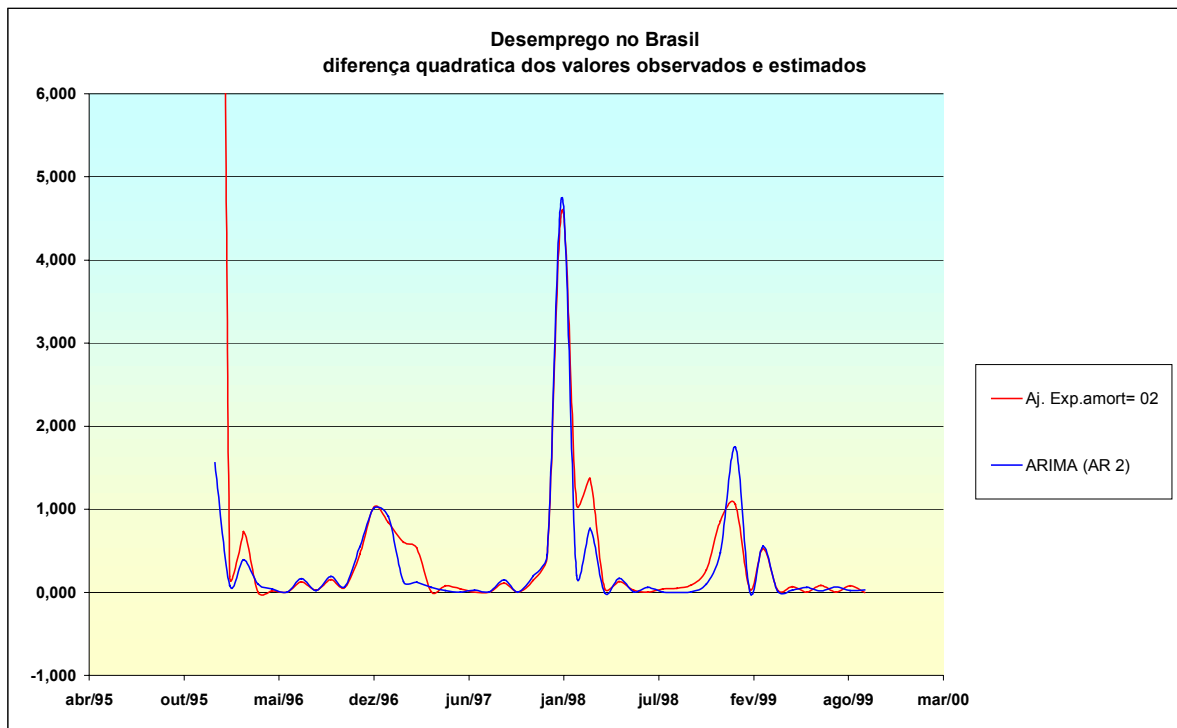
## Séries Temporais

67

Prof. Daniel Neyra Castañeda.

1998	25	jan/98	7,25	4,96	5,261	5,105	4,601	5,071	4,749	5,045	4,860
	26	fev/98	7,42	6,79	0,395	6,392	1,057	6,970	0,202	7,159	0,068
	27	mar/98	8,18	7,29	0,785	7,009	1,372	7,303	0,769	7,308	0,760
	28	abr/98	7,94	8,00	0,004	7,712	0,052	7,929	0,000	7,975	0,001
	29	mai/98	8,21	7,95	0,066	7,849	0,131	7,797	0,170	7,764	0,199
	30	jun/98	7,91	8,16	0,062	8,065	0,024	7,995	0,007	8,001	0,008
	31	jul/98	8,02	7,96	0,004	7,972	0,002	7,775	0,060	7,738	0,079
	32	ago/98	7,80	8,01	0,043	8,001	0,040	7,840	0,002	7,835	0,001
	33	set/98	7,66	7,84	0,033	7,880	0,049	7,671	0,000	7,642	0,000
	34	out/98	7,45	7,70	0,061	7,748	0,089	7,541	0,008	7,519	0,005
	35	nov/98	7,05	7,50	0,202	7,569	0,270	7,360	0,096	7,335	0,081
	36	dez/98	6,33	7,14	0,656	7,258	0,861	7,021	0,478	6,984	0,427
1999	37	jan/99	7,73	6,49	1,533	6,701	1,059	6,409	1,746	6,352	1,898
	38	fev/99	7,51	7,48	0,001	7,318	0,037	7,478	0,001	7,580	0,005
	39	mar/99	8,16	7,50	0,430	7,433	0,528	7,414	0,556	7,387	0,597
	40	abr/99	8,03	8,03	0,000	7,869	0,026	7,920	0,012	7,957	0,005
	41	mai/99	7,70	8,03	0,109	7,966	0,071	7,868	0,028	7,843	0,021
	42	jun/99	7,84	7,77	0,005	7,806	0,001	7,592	0,062	7,554	0,082
	43	jul/99	7,54	7,83	0,081	7,827	0,082	7,678	0,019	7,677	0,019
	44	ago/99	7,70	7,60	0,011	7,655	0,002	7,447	0,064	7,414	0,082
	45	set/99	7,40	7,68	0,078	7,682	0,079	7,552	0,023	7,554	0,024
	46	out/99	7,50	7,46	0,002	7,513	0,000	7,323	0,031	7,291	0,044
Previsão		nov/99	7,49		7,505		7,379		7,378		
EQM				0,922		0,950		0,334		0,336	

EQM: Erro quadrático Médio



Uma maneira alternativa para verificar qual a ordem do processo autoregressivo é por meio da função de autocorrelação parcial da variável em questão.

$$FAP(Y_t, Y_{t-2}) = \text{cor}(Y_t, Y_{t-2}) \text{ controlando o efeito de } Y_{t-1} \text{ sobre } Y_t.$$

A ordem do modelo será definido pelo coeficiente de autocorrelação parcial de maior ordem que é estatisticamente diferente de zero. Tal verificação pode ser realizada pelo gráfico das autocorrelações chamado de correlograma, que pode ser visto na última página. Pela função de autocorrelação parcial, o modelo será de ordem um. A correlação parcial

## Capítulo 4. Sobre as previsões

### 4.1 Considerações sobre os Métodos de Previsão de Séries Temporais

Makridakis e Wheelwright, investigaram o poder preditivo de vários métodos comumente utilizados na previsão de séries temporais. Pela comparação dos resultados alcançados, identificaram situações onde esses métodos apresentam melhor desempenho e definiram alguns critérios para a seleção dos procedimentos de previsão, pela confrontação dos objetivos a atingir.

O estudo desenvolvido por tais pesquisadores constatou que o incremento da complexidade e da sofisticação estatística dos métodos de previsão de séries temporais não implica, necessariamente, em uma melhora na acuidade da previsão. *"Métodos simples de previsão podem apresentar desempenho extremamente satisfatório sob certas condições"* . Além disso, métodos de previsão menos complexos normalmente permitem alcançar total compreensão de suas suposições e limitações, e de interpretação de seus resultados. Assim, antes de se adotar um método de previsão mais complexo, é necessário avaliar os benefícios que um método dessa natureza pode gerar em relação ao custo de sua aplicação.

Segundo, como a acuidade de uma previsão fica determinada não apenas pelo horizonte de previsão especificado, mas também pelas características das observações da série temporal investigada, a otimização desse critério pode ser alcançada com a aplicação de mais de um método de previsão. A combinação de previsões ou a verificação da consistência desses valores permite aumentar a confiabilidade da previsão e reduzir a possibilidade de grandes desvios.

Mentzer e Cox também analisaram a frequência de aplicação de diferentes métodos de previsão de séries temporais considerando como critérios o horizonte de previsão, familiaridade com os procedimentos preditivos e nível de satisfação decorrente do emprego desses procedimentos. Em relação ao horizonte de previsão (tabela 1), o Alisamento Exponencial Simples e a Média Móvel são mais frequentemente usados para períodos curtos, menos para médio prazo e menos ainda para longo prazo, comprovando evidências empíricas que indicam melhor desempenho destes métodos para horizonte de previsão menores .

Independente do horizonte de previsão adotado, pouca aplicabilidade do método de Box e Jenkins pode ser observada.

Tabela 1. Frequência relativa de uso de técnicas de previsão em relação ao horizonte de previsão (em %).

Método	Horizonte de Previsão		
	Curto	Médio	Longo
Alisamento Exponencial Simples	24	17	6
Média Móvel	24	22	5
Decomposição	9	13	5
Box e Jenkins	5	6	2

Fonte : [MENTZER1984].

A tabela 2 sintetiza os estudos de Mentzer quanto à familiaridade do grupo entrevistado com diferentes métodos de previsão. Analisando os resultados dessa pesquisa, fica evidente que métodos mais simples de previsão são mais conhecidos que métodos avançados de séries temporais.

Tabela 2. Frequência relativa dos graus de familiaridade com os diferentes métodos de previsão (em %).

Método	Familiaridade		
	Muito familiar	Vagamente familiar	Completamente não familiar
Média Móvel	85	7	8
Alisamento Exponencial Simples	73	12	15
Decomposição	42	9	49
Box e Jenkins	26	9	65

Fonte : [MENTZER1984].

A satisfação ao se adotar algum desses métodos de previsão também foi estudada e os resultados descritos na tabela 3.

Tabela 3. Frequência relativa dos graus de satisfação com os diferentes métodos de previsão (em %).

Método	Satisfação		
	Satisfatório	Indiferente	Insatisfatório
Alisamento Exponencial Simples	60	19	21
Média Móvel	58	21	21
Decomposição Clássica	55	14	31
Box e Jenkins	30	13	57

Fonte : [MENTZER1984].

Analisando a tabela 3, pode ser observado que o Alisamento Exponencial Simples foi considerado o mais satisfatório entre os métodos de previsão avaliados. Esse resultado é consistente com estudos empíricos, que constataram ser esse método capaz de alcançar considerável acuidade, além de ser de fácil compreensão e utilização. O método de Médias Móveis também apresentou alto nível de satisfação, enquanto que o método de Box e Jenkins ficou caracterizado como o de menor satisfação e maior insatisfação.



Tabela 1. Produção de Leite no estado de São Paulo (em Milhões de Litros)

ANO	MESES											
	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1975												152,79
1976	155,48	141,92	137,73	132,77	126,00	120,27	116,59	120,72	125,37	132,93	142,9	144,20
1977	145,28	135,33	134,19	130,65	124,89	119,98	117,65	116,99	125,19	135,92	152,42	158,33
1978	159,66	148,53	147,67	138,87	132,48	129,02	133,07	134,02	131,47	138,53	152,57	159,60
1979	158,85	146,35	143,06	136,60	131,63	128,05	121,26	123,18	137,34	141,57	151,22	149,28
1980	149,76	145,27	142,80	132,88	129,91	127,50	134,06	135,97	138,43	144,82	151,56	162,45
1981	163,94	151,24	149,30	155,80	153,81	148,24	136,77	134,10	142,02	173,60	195,40	190,50
1982	186,90	171,20	172,70	164,60								

**Saída de Regressão Linear** (considero-se para análise o período de dez/75 a dez/81)

## RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0,449725
R-Quadrado	0,202252
R-quadrado ajustado	0,191016
Erro padrão	13,75186
Observações	73

## ANOVA

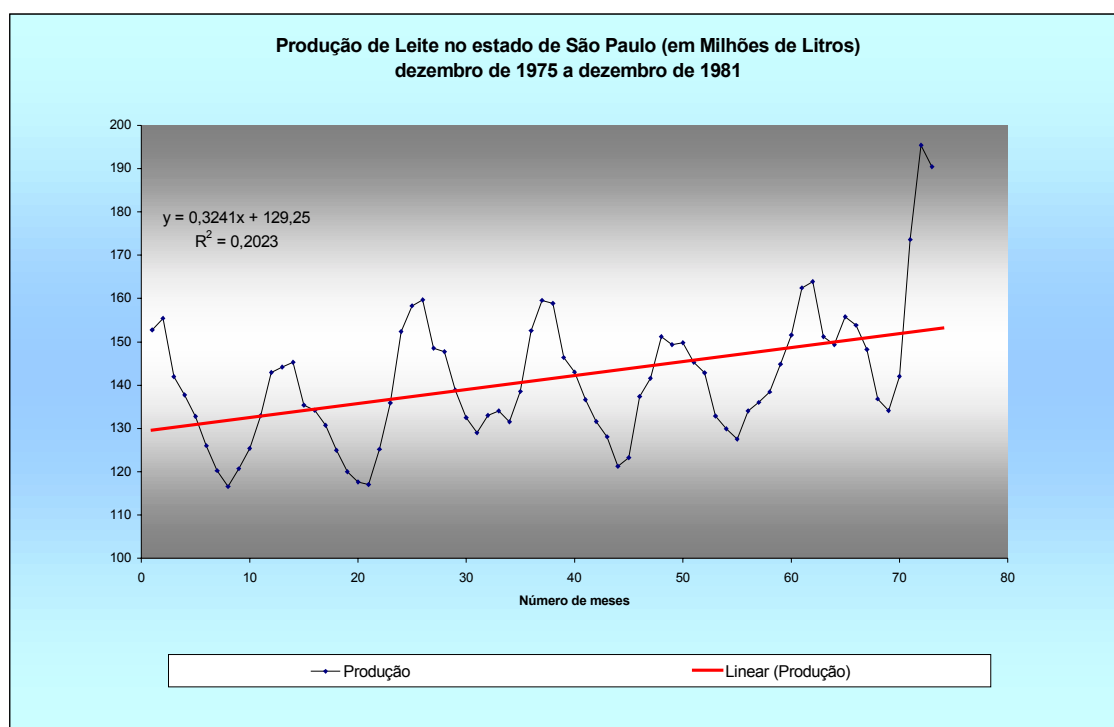
	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	3404,155	3404,155	18,00057	6,56E-05
Resíduo	71	13427,08	189,1137		
Total	72	16831,23			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	129,2488	3,252427	39,73918	3,03E-50	122,7636	135,7339	122,7636	135,7339
Variável X 1	0,32408	0,076385	4,242708	6,56E-05	0,171772	0,476387	0,171772	0,476387

## Previsões

Tabela 1.2

Mês/ano	Produção Real	Previsão	Erro
jan/82	186,9	153,23	33,67
fev/82	171,2	153,55	17,65
mar/82	172,7	153,88	18,82
abr/82	164,4	154,20	10,20



**Saída de Regressão Quadrática** (considero-se para análise o período de dez/75 a dez/81)

## RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0,52728
R-Quadrado	0,27802
R-quadrado ajustado	0,25740
Erro padrão	13,1756
Observações	73

## ANOVA

	<i>Gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	2	4679,486	2339,743	13,47806	1,12E-05
Resíduo	70	12151,75	173,5964		
Total	72	16831,23			

	<i>Coefficient es</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	138,9871	4,755967	29,22374	5,94E-41	129,5017	148,4726	129,5017	148,4726
Variável X 1	-0,45499	0,296602	-1,534	0,129536	-1,04654	0,136565	-1,04654	0,136565
Variável X 2	0,010528	0,003884	2,710448	0,008444	0,002781	0,018275	0,002781	0,018275

**Previsões**

Tabela 1.3

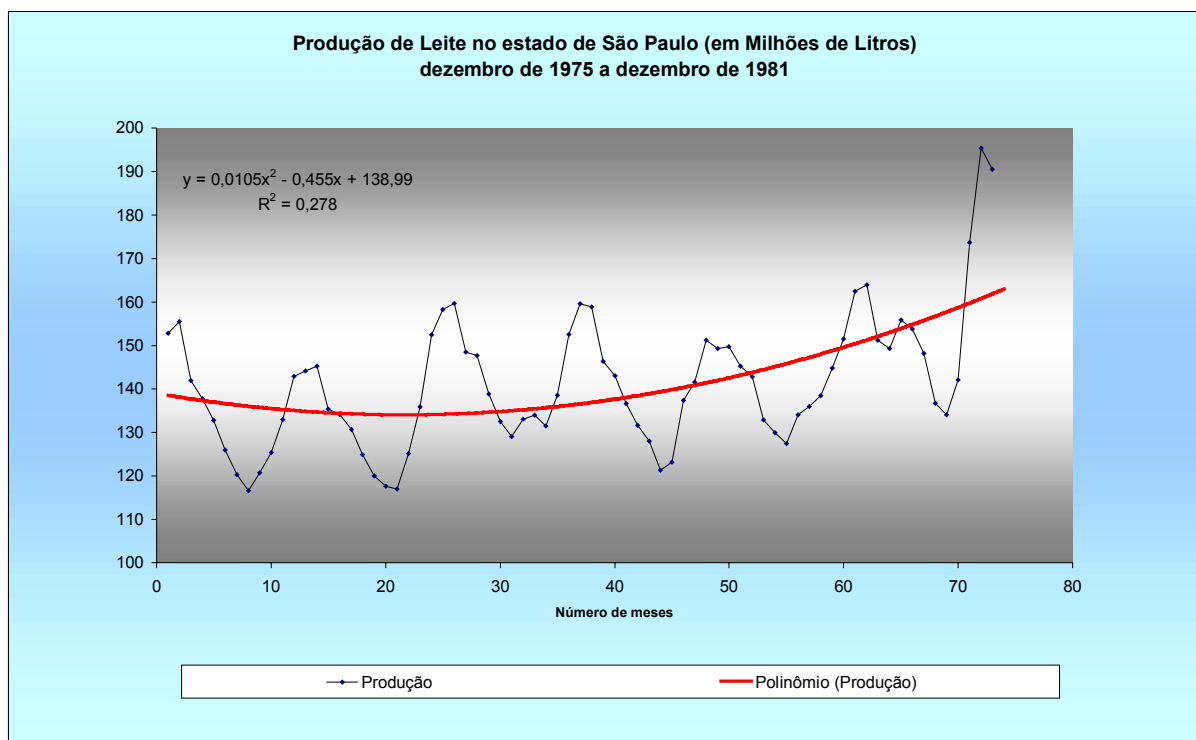
Mês/ano	Produção Real	Previsão	Erro
jan/82	186,9	162,82	24,08
fev/82	171,2	163,93	7,27
mar/82	172,7	165,06	7,64
abr/82	164,4	166,21	-1,81

Tabela 1.4 Ajuste de mínimos quadrados

Mês/Ano	Produção	Ajuste Linear	Ajuste Quadrático
Dez/75	152,79	129,57	138,55
jan/76	155,48	129,90	138,12
fev/76	141,92	130,22	137,72
mar/76	137,73	130,55	137,34
abr/76	132,77	130,87	136,98
mai/76	126,00	131,19	136,64
jun/76	120,27	131,52	136,32
jul/76	116,59	131,84	136,02
ago/76	120,72	132,17	135,75
set/76	125,37	132,49	135,49
out/76	132,93	132,81	135,26
nov/76	142,90	133,14	135,04
dez/76	144,20	133,46	134,85
jan/77	145,28	133,79	134,68
fev/77	135,33	134,11	134,53
mar/77	134,19	134,43	134,40
abr/77	130,65	134,76	134,29
mai/77	124,89	135,08	134,20
jun/77	119,98	135,41	134,14
jul/77	117,65	135,73	134,09
ago/77	116,99	136,05	134,07
set/77	125,19	136,38	134,06
out/77	135,92	136,70	134,08
nov/77	152,42	137,03	134,12
dez/77	158,33	137,35	134,18
jan/78	159,66	137,67	134,26

fev/78	148,53	138,00	134,36
mar/78	147,67	138,32	134,48
abr/78	138,87	138,65	134,63
mai/78	132,48	138,97	134,79
jun/78	129,02	139,30	134,98
jul/78	133,07	139,62	135,18
ago/78	134,02	139,94	135,41
set/78	131,47	140,27	135,66
out/78	138,53	140,59	135,93
nov/78	152,57	140,92	136,22
dez/78	159,60	141,24	136,53
jan/79	158,85	141,56	136,86
fev/79	146,35	141,89	137,22
mar/79	143,06	142,21	137,59
abr/79	136,60	142,54	137,99
mai/79	131,63	142,86	138,40
jun/79	128,05	143,18	138,84
jul/79	121,26	143,51	139,30
ago/79	123,18	143,83	139,78
set/79	137,34	144,16	140,28
out/79	141,57	144,48	140,80
nov/79	151,22	144,80	141,34
dez/79	149,28	145,13	141,91
jan/80	149,76	145,45	142,49
fev/80	145,27	145,78	143,10
mar/80	142,80	146,10	143,72
abr/80	132,88	146,43	144,37
mai/80	129,91	146,75	145,04
jun/80	127,50	147,07	145,73
jul/80	134,06	147,40	146,44

ago/80	135,97	147,72	147,17
set/80	138,43	148,05	147,92
out/80	144,82	148,37	148,70
nov/80	151,56	148,69	149,49
dez/80	162,45	149,02	150,31
jan/81	163,94	149,34	151,14
fev/81	151,24	149,67	152,00
mar/81	149,30	149,99	152,88
abr/81	155,80	150,31	153,78
mai/81	153,81	150,64	154,70
jun/81	148,24	150,96	155,64
jul/81	136,77	151,29	156,60
ago/81	134,10	151,61	157,59
set/81	142,02	151,93	158,59
out/81	173,60	152,26	159,62
nov/81	195,40	152,58	160,66
dez/81	190,50	152,91	161,73



**Modelos com Tendência + Sazonalidade**

Tabela 1.5. Percentuais da produção em relação à média anual.

Produção de Leite no estado de São Paulo													
MESES %													Médias
ANO	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	Anuais
1975												100	152,79
1976	1,1684	1,0665	1,0350	0,9977	0,9468	0,9038	0,8761	0,9072	0,9421	0,9989	1,0738	1,0836	133,07
1977	1,0918	1,0170	1,0084	0,9818	0,9385	0,9016	0,8841	0,8792	0,9408	1,0214	1,1454	1,1898	133,07
1978	1,1234	1,0451	1,0390	0,9771	0,9321	0,9078	0,9363	0,9430	0,9250	0,9747	1,0735	1,1230	142,12
1979	1,1425	1,0526	1,0290	0,9825	0,9468	0,9210	0,8722	0,8860	0,9878	1,0183	1,0877	1,0737	139,03
1980	1,0600	1,0282	1,0107	0,9405	0,9195	0,9024	0,9489	0,9624	0,9798	1,0250	1,0727	1,1498	141,28
1981	1,0383	0,9579	0,9456	0,9867	0,9741	0,9389	0,8662	0,8493	0,8995	1,0995	1,2375	1,2065	157,89
1982	1,0751	0,9848	0,9934	0,9468									173,85
Médias %	1,0999	1,0217	1,0087	0,9733	0,9430	0,9126	0,8973	0,9045	0,9458	1,0230	1,1151	1,1377	

Com estes valores aprecia-se a produção em % para cada mês.

O seguinte modelo pode ser analisado

$$Y_t = \hat{Y}_t(\text{Linear}) + \text{Sazonalidade} + \text{erro aleatorio}$$

onde:

$\hat{Y}_t(\text{Linear})$  : Corresponde a previsão da regressão linear

Sazonalidade : Corresponde a multiplicação do valor da previsão linear com a % a mais por cada mês.

Tabela 1.6 Ajuste do modelo de mínimos quadrados( linear) mais sazonalidade

mês/ano	Produção	T(Linear)+S	mês/ano	Produção	T(Linear)+S
Dez/75	152,79	151,39	Jan/79	158,85	161,74
jan/76	155,48	151,77	Fev/79	146,35	149,36
fev/76	141,92	138,88	mar/79	143,06	146,33
mar/76	137,73	135,11	abr/79	136,60	140,04
abr/76	132,77	130,57	Mai/79	131,63	135,25
mai/76	126,00	124,22	Jun/79	128,05	131,87
jun/76	120,27	118,86	jul/79	121,26	125,16
jul/76	116,59	115,51	ago/79	123,18	127,43
ago/76	120,72	119,90	set/79	137,34	142,40
set/76	125,37	124,82	out/79	141,57	147,12
out/76	132,93	132,67	nov/79	151,22	157,50
nov/76	142,90	142,97	dez/79	149,28	155,83
dez/76	144,20	144,62	jan/80	149,76	154,18
jan/77	145,28	146,06	fev/80	145,27	149,89
fev/77	135,33	136,39	mar/80	142,80	147,67
mar/77	134,19	135,57	abr/80	132,88	137,72
abr/77	130,65	132,31	mai/80	129,91	134,93
mai/77	124,89	126,78	jun/80	127,50	132,72
jun/77	119,98	122,09	jul/80	134,06	139,86
jul/77	117,65	120,00	ago/80	135,97	142,17
ago/77	116,99	119,62	set/80	138,43	145,05
set/77	125,19	128,30	out/80	144,82	152,08
out/77	135,92	139,63	nov/80	151,56	159,51
nov/77	152,42	156,95	dez/80	162,45	171,34
dez/77	158,33	163,43	jan/81	163,94	155,06
jan/78	159,66	154,66	fev/81	151,24	143,36
fev/78	148,53	144,22	mar/81	149,30	141,83

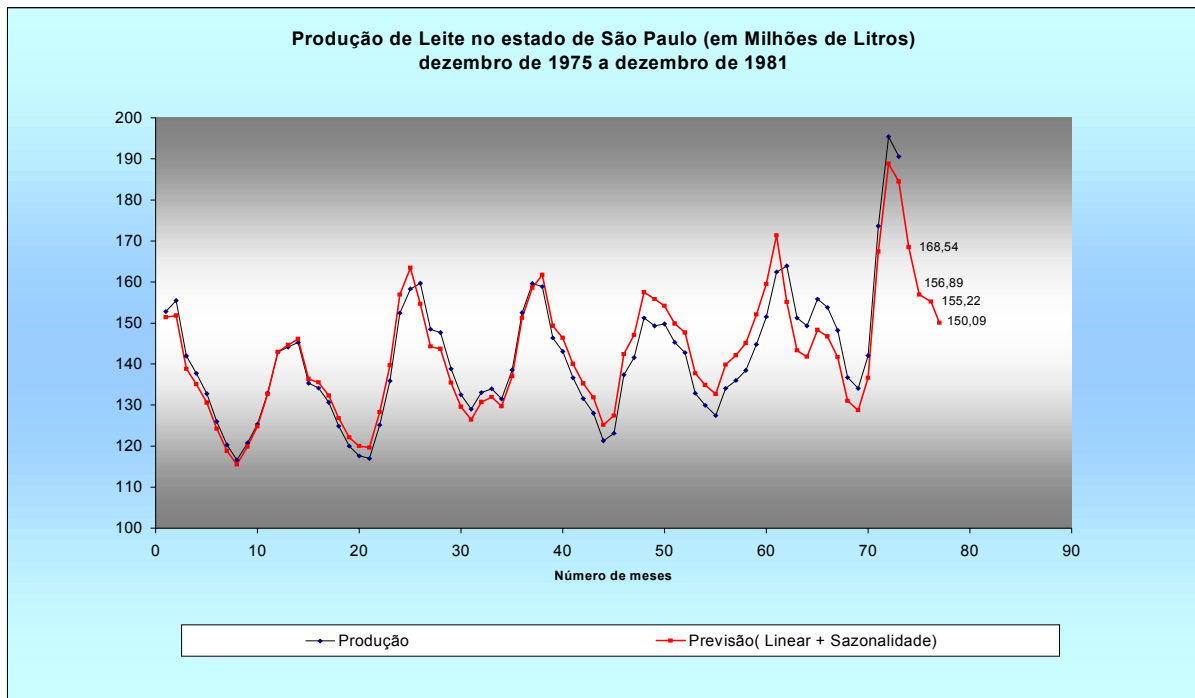


mar/78	147,67	143,72	abr/81	155,80	148,32
abr/78	138,87	135,47	mai/81	153,81	146,74
mai/78	132,48	129,54	jun/81	148,24	141,73
jun/78	129,02	126,45	jul/81	136,77	131,05
jul/78	133,07	130,72	ago/81	134,10	128,76
ago/78	134,02	131,96	set/81	142,02	136,66
set/78	131,47	129,75	out/81	173,60	167,40
out/78	138,53	137,04	nov/81	195,40	188,83
nov/78	152,57	151,27	dez/81	190,50	184,48
dez/78	159,60	158,61			

## Previsões

Tabela 1.7

Mês/ano	Produção Real	Previsão	Erro
jan/82	186,9	168,54	18,36
fev/82	171,2	156,89	14,31
mar/82	172,7	155,22	17,48
abr/82	164,4	150,09	14,31



$$Y_t = \hat{Y}_t(\text{Quadrático}) + \text{Sazonalidade} + \text{erro aleatório}$$

onde:

$\hat{Y}_t(\text{Quadrático})$ : Corresponde a previsão da regressão quadrática.

Sazonalidade: Corresponde a multiplicação do valor da previsão quadrática com a % a mais por cada mês.

Tabela 1.8 Ajuste do modelo de mínimos quadrados( linear) mais sazonalidade

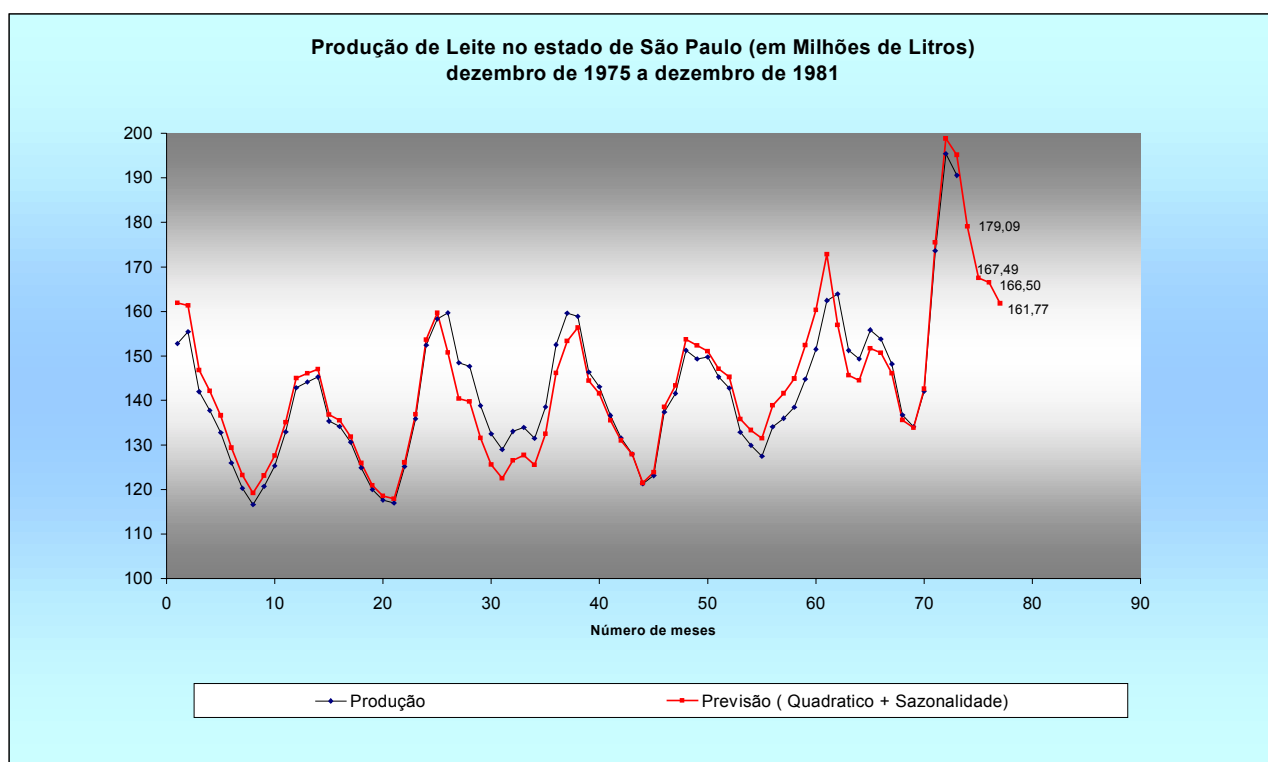
mês/ano	Produção	T(quadrático)+S	mês/ano	Produção	T(quadrático)+S
dez/75	152,79	161,87	jan/79	158,85	156,37
jan/76	155,48	161,38	fev/79	146,35	144,44
fev/76	141,92	146,88	mar/79	143,06	141,58
mar/76	137,73	142,14	abr/79	136,60	135,57
abr/76	132,77	136,67	mai/79	131,63	131,03
mai/76	126,00	129,38	jun/79	128,05	127,87
jun/76	120,27	123,20	jul/79	121,26	121,49

jul/76	116,59	119,17	ago/79	123,18	123,84
ago/76	120,72	123,14	set/79	137,34	138,57
set/76	125,37	127,65	out/79	141,57	143,37
out/76	132,93	135,11	nov/79	151,22	153,73
nov/76	142,90	145,01	dez/79	149,28	152,36
dez/76	144,20	146,12	jan/80	149,76	151,04
jan/77	145,28	147,04	fev/80	145,27	147,13
fev/77	135,33	136,81	mar/80	142,80	145,26
mar/77	134,19	135,53	abr/80	132,88	135,78
abr/77	130,65	131,85	mai/80	129,91	133,36
mai/77	124,89	125,95	jun/80	127,50	131,51
jun/77	119,98	120,94	jul/80	134,06	138,95
jul/77	117,65	118,55	ago/80	135,97	141,63
ago/77	116,99	117,87	set/80	138,43	144,93
set/77	125,19	126,12	out/80	144,82	152,42
out/77	135,92	136,95	nov/80	151,56	160,36
nov/77	152,42	153,62	dez/80	162,45	172,82
dez/77	158,33	159,65	jan/81	163,94	156,93
jan/78	159,66	150,82	fev/81	151,24	145,59
fev/78	148,53	140,42	mar/81	149,30	144,56
mar/78	147,67	139,73	abr/81	155,80	151,74
abr/78	138,87	131,54	mai/81	153,81	150,70
mai/78	132,48	125,64	jun/81	148,24	146,12
jun/78	129,02	122,53	jul/81	136,77	135,65
jul/78	133,07	126,57	ago/81	134,10	133,84
ago/78	134,02	127,69	set/81	142,02	142,65
set/78	131,47	125,49	out/81	173,60	175,49
out/78	138,53	132,49	nov/81	195,40	198,83
nov/78	152,57	146,23	dez/81	190,50	195,13
dez/78	159,60	153,32			

**Previsões**

Tabela 1.9

Mês/ano	Produção Real	Previsão	Erro
jan/82	186,9	179,09	7,81
fev/82	171,2	167,49	3,71
mar/82	172,7	166,50	6,20
abr/82	164,4	161,77	2,63



**Modelo sobre uma abordagem Arima**

Tabela 1.10 Ajuste da produção de Leite em São Paulo a partir de um modelo ARIMA.

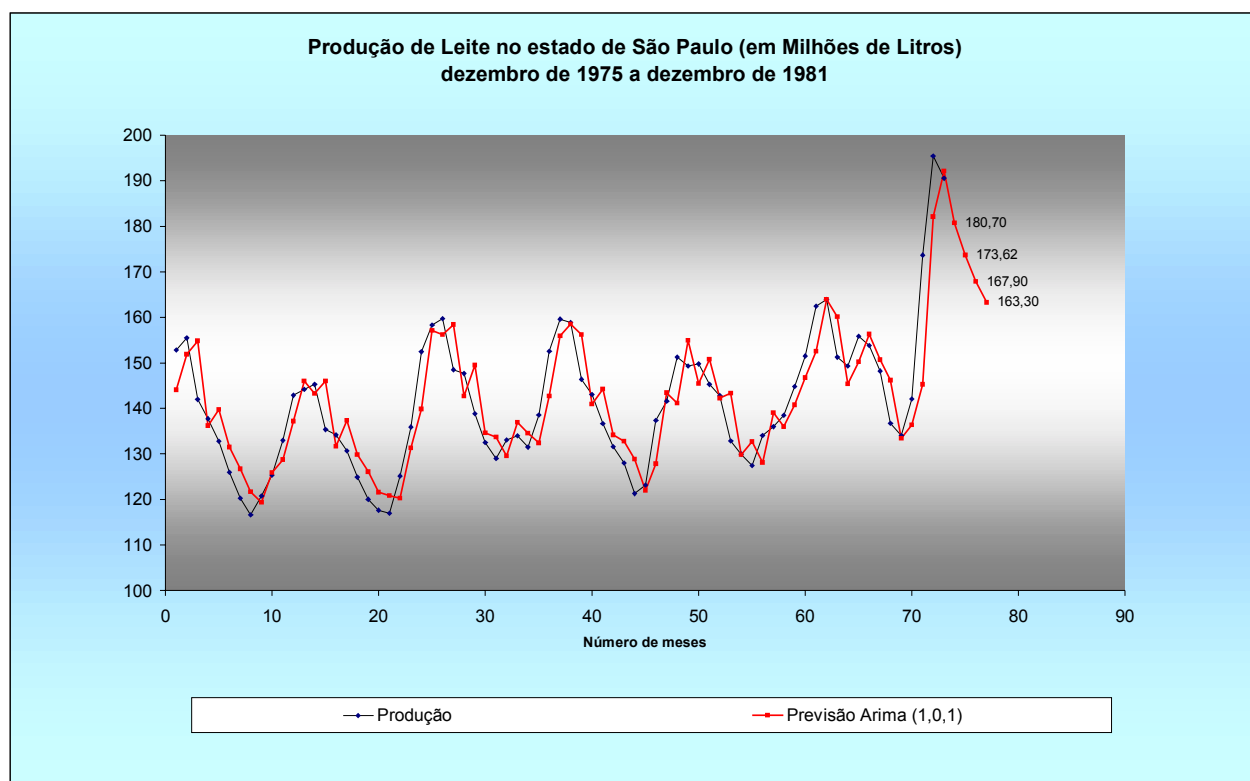
mês/ano	Produção	Arima(1,0,1)	mês/ano	Produção	Arima(1,0,1)
dez/75	152,79	144,10	Jan/79	158,85	158,46
jan/76	155,48	151,85	Fev/79	146,35	156,19
fev/76	141,92	154,78	Mar/79	143,06	140,98
mar/76	137,73	136,16	Abr/79	136,60	144,30
abr/76	132,77	139,74	Mai/79	131,63	134,19
mai/76	126,00	131,48	Jun/79	128,05	132,76
jun/76	120,27	126,76	Jul/79	121,26	128,80
jul/76	116,59	121,63	Ago/79	123,18	121,90
ago/76	120,72	119,39	Set/79	137,34	127,87
set/76	125,37	125,91	Out/79	141,57	143,39
out/76	132,93	128,72	Nov/79	151,22	141,15
nov/76	142,90	137,20	dez/79	149,28	154,89
dez/76	144,20	145,99	jan/80	149,76	145,47
jan/77	145,28	143,28	fev/80	145,27	150,82
fev/77	135,33	146,05	mar/80	142,80	142,26
mar/77	134,19	131,65	abr/80	132,88	143,32
abr/77	130,65	137,38	mai/80	129,91	129,82
mai/77	124,89	129,88	jun/80	127,50	132,70
jun/77	119,98	126,11	jul/80	134,06	128,11
jul/77	117,65	121,58	ago/80	135,97	138,99
ago/77	116,99	120,80	set/80	138,43	136,03
set/77	125,19	120,33	out/80	144,82	140,73
out/77	135,92	131,29	nov/80	151,56	146,73
nov/77	152,42	139,82	dez/80	162,45	152,54
dez/77	158,33	157,12	jan/81	163,94	163,87
jan/78	159,66	156,18	fev/81	151,24	160,14

fev/78	148,53	158,39	mar/81	149,30	145,40
mar/78	147,67	142,73	abr/81	155,80	150,25
abr/78	138,87	149,45	mai/81	153,81	156,32
mai/78	132,48	134,58	jun/81	148,24	150,67
jun/78	129,02	133,68	jul/81	136,77	146,22
jul/78	133,07	129,60	ago/81	134,10	133,45
ago/78	134,02	136,94	set/81	142,02	136,36
set/78	131,47	134,51	out/81	173,60	145,26
out/78	138,53	132,39	nov/81	195,40	182,09
nov/78	152,57	142,68	dez/81	190,50	192,14
dez/78	159,60	155,88			

## Previsões

Tabela 1.9

Mês/ano	Produção Real	Previsão	Erro
jan/82	186,9	180,70	6,20
fev/82	171,2	173,62	-2,42
mar/82	172,7	167,90	4,80
abr/82	164,4	163,30	1,10



**Saída SPSS do modelo ARIMA**

MODEL: MOD\_1.

Autocorrelations: P.LEITE

Auto- Stand.

Lag	Corr.	Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
1	,784	,115				.	↔****.*****					46,717	,000
2	,415	,114				.	↔****.***					60,004	,000
3	,096	,113				.	↔**	.				60,717	,000
4	-,075	,112				.	**↔	.				61,164	,000
5	-,161	,111				.	***↔	.				63,238	,000
6	-,186	,111				.	****↔	.				66,059	,000
7	-,164	,110				.	***↔	.				68,303	,000
8	-,057	,109				.	*↔	.				68,576	,000
9	,120	,108				.	↔**	.				69,800	,000
10	,336	,107				.	↔***.***					79,600	,000
11	,463	,106				.	↔***.*****					98,560	,000
12	,455	,106				.	↔***.*****					117,145	,000
13	,341	,105				.	↔***.***					127,743	,000
14	,175	,104				.	↔****					130,591	,000
15	-,027	,103				.	*↔	.				130,660	,000
16	-,223	,102				.	****↔	.				135,458	,000

Plot Symbols: Autocorrelations \* Two Standard Error Limits .

Total cases: 73 Computable first lags: 72

Partial Autocorrelations: P.LEITE

Pr-Aut- Stand.

Lag	Corr.	Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	
												□□□□□↑□□□□□↑□□□□□↑□□□□□↑□□□□□↑□□□□□↑□□□□□↑□□□□□↑□□□□□↑□□□□□
1	,784	,117				.	↔****.*****					
2	-,517	,117				.	****.***↔	.				
3	,025	,117				.	*	.				
4	,062	,117				.	↔*	.				
5	-,176	,117				.	****↔	.				



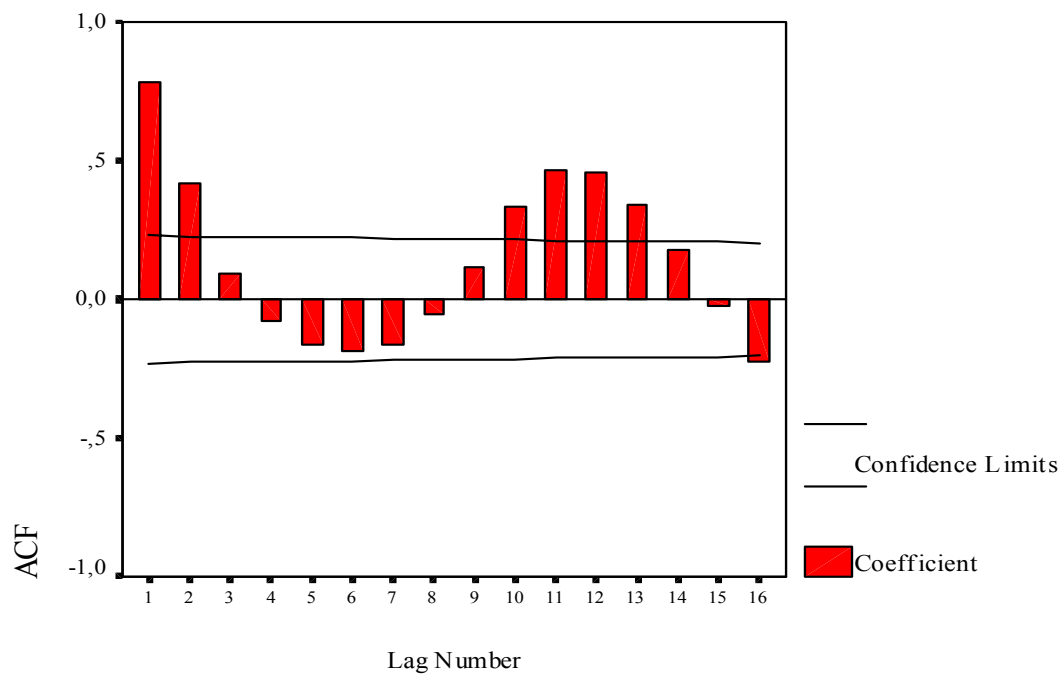
6	,027	,117	.	↔*	.
7	,016	,117	.	*	.
8	,175	,117	.	↔****	.
9	,161	,117	.	↔***	.
10	,244	,117	.	↔*****	.
11	-,018	,117	.	*	.
12	,029	,117	.	↔*	.
13	,052	,117	.	↔*	.
14	-,093	,117	.	**↔	.
15	-,163	,117	.	***↔	.
16	-,112	,117	.	**↔	.

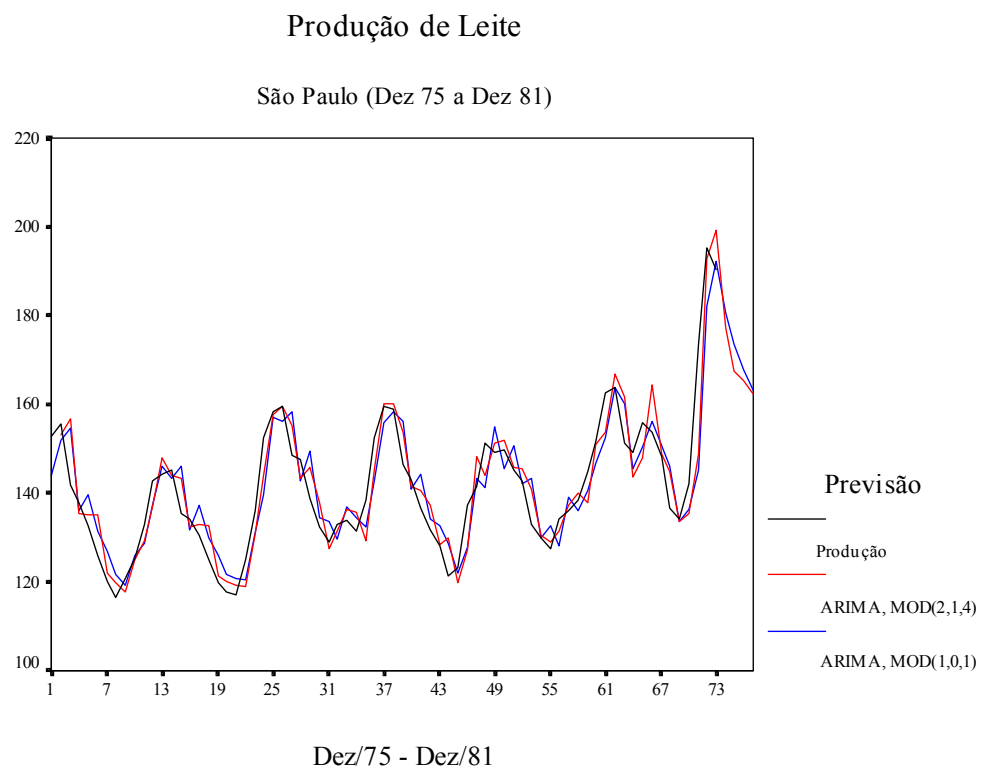
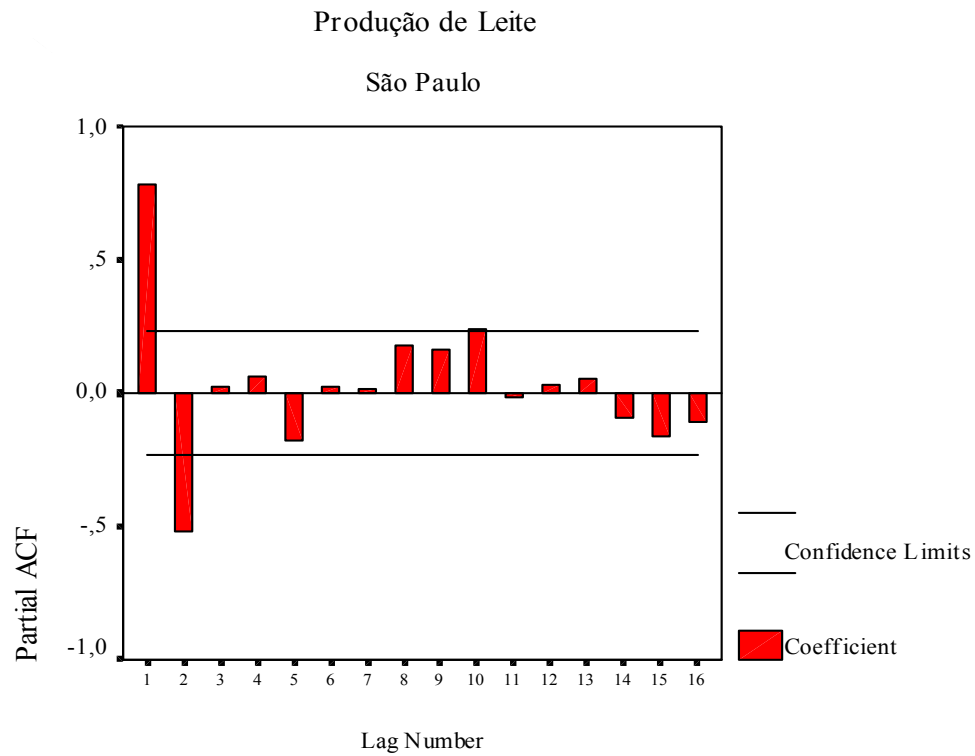
Plot Symbols:      Autocorrelations \*      Two Standard Error Limits .

Total cases: 73      Computable first lags: 72

## Produção de Leite

São Paulo





MODEL: MOD\_(1,0,1)

Model Description:

Variable: P.LEITE

Regressors: NONE

Non-seasonal differencing: 0

No seasonal component in model.

Parameters:

AR1 \_\_\_\_\_ < value originating from estimation >

MA1 \_\_\_\_\_ < value originating from estimation >

CONSTANT \_\_\_\_\_ < value originating from estimation >

95,00 percent confidence intervals will be generated.

Split group number: 1 Series length: 73

Number of cases skipped at end because of missing values: 4

Melard's algorithm will be used for estimation.

Termination criteria:

Parameter epsilon: ,001

Maximum Marquardt constant: 1,00E+09

SSQ Percentage: ,001

Maximum number of iterations: 10

Initial values:

AR1 ,52958

MA1 -,56932

CONSTANT 142,2340

Marquardt constant = ,001

Adjusted sum of squares = 4018,454

—

Iteration History:

Iteration	Adj. Sum of Squares	Marquardt Constant
-----------	---------------------	--------------------

1	3539,9321	,0010000
2	3496,6459	1,0000000
3	3490,8398	,1000000
4	3490,6346	,0100000

Conclusion of estimation phase.

Estimation terminated at iteration number 5 because:

Sum of squares decreased by less than ,001 percent.

#### FINAL PARAMETERS:

Number of residuals	73
Standard error	6,9645946
Log likelihood	-244,74526
AIC	495,49053
SBC	502,36191

#### Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	70	3490,6047	48,505578

#### Variables in the Model:

B	SEB	T-RATIO	APPROX. PROB.
---	-----	---------	---------------

AR1	,80649	,0853397	9,450330	,00000000
MA1	-,50105	,1133159	-4,421725	,00003505
CONSTANT	144,09694	5,9658241	24,153736	,00000000

Covariance Matrix:

	AR1	MA1
AR1	,00728287	,00408882
MA1	,00408882	,01284050

Correlation Matrix:

	AR1	MA1
AR1	1,0000000	,4228206
MA1	,4228206	1,0000000

Regressor Covariance Matrix:

	CONSTANT
CONSTANT	35,591058

Regressor Correlation Matrix:

	CONSTANT
--	----------

CONSTANT 1,0000000

The following new variables are being created:

Name	Label
FIT_1	Fit for P.LEITE from ARIMA, MOD_1 CON
ERR_1	Error for P.LEITE from ARIMA, MOD_1 CON
LCL_1	95% LCL for P.LEITE from ARIMA, MOD_1 CON
UCL_1	95% UCL for P.LEITE from ARIMA, MOD_1 CON
SEP_1	SE of fit for P.LEITE from ARIMA, MOD_1 Com

Tabela. Ajuste e Previsão para a produção a partir de dois modelos Arima. ARIMA(2,1,4) com transformação logarítmica, e ARIMA(1,0,1)

Mês/Ano	Produção	FIT_(2,1,4)	FIT_1	ERR_1	LCL_1	UCL_1	SEP_1
dez/75	152,79	-----	144,10	8,69	108,35	179,85	17,93
jan/76	155,48	153,13	151,85	3,63	136,52	167,17	7,68
fev/76	141,92	156,88	154,78	-12,86	140,49	169,08	7,17
mar/76	137,73	135,38	136,16	1,57	122,13	150,20	7,04
abr/76	132,77	135,27	139,74	-6,97	125,74	153,74	7,02
mai/76	126	135,25	131,48	-5,48	117,50	145,46	7,01
jun/76	120,27	121,94	126,76	-6,49	112,78	140,74	7,01
jul/76	116,59	119,77	121,63	-5,04	107,66	135,60	7,01
ago/76	120,72	117,85	119,39	1,33	105,41	133,36	7,01
set/76	125,37	125,19	125,91	-0,54	111,94	139,89	7,01
out/76	132,93	129,43	128,72	4,21	114,75	142,70	7,01
nov/76	142,9	137,03	137,20	5,70	123,22	151,17	7,01
dez/76	144,2	148,04	145,99	-1,79	132,01	159,96	7,01
jan/77	145,28	143,89	143,28	2,00	129,31	157,26	7,01

fev/77	135,33	143,25	146,05	-10,72	132,08	160,03	7,01
mar/77	134,19	132,28	131,65	2,54	117,68	145,63	7,01
abr/77	130,65	132,99	137,38	-6,73	123,40	151,35	7,01
mai/77	124,89	132,65	129,88	-4,99	115,91	143,86	7,01
jun/77	119,98	121,43	126,11	-6,13	112,13	140,08	7,01
jul/77	117,65	120,12	121,58	-3,93	107,60	135,55	7,01
ago/77	116,99	119,31	120,80	-3,81	106,83	134,77	7,01
set/77	125,19	119,08	120,33	4,86	106,35	134,30	7,01
out/77	135,92	130,91	131,29	4,63	117,31	145,26	7,01
nov/77	152,42	144,04	139,82	12,60	125,85	153,80	7,01
dez/77	158,33	157,57	157,12	1,21	143,15	171,10	7,01
jan/78	159,66	159,49	156,18	3,48	142,21	170,16	7,01
fev/78	148,53	155,15	158,39	-9,86	144,42	172,37	7,01
mar/78	147,67	143,38	142,73	4,94	128,76	156,71	7,01
abr/78	138,87	145,81	149,45	-10,58	135,48	163,43	7,01
mai/78	132,48	137,73	134,58	-2,10	120,60	148,55	7,01
jun/78	129,02	127,47	133,68	-4,66	119,70	147,65	7,01
jul/78	133,07	131,53	129,60	3,47	115,63	143,58	7,01
ago/78	134,02	136,38	136,94	-2,92	122,97	150,92	7,01
set/78	131,47	135,88	134,51	-3,04	120,53	148,48	7,01
out/78	138,53	129,19	132,39	6,14	118,42	146,37	7,01
nov/78	152,57	145,11	142,68	9,89	128,71	156,66	7,01
dez/78	159,6	160,27	155,88	3,72	141,91	169,86	7,01
jan/79	158,85	160,01	158,46	0,39	144,49	172,44	7,01
fev/79	146,35	153,70	156,19	-9,84	142,21	170,16	7,01
mar/79	143,06	141,50	140,98	2,08	127,01	154,96	7,01
abr/79	136,6	140,76	144,30	-7,70	130,33	158,28	7,01
mai/79	131,63	137,35	134,19	-2,56	120,22	148,17	7,01
jun/79	128,05	128,31	132,76	-4,71	118,78	146,73	7,01
jul/79	121,26	129,98	128,80	-7,54	114,82	142,77	7,01

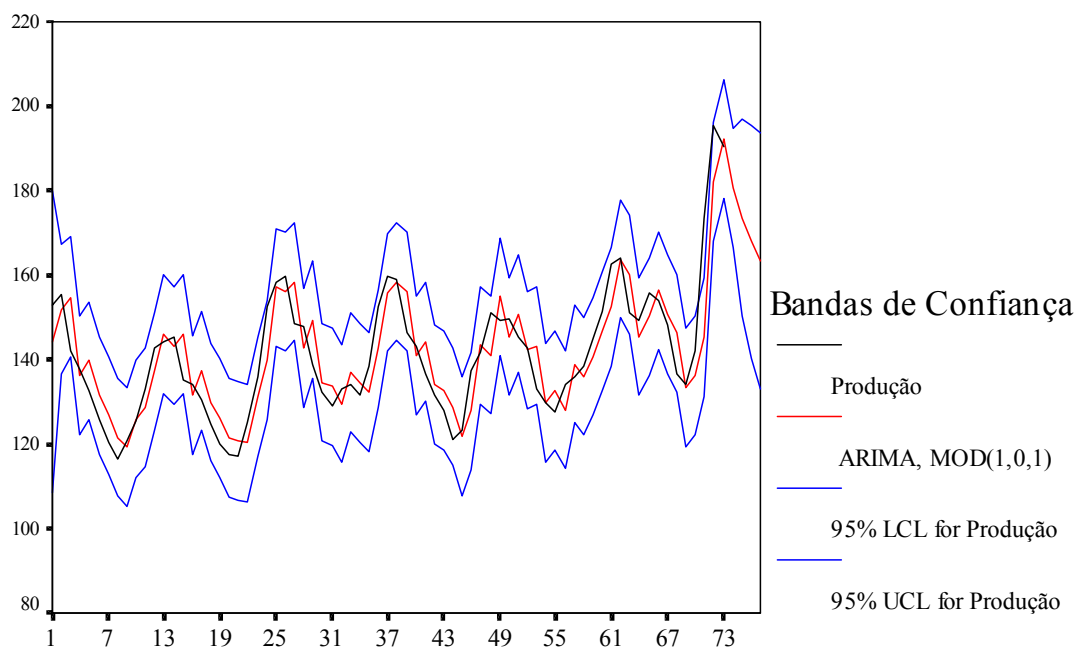
ago/79	123,18	119,80	121,90	1,28	107,93	135,88	7,01
set/79	137,34	127,06	127,87	9,47	113,89	141,84	7,01
out/79	141,57	148,30	143,39	-1,82	129,42	157,37	7,01
nov/79	151,22	143,91	141,15	10,07	127,17	155,12	7,01
dez/79	149,28	151,18	154,89	-5,61	140,91	168,86	7,01
jan/80	149,76	151,82	145,47	4,29	131,49	159,44	7,01
fev/80	145,27	145,93	150,82	-5,55	136,84	164,79	7,01
mar/80	142,8	145,36	142,26	0,54	128,29	156,24	7,01
abr/80	132,88	140,99	143,32	-10,44	129,34	157,29	7,01
mai/80	129,91	130,48	129,82	0,09	115,85	143,79	7,01
jun/80	127,5	129,02	132,70	-5,20	118,73	146,68	7,01
jul/80	134,06	131,58	128,11	5,95	114,13	142,08	7,01
ago/80	135,97	137,32	138,99	-3,02	125,01	152,96	7,01
set/80	138,43	140,00	136,03	2,40	122,06	150,01	7,01
out/80	144,82	137,98	140,73	4,09	126,75	154,70	7,01
nov/80	151,56	151,02	146,73	4,83	132,76	160,71	7,01
dez/80	162,45	153,78	152,54	9,91	138,56	166,51	7,01
jan/81	163,94	166,92	163,87	0,07	149,89	177,84	7,01
fev/81	151,24	161,67	160,14	-8,90	146,16	174,11	7,01
mar/81	149,3	143,80	145,40	3,90	131,42	159,37	7,01
abr/81	155,8	147,99	150,25	5,55	136,27	164,22	7,01
mai/81	153,81	164,35	156,32	-2,51	142,34	170,29	7,01
jun/81	148,24	148,94	150,67	-2,43	136,70	164,65	7,01
jul/81	136,77	144,92	146,22	-9,45	132,24	160,19	7,01
ago/81	134,1	133,63	133,45	0,65	119,48	147,43	7,01
set/81	142,02	135,55	136,36	5,66	122,38	150,33	7,01
out/81	173,6	148,71	145,26	28,34	131,28	159,23	7,01
nov/81	195,4	192,97	182,09	13,31	168,12	196,07	7,01
dez/81	190,5	199,32	192,14	-1,64	178,17	206,12	7,01
jan/82	186,9	177,36	180,70 ,		166,72	194,67	7,01



fev/82	171,2	167,45	173,62 ,	150,48	196,75	11,60
mar/82	172,7	165,35	167,90 ,	140,26	195,54	13,86
abr/82	164,4	162,39	163,30 ,	132,99	193,61	15,20

### Produção de Leite

São Paulo (Dez 75 a Dez 81)



Dez 75 - Dez 81

## 4.2. Medidas de Acuidade dos Métodos de Previsão de Séries Temporais

De acordo com Makridakis, a suposição básica de qualquer técnica de previsão de séries temporais é que o valor observado na série fica determinado por um padrão que se repete no tempo e por alguma influência aleatória. Isto significa dizer que mesmo quando o padrão exato que caracteriza o comportamento da série temporal tenha sido isolado, algum desvio ainda existirá entre os valores da previsão e os valores realmente observados. Essa aleatoriedade não pode ser prevista; entretanto, se isolada, sua magnitude pode ser estimada e usada para determinar a variação ou erro entre as observações e previsões realizadas.

A acuidade de um método de previsão pode ser mensurada através de muitas medidas de erro, dentre as quais pode-se citar :

erro médio =

$$\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n} \quad (1.15)$$

onde  $x_i$  é o valor observado no instante  $i$ ;

$\hat{x}_i$  é o valor previsto no instante  $i$  e

$n$  corresponde ao número de previsões efetuadas.

$$\text{erro absoluto médio} = \frac{\sum_{i=1}^n |(x_i - \hat{x}_i)|}{n} \quad (1.16)$$

erro quadrado médio =

$$\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n} \quad (1.17)$$

erro percentual absoluto =

$$\left| \frac{(x_i - \hat{x}_i)}{x_i} \right| (100) \quad (1.18)$$

erro percentual médio =

$$\frac{\sum_{i=1}^n \frac{(x_i - \hat{x}_i)}{x_i}}{n} (100) \quad (1.19)$$

erro percentual absoluto médio =

$$\frac{\sum_{i=1}^n \left| \frac{(x_i - \hat{x}_i)}{x_i} \right|}{n} (100) \quad (1.20)$$

Dessa forma, a verificação da adequação de um determinado modelo supostamente representativo da série histórica de dados é dependente da medida de erro adotada para efetuar essa validação.

### Bibliografia

- S.SOUZA, R.C. CAMARGO, M.E. Análise e Previsão de Séries Temporais. SEDIGRAF, 1996.
- MAKRIDAKIS, WHEELWRIGTH, MCGEE :Forecasting. Methods and Aplications. Wiley 1983.
- MORETTIN, PEDRO A. TOLOI, CLÉLIA M. Previsão de Séries Temporais, Edit. Atual. São Paulo 1985.