Introdução à Previsão de Demanda

PNV-3421 – Processos Estocásticos

Prof. Dr. Joao Ferreira Netto

Bibliografia Principal

- Hanke, J.E & Reitsch A.G. (1998) Business Forecasting. 6th Edition, Prentice Hall, Upper Sadle River, NJ.

Por quê Prever?

- Todas as organizações operam em uma atmosfera de incerteza.
- Apesar das incertezas, há a necessidade de tomar decisões sobre o futuro.
- Necessidade de reação rápida face a um ambiente externo dinâmico.
- Empregar métodos lógicos e científicos que se baseiam em dados gerados por eventos históricos.
- Decisões com respaldo x Decisões intuitivas.

Exemplos

- Qual será a arrecadação de tributos federais nos próximos 2 anos?
- Como as vendas serão afetadas se o orçamento com publicidade aumentar em 10%?
- Qual será o montante total concedido nas diversas linhas de empréstimo de um banco nos próximos 10 anos?
- Haverá recessão na economia nos próximos 5 anos? Quando terá início? Por quanto tempo durará? Quão severa ela será?
- Qual o futuro do sistema bancário?

Metodologia

1. **Coleta de Dados** - refere-se à coleta apropriada dos dados necessários;
2. **Preparação dos Dados** - etapa em que os dados serão adaptados para emprego dos modelos de previsão. Poderá haver excesso ou falta de dados, ou os dados poderão ser válidos apenas para determinados períodos, etc.;
3. **Elaboração do Modelo** - consiste em aderir os dados coletados a um modelo de previsão visando a minimização do erro de previsão ("forecasting error");

Metodologia

3. **Elaboração do Modelo** - ... quanto mais simples o modelo, maior a facilidade de compreensão e aceitação dos resultados;
4. **Extrapolação dos Resultados** - consiste da previsão propriamente dita, ou seja, da obtenção, interpretação e aplicação dos resultados obtidos.

Origem dos Dados

“Garbage in, garbage out” (GIGO)

Uma previsão não oferece mais precisão do que os dados nos quais ela se baseia.

Os dados devem ser...

- **Confiáveis e precisos**
  Coletados de uma fonte confiável e de maneira precisa
- **Consistentes**
  Representativos do problema em questão
- **Relevantes**
  Conseguem expressar ou traduzir o evento estudado
- **Temporais**
  Coletados ao longo de um período cronológico: **Série temporal**

Origem dos Dados

1. **Fontes Primárias** – requer métodos de coleta (levantamento) dos dados brutos.
   - Especificação do universo
   - Dimensionamento da amostra
   - Técnicas de amostragem
2. **Fontes Secundárias** – dados coletados já publicados e divulgados (fontes confiáveis!?).
   - Internet
   - Órgãos governamentais
   - Associações de classe

Abordagens de Previsão

Abordagens

Quantitativa

É empregada quando há dados históricos suficientes e adequados para prever o futuro

Qualitativa

Baseia-se no julgamento humano e na intuição, ao invés de manipulação de dados históricos

Modelos Causais

Identificação e determinação da relação entre a variável a ser prevista e as demais variáveis relevantes. Principais técnicas: ***Regressão simples****,* ***regressão multivariada****, modelos econométricos, "leading indicators"*

Séries Temporais

Identificação da forma dos dados ao longo do tempo, da eventual mudança de padrão e dos distúrbios introduzidos por influências aleatórias. Principais técnicas: ***Média móvel****,* ***suavizamento exponencial****,* ***decomposição de série temporal, ARIMA,*** *Box-Jenkins*

Modelos Qualitativos

Tipo	Características	Aspectos +	Aspectos -
Opinião Executivos	Um grupo de executivos reunem-se e propõem uma previsão	Bom para decisões estratégicas ou previsão sobre novos produtos	Uma opinião poderá ser dominante
Pesquisa de Mercado	Uso de pesquisas e entrevistas para identificar a preferência dos clientes	Bom determinante da preferência dos consumidores	Poderá ser difícil elaborar um bom questionário
Método Delphi	Busca encontrar um consenso entre diversos especialistas	Excelente para previsão de longo prazo, considerando mudanças tecnológicas e avanços científicos	Processo demorado e caro

Série Temporal

- Uma série temporal consiste em dados que são coletados, registrados ou observados em sucessivos incrementos de tempo.
- A análise de séries temporais requer a aplicação de um procedimento sistematizado.
- Decomposição da série nas componentes: tendência, cíclica, sazonalidade e irregular.

Componente de Tendência

- A tendência é uma componente de longo prazo que representa o crescimento ou o decréscimo de uma série temporal sobre um período estendido de tempo.
- Exemplo de influências: crescimento populacional, inflação, mudanças tecnológicas, aumento de produtividade.

Componente Cíclica

- A componente cíclica é uma flutuação 'em forma de onda' em torno de uma tendência, e usualmente é afetada por questões econômicas (ciclos econômicos). Padrões cíclicos tendem a se repetir a cada 2, 3 ou mais anos.

Componente Sazonal

- A componente sazonal se refere a um padrão de mudança que se repete a cada ano. Para uma série mensal, por exemplo, deve ser determinada a variação que é observada a cada mês.

Componente Irregular

- A componente irregular mede a variabilidade da série temporal depois que as outras componentes foram removidas.

Decomposição de Série Temporal

- Uma série temporal pode ser representadas pelas diferentes componentes, consideradas de forma aditiva ou multiplicativa.
- *Aditiva*:

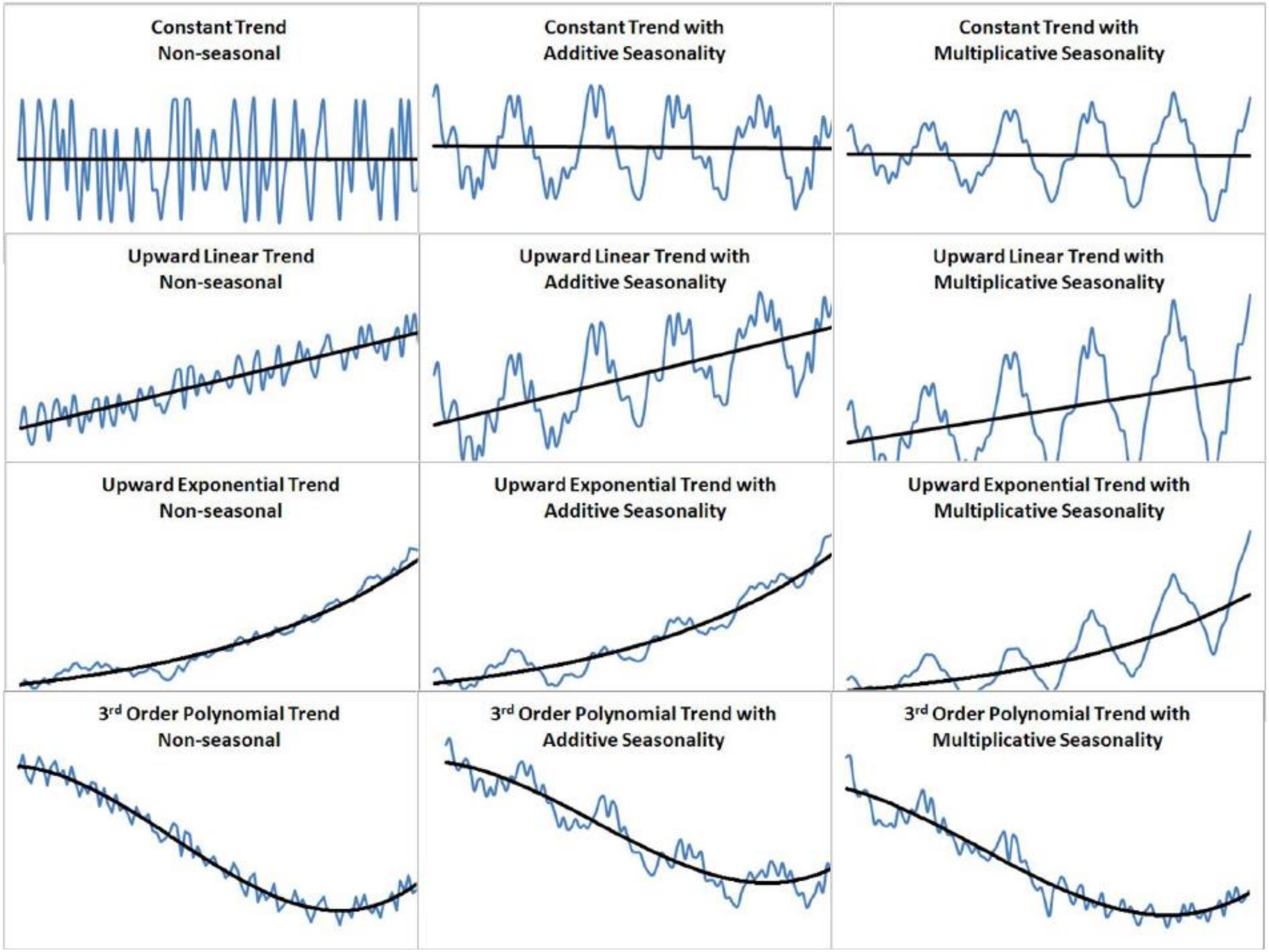
  Y_t = tendência + sazonalidade + ruído
- *Multiplicativa*:

  Y_t= tendência * sazonalidade * ruído
- OBS: i) 'Ruído' é a componente irregular; ii) a componente cíclica, usualmente de difícil identificação, pode ser incorporada na componente de tendência; iii) alguns autores representam Y_t ajustada por uma 'média'.

Decomposição de Série Temporal

- Um modelo multiplicativo é indicado nos casos em que as flutuações sazonais aumentam ou diminuem proporcionalmente com o aumento ou decréscimo do nível da série temporal.

Decomposição de Série Temporal



http://www.datasciencecentral.com/profiles/blogs/selecting-forecasting-methods-in-data-science

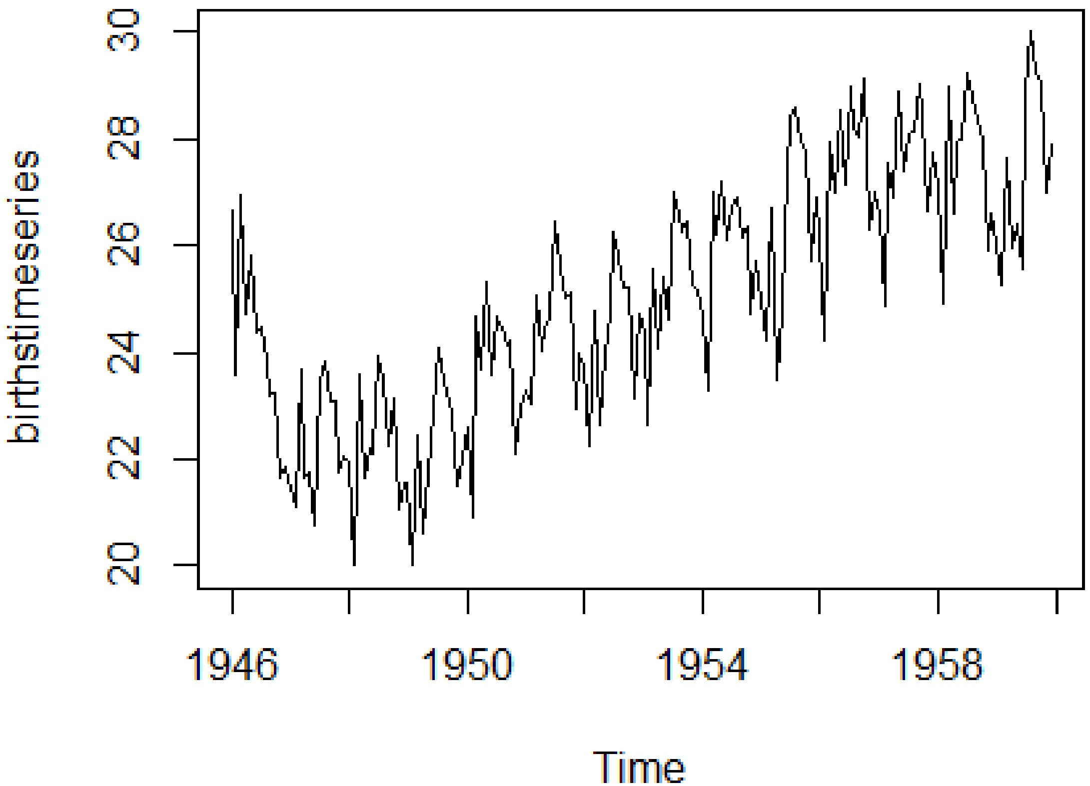
Exemplo

- Considere o registro do número de nascimentos em Nova York, entre 1946 e 1959, mês a mês.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1946	26.663	23.598	26.931	24.740	25.806	24.364	24.477	23.901	23.175	23.227	21.672	21.870
1947	21.439	21.089	23.709	21.669	21.752	20.761	23.479	23.824	23.105	23.110	21.759	22.073
1948	21.937	20.035	23.590	21.672	22.222	22.123	23.950	23.504	22.238	23.142	21.059	21.573
1949	21.548	20.000	22.424	20.615	21.761	22.874	24.104	23.748	23.262	22.907	21.519	22.025
1950	22.604	20.894	24.677	23.673	25.320	23.583	24.671	24.454	24.122	24.252	22.084	22.991
1951	23.287	23.049	25.076	24.037	24.430	24.667	26.451	25.618	25.014	25.110	22.964	23.981
1952	23.798	22.270	24.775	22.646	23.988	24.737	26.276	25.816	25.210	25.199	23.162	24.707
1953	24.364	22.644	25.565	24.062	25.431	24.635	27.009	26.606	26.268	26.462	25.246	25.180
1954	24.657	23.304	26.982	26.199	27.210	26.122	26.706	26.878	26.152	26.379	24.712	25.688
1955	24.990	24.239	26.721	23.475	24.767	26.219	28.361	28.599	27.914	27.784	25.693	26.881
1956	26.217	24.218	27.914	26.975	28.527	27.139	28.982	28.169	28.056	29.136	26.291	26.987
1957	26.589	24.848	27.543	26.896	28.878	27.390	28.065	28.141	29.048	28.484	26.634	27.735
1958	27.132	24.924	28.963	26.589	27.931	28.009	29.229	28.759	28.405	27.945	25.912	26.619
1959	26.076	25.286	27.660	25.951	26.398	25.565	28.865	30.000	29.261	29.012	26.992	27.897

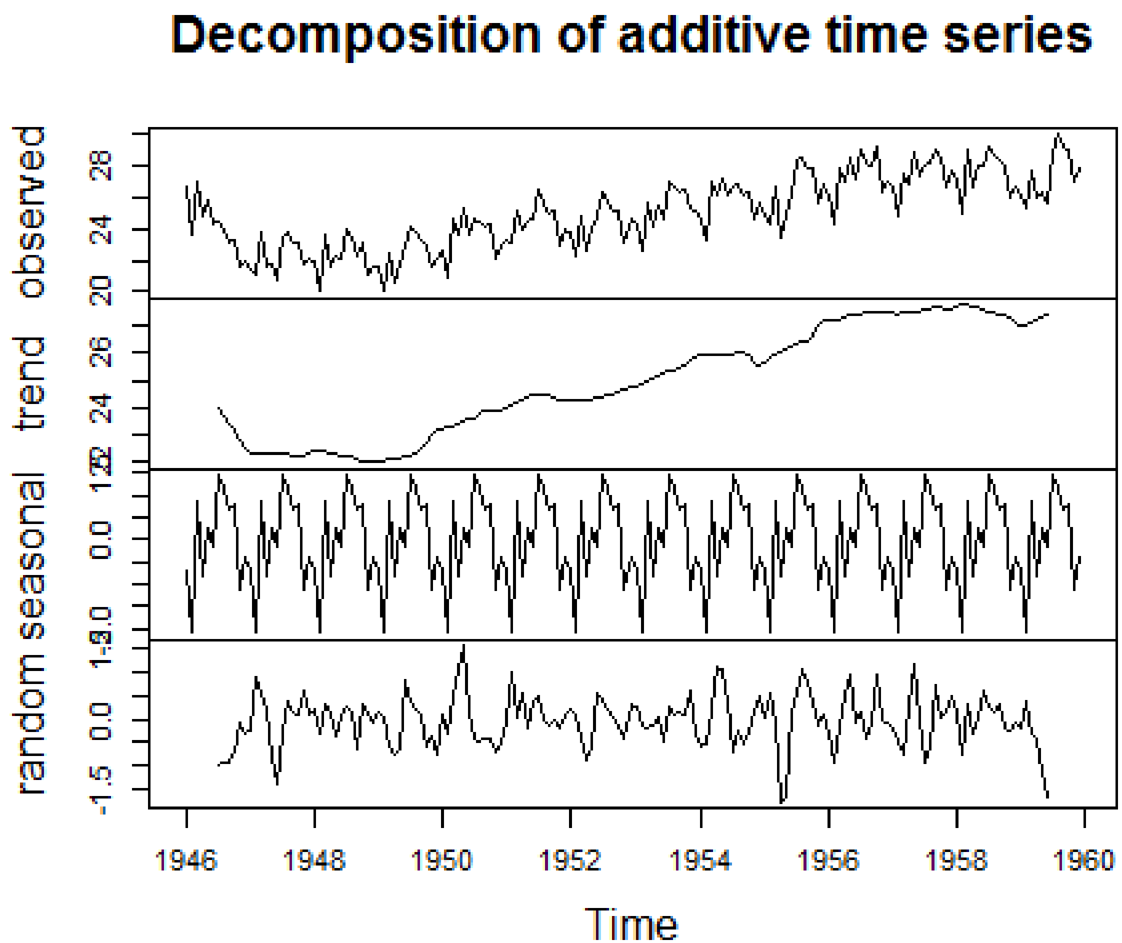
Fonte: http://a-little-book-of-r-for-time-series.readthedocs.io/en/latest/src/timeseries.html

Exemplo



Fonte: http://a-little-book-of-r-for-time-series.readthedocs.io/en/latest/src/timeseries.html

Exemplo



Fonte: http://a-little-book-of-r-for-time-series.readthedocs.io/en/latest/src/timeseries.html

Análise de Autocorrelação

Autocorrelação

- Autocorrelação de uma série temporal (variável medida ao longo do tempo), é a correlação desta variável com ela mesma, porém defasada de 1 ou mais períodos.

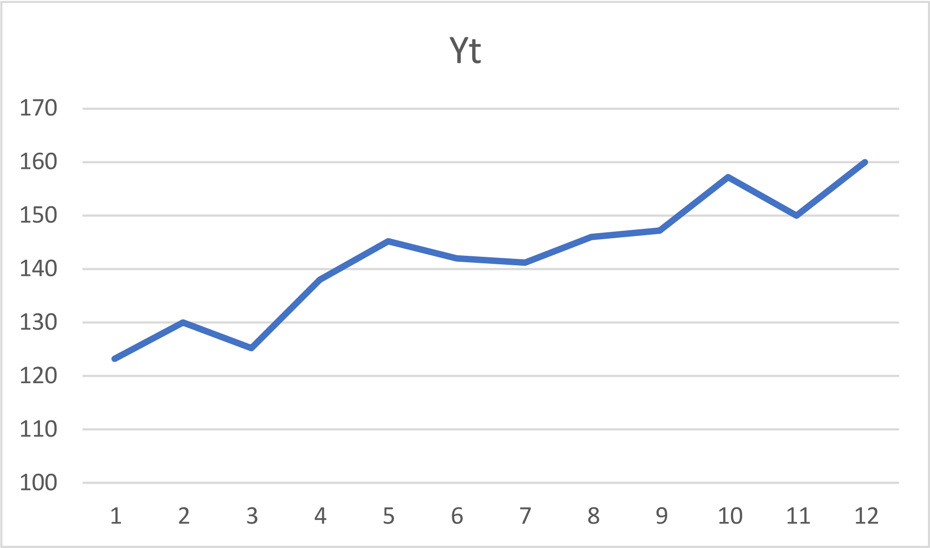
Autocorrelação

- Coeficiente de autocorrelação:
- $r_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$
- Onde:
  - r_k - coeficiente de autocorrelação para defasagem de k períodos
  - Y_t - observação da série temporal no instante t
  - Y_{t-k} - observação da série temporal no instante $t - k$
  - $\bar{Y}$ - valor médio da série temporal

Autocorrelação – Exemplo 1

- Para os dados da tabela, calcular o coeficiente de autocorrelação, para k=1 (defasagem de 1 período).

t	Mês	Y_t
1	Janeiro	123
2	Fevereiro	130
3	Março	125
4	Abril	138
5	Maio	145
6	Junho	142
7	Julho	141
8	Agosto	146
9	Setembro	147
10	Outubro	157
11	Novembro	150
12	Dezembro	160



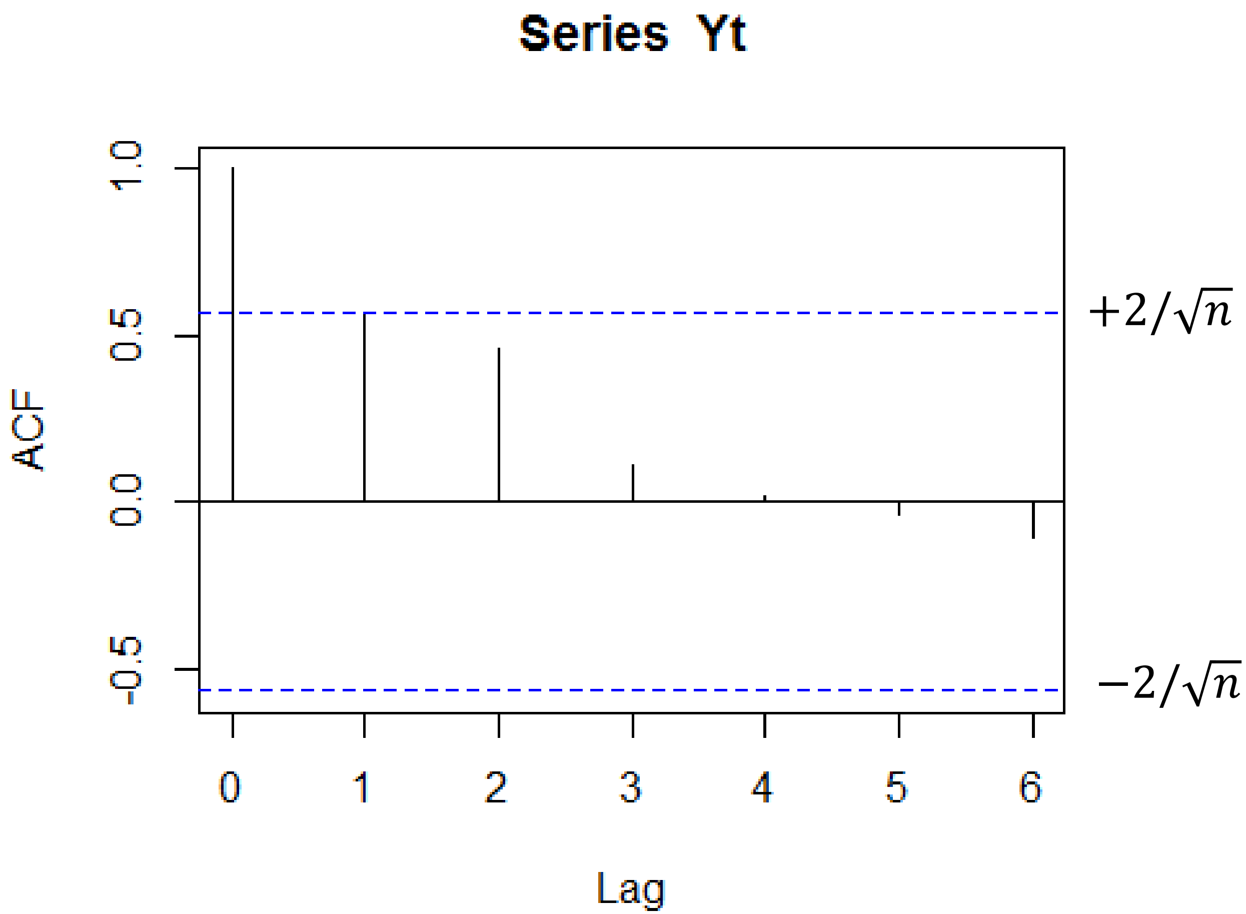
Autocorrelação – Exemplo 1

- Para os dados da tabela, calcular o coeficiente de autocorrelação, para k=1 (defasagem de 1 período).

t	Mês	Y_t	Y_{t-1}	$(Y_t - \bar{Y})$	$(Y_{t-1} - \bar{Y})$	$(Y_t - \bar{Y})^2$	$(Y_t - \bar{Y})(Y_{t-1} - \bar{Y})$
1	Janeiro	123		-19		361	
2	Fevereiro	130	123	-12	-19	144	228
3	Março	125	130	-17	-12	289	204
4	Abril	138	125	-4	-17	16	68
5	Maio	145	138	3	-4	9	-12
6	Junho	142	145	0	3	0	0
7	Julho	141	142	-1	0	1	0
8	Agosto	146	141	4	-1	16	-4
9	Setembro	147	146	5	4	25	20
10	Outubro	157	147	15	5	225	75
11	Novembro	150	157	8	15	64	120
12	Dezembro	160	150	18	8	324	144
	$\bar{Y} =$	142				1474	843
						$r_1 =$	0,57

Autocorrelação

- Autocorrelograma ($r_1 = 0{,}57; r_2 = 0{,}46$)



Autocorrelação

- A análise de autocorrelação fornece elementos para compreender a série de dados.
- Se a série é aleatória, a correlação entre Y_t e Y_{t-1} é próxima de zero.
- Se a série tem tendência, então Y_t e Y_{t-1} são altamente correlacionados, sendo r_k *significativamente diferente de zero* para $k = 1$ e $k = 2$..., e progressivamente diminui para defasagens maiores.
- Se a série for sazonal, então a correlação será mensal, quadrimestral, etc.

Autocorrelação

- Como saber se os coeficientes de autocorrelação, para diferentes defasagens, são significativamente diferente de zero?
- É possível demostrar que os coeficientes de autocorrelação de dados aleatórios podem ser aproximados por uma curva normal com média zero e desvio padrão $1/\sqrt{n}$.
- Podemos assim aplicar um teste de hipótese para determinar se os coeficientes de autocorrelação de uma amostra, para uma determinada defasagem, possuem média zero.

Autocorrelação

- Para cálculo do erro padrão dos coeficientes de autocorrelação:
- $$SEr(k) = \sqrt{\frac{1+2\sum_{i=1}^{k-1}(r_i)^2}{n}}$$
- Onde:
  - $SEr(k)$ - erro padrão dos coeficientes de autocorrelação
  - r_i - autocorrelação com defasagem i
  - n – número de observações na série temporal

Autocorrelação

- Aproximando a curva normal pela distribuição de Student, um valor será considerado igual a zero se estiver no intervalo dado por $[-t * SEr(k), +t * SEr(k)]$, onde t é o valor da estatística da distribuição de Student para um dado nível de significância.

Autocorrelação

- Teste de hipótese: $\begin{cases} H_0: \rho_1 = 0 \\ H_1: \rho_1 \neq 0 \end{cases}$
- $t = \frac{r_k - \rho_k}{SEr(k)}$
- Para o exemplo 1, verificar se r_1 e r_2 são significativamente diferentes de zero, para nível de significância $\alpha = 5\%$, e $n - 1 = 11$ graus de liberdade.
- O valor tabelado de t é 2,2.
- Assim, se $t < -2{,}2$ ou $t > +2{,}2$, rejeita-se a hipótese nula.

Autocorrelação

- $SEr(1) = \sqrt{\frac{1+2\sum_{i=1}^{k-1}(r_i)^2}{n}} = \sqrt{\frac{1+2\sum_{i=1}^{1-1}(r_i)^2}{12}} = \sqrt{\frac{1}{12}} =$ 0,2887
- $t = \frac{r_1 - \rho_1}{SEr(1)} = \frac{0{,}572 - 0{,}0}{0{,}2887} = 1{,}98 < 2{,}2$
- Não rejeitar a hipótese nula para k=1.
- $SEr(2) = \sqrt{\frac{1+2\sum_{i=1}^{k-1}(r_i)^2}{n}} = \sqrt{\frac{1+2\sum_{i=1}^{2-1}(r_i)^2}{12}} = \sqrt{\frac{1{,}6544}{12}} =$ 0,371
- $t = \frac{r_2 - \rho_2}{SEr(2)} = \frac{0{,}463 - 0{,}0}{0{,}371} = 1{,}25 < 2{,}2$
- Não rejeitar a hipótese nula para k=2.

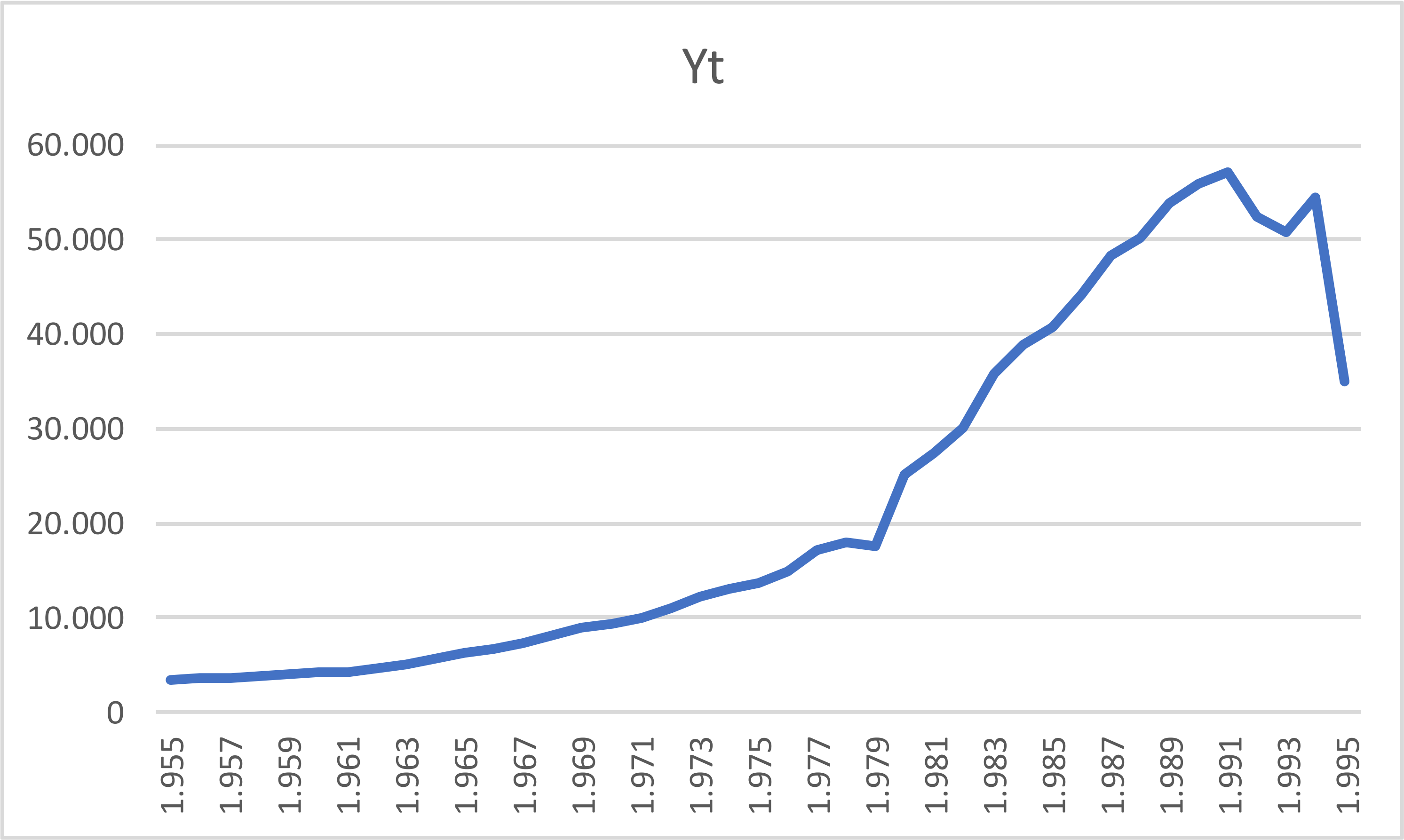
Autocorrelação – Exemplo 2

- Para os dados da tabela, calcular o coeficiente de autocorrelação, para k=1, k=2, k=3 e k=4.

t	Y_t	t	Y_t	t	Y_t	t	Y_t
1.955	3.307	1.966	6.769	1.976	14.950	1.986	44.282
1.956	3.556	1.967	7.296	1.977	17.224	1.987	48.440
1.957	3.601	1.968	8.178	1.978	17.946	1.988	50.251
1.958	3.721	1.969	8.844	1.979	17.514	1.989	53.794
1.959	4.036	1.970	9.251	1.980	25.195	1.990	55.972
1.960	4.134	1.971	10.006	1.981	27.357	1.991	57.242
1.961	4.268	1.972	10.991	1.982	30.020	1.992	52.345
1.962	4.578	1.973	12.306	1.983	35.883	1.993	50.838
1.963	5.093	1.974	13.101	1.984	38.828	1.994	54.559
1.964	5.716	1.975	13.639	1.985	40.715	1.995	34.925
1.965	6.357						

Autocorrelação – Exemplo 2

- Para os dados da tabela, calcular o coeficiente de autocorrelação, para k=1, k=2, k=3 e k=4.

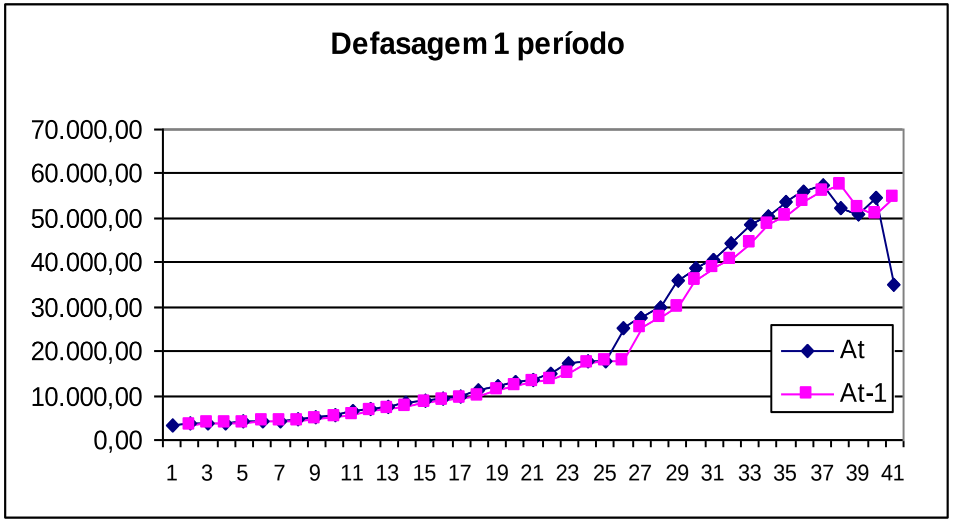


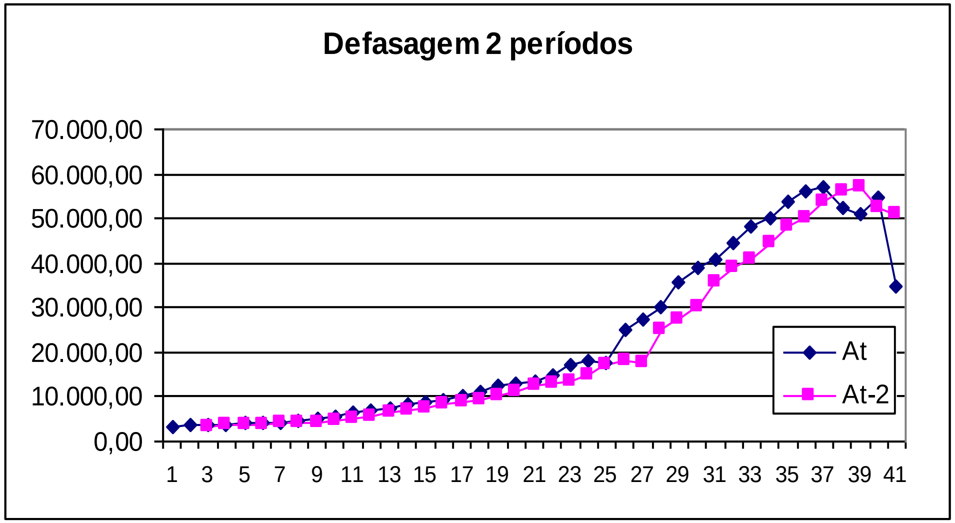
Autocorrelação – Exemplo 2

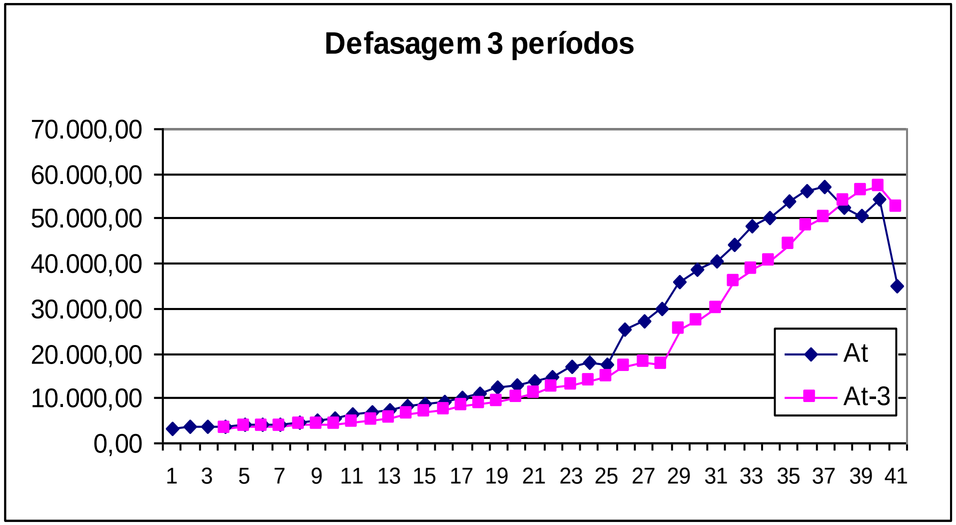
Defasagem	Coeficiente de Autocorrelação	SEr(k)	t
$k = 1$	0,960	0,156	6,150
$k = 2$	0,901	0,263	3,422
$k = 3$	0,838	0,330	2,538
$k = 4$	0,758	0,379	2,002

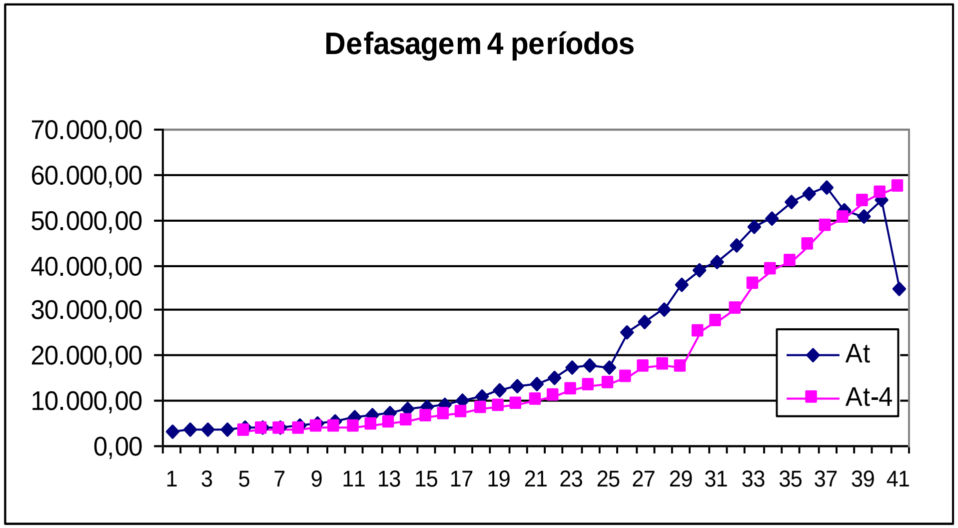
- Para nível de significância $\alpha = 5\%$, e $n - 1 = 40$ graus de liberdade, o valor crítico é $t = 2{,}02$ (no Excel: =INV.T(97,5%;40)). Assim, r_1, r_2 e r_3 podem ser considerados diferentes de 0.

Autocorrelação – Exemplo 2

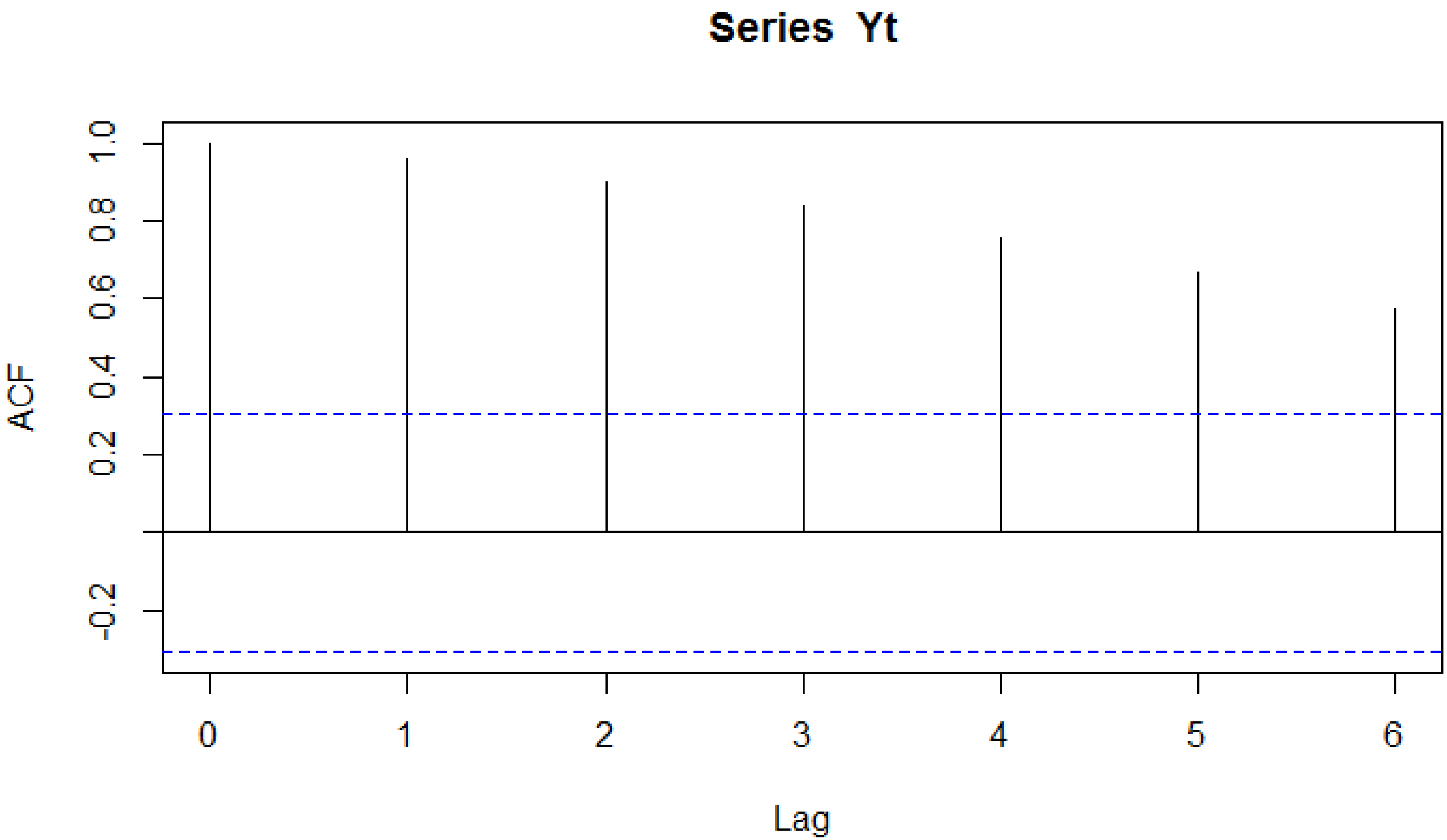








Autocorrelação – Exemplo 2



Autocorrelação – Exemplo 3

- Para os dados da tabela, calcular o coeficiente de autocorrelação, para k=1, k=2, k=3 e k=4.

Ano	Trim.	Y_t
1984	1	147,6
1984	2	251,8
1984	3	273,1
1984	4	249,1
1985	1	139,3
1985	2	221,2
1985	3	260,2
1985	4	259,5
1986	1	140,5
1986	2	245,5
1986	3	298,8
1986	4	287,0
1987	1	168,8

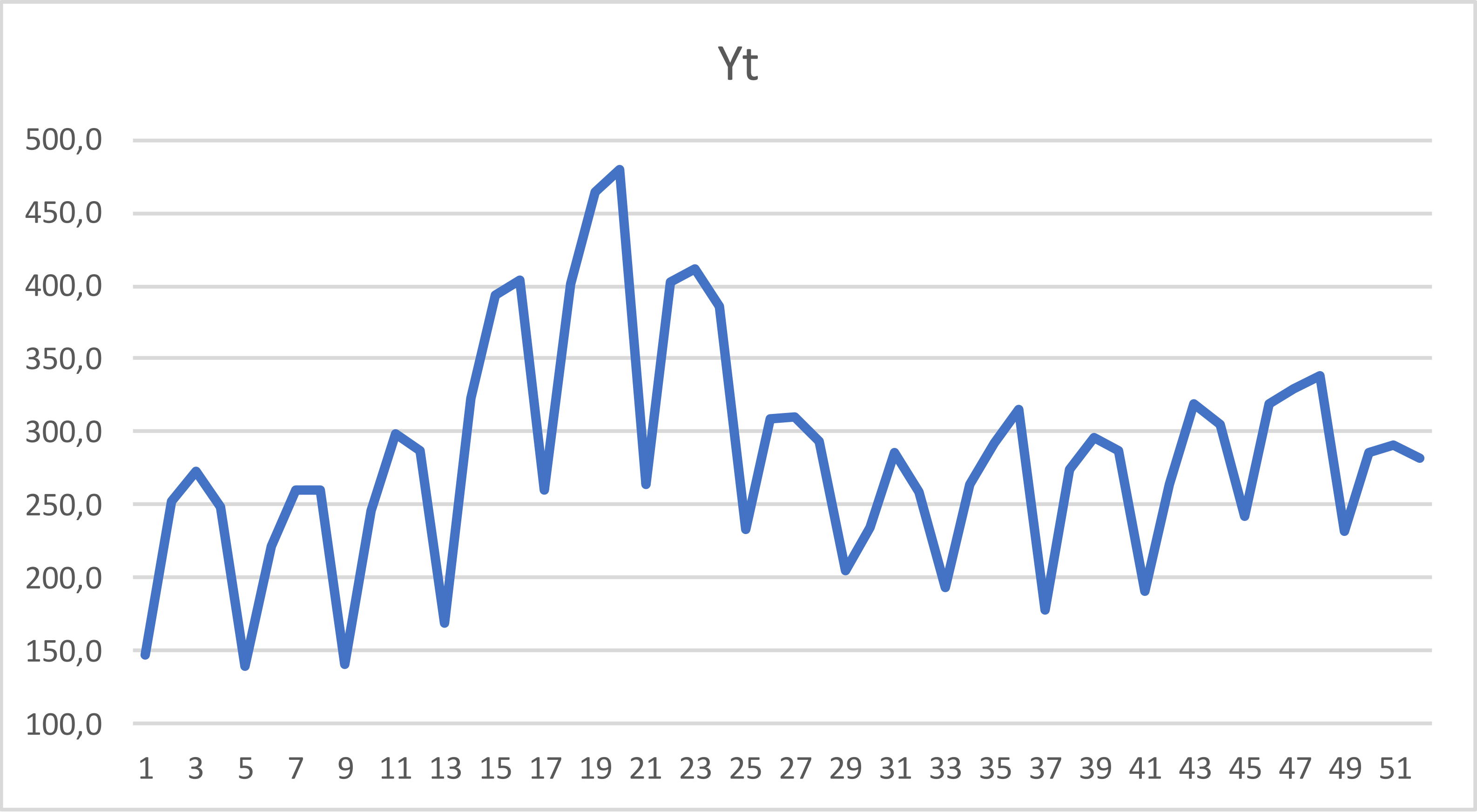
Ano	Trim.	Y_t
1987	2	322,6
1987	3	393,5
1987	4	404,3
1988	1	259,7
1988	2	401,1
1988	3	464,6
1988	4	479,7
1989	1	264,4
1989	2	402,6
1989	3	411,3
1989	4	385,9
1990	1	232,7
1990	2	309,2

Ano	Trim.	Y_t
1990	3	310,7
1990	4	293,0
1991	1	205,1
1991	2	234,4
1991	3	285,4
1991	4	258,7
1992	1	193,2
1992	2	263,7
1992	3	292,5
1992	4	315,2
1993	1	178,3
1993	2	274,5
1993	3	295,4

Ano	Trim.	Y_t
1993	4	286,4
1994	1	190,8
1994	2	263,5
1994	3	318,8
1994	4	305,5
1995	1	242,6
1995	2	318,8
1995	3	329,6
1995	4	338,2
1996	1	232,1
1996	2	285,6
1996	3	291,0
1996	4	281,4

Autocorrelação – Exemplo 3

- Para os dados da tabela, calcular o coeficiente de autocorrelação, para k=1, k=2, k=3 e k=4.

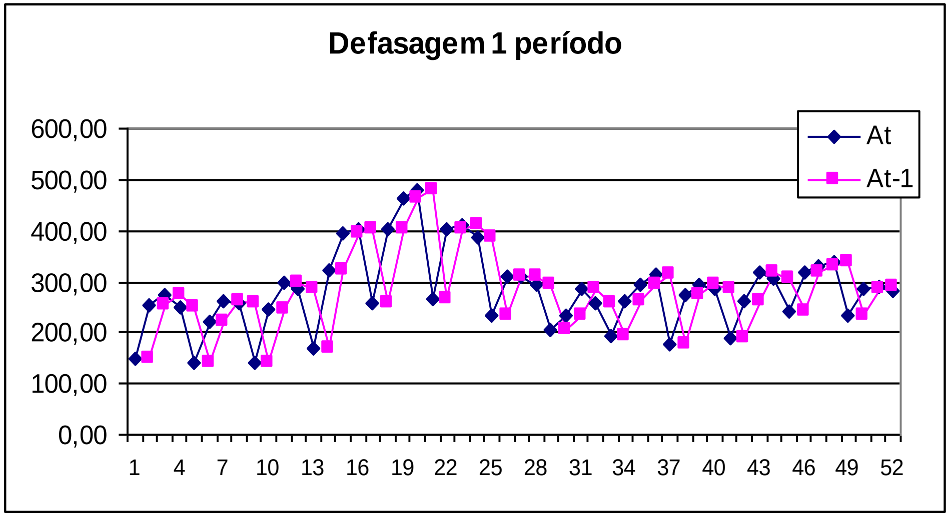


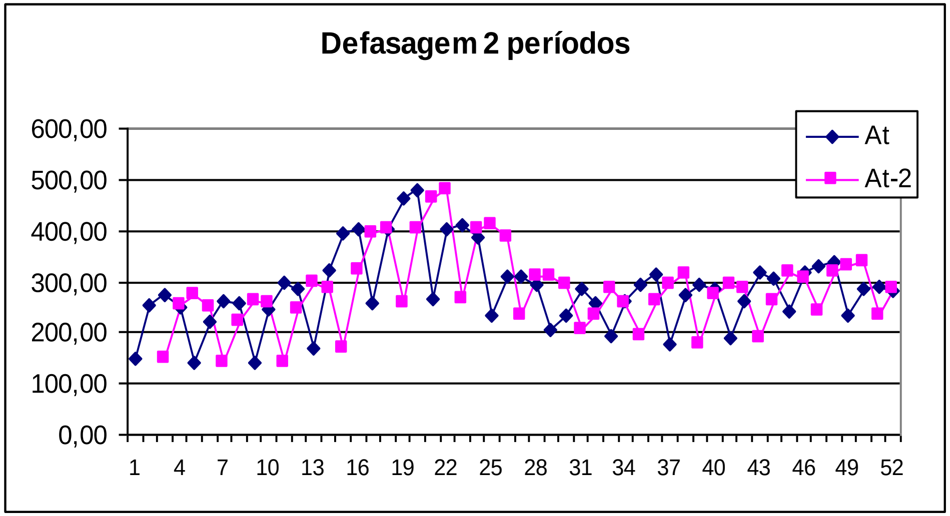
Autocorrelação – Exemplo 3

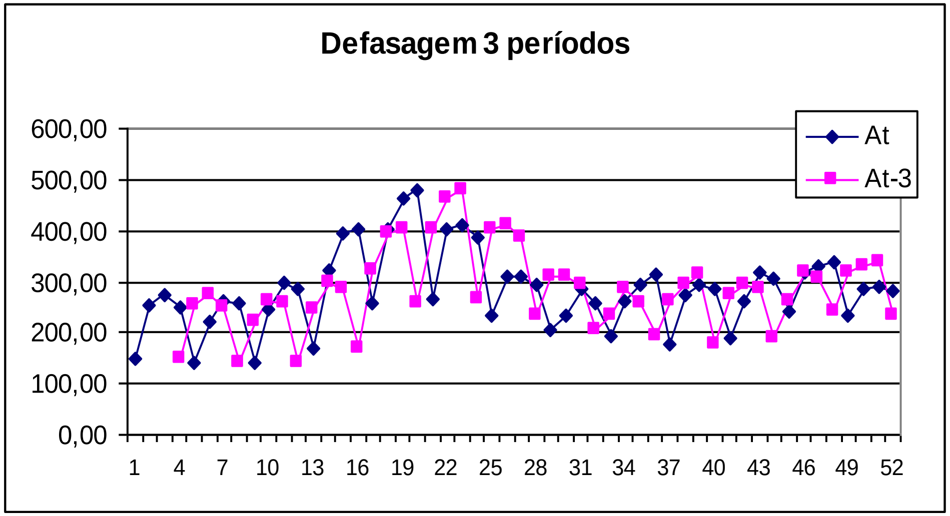
Defasagem	Coeficiente de Autocorrelação	SEr(k)	t
$k = 1$	0,393	0,139	2,833
$k = 2$	0,154	0,159	0,970
$k = 3$	0,294	0,161	1,819
$k = 4$	0,744	0,171	4,337

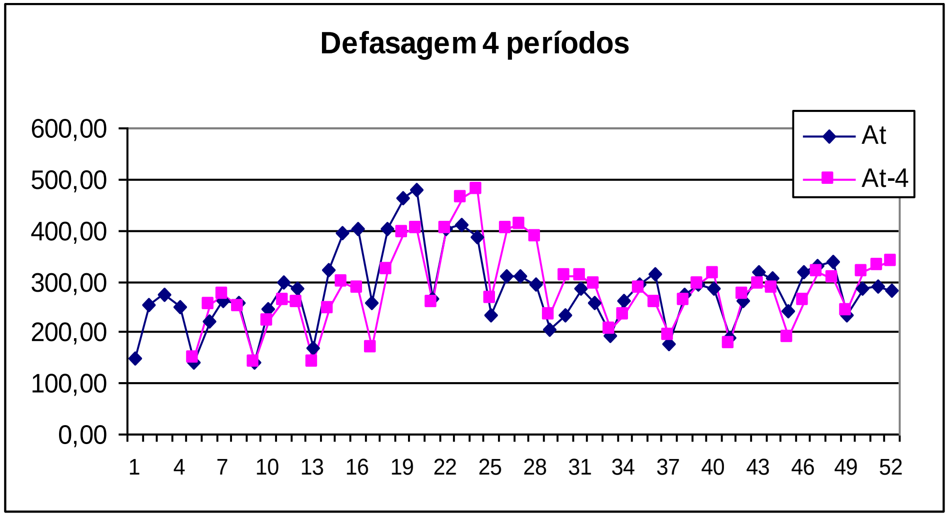
- Para nível de significância $\alpha = 5\%$, e $n - 1 = 51$ graus de liberdade, o valor crítico é $t = 2{,}01$ (no Excel: =INV.T(97,5%;51)). Assim, r_2 e r_3 não podem ser considerados diferentes de 0.

Autocorrelação – Exemplo 3

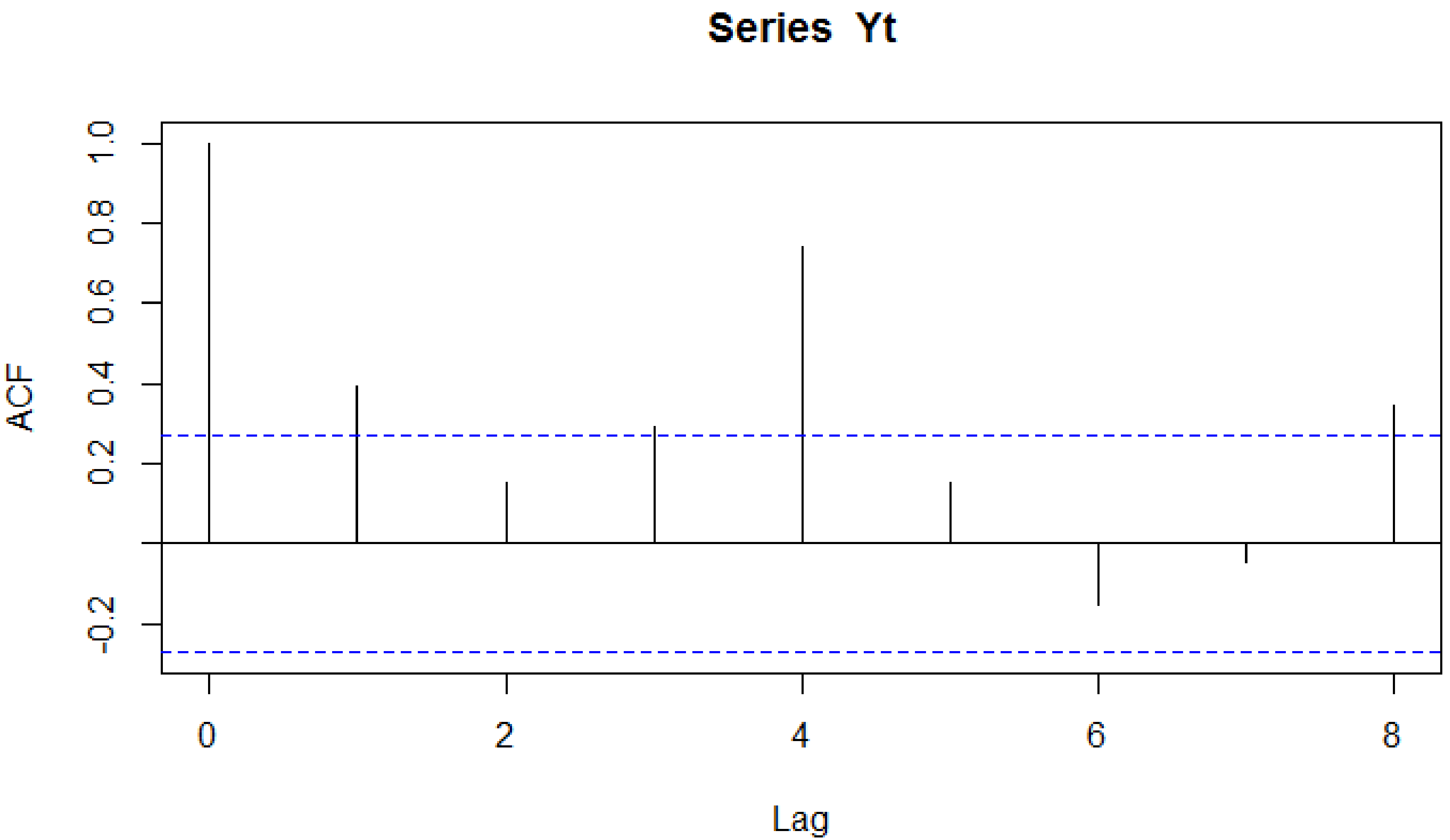






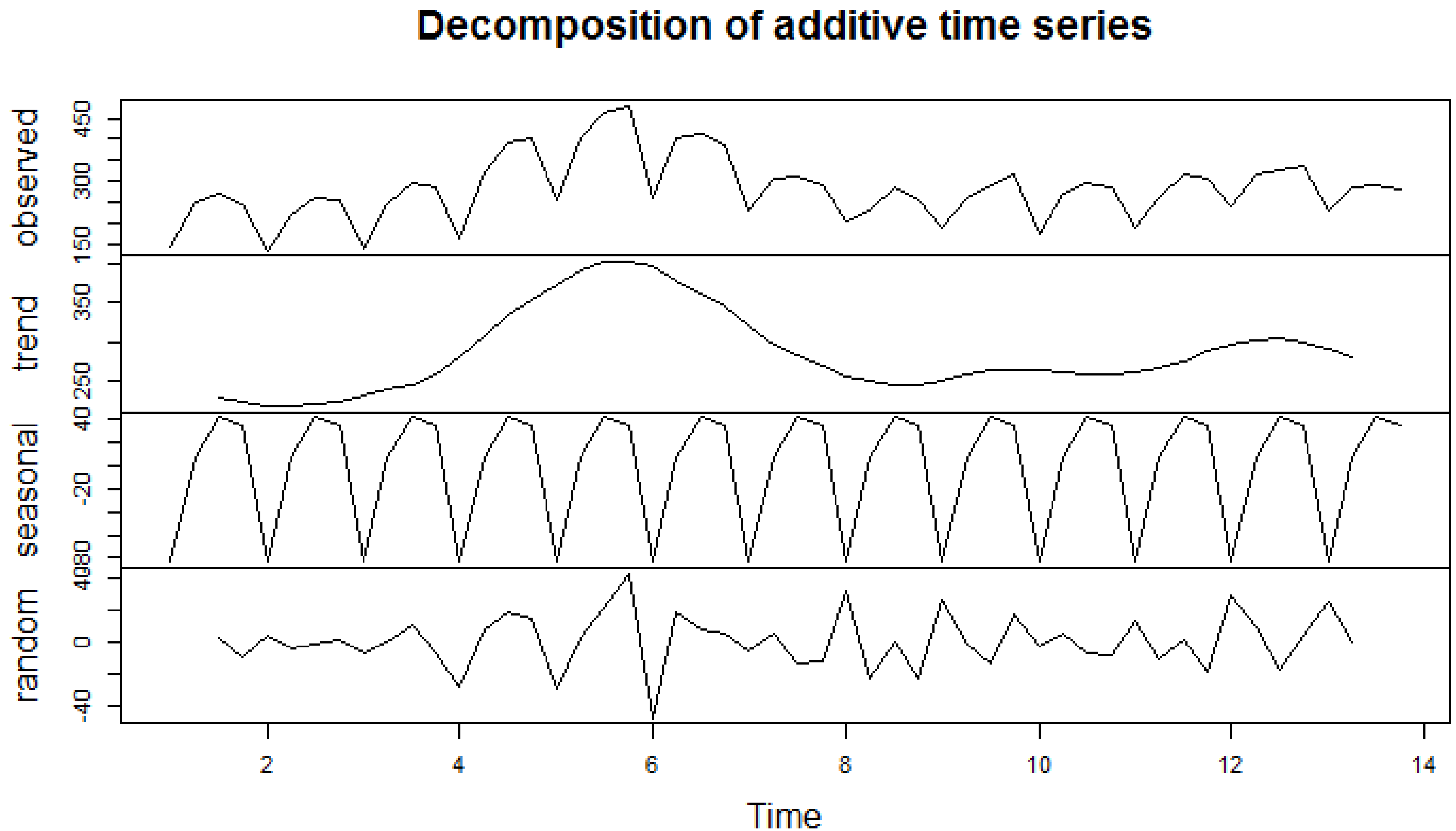


Autocorrelação – Exemplo 3



A série de dados possui correlação sazonal, com defasagem de 4 períodos.

Autocorrelação – Exemplo 3



Escolha de uma Técnica de Previsão

Escolha de uma Técnica de Previsão

- Perguntas essenciais:
  - Por quê a previsão é necessária?
  - Quem fara uso?
  - Quais as características dos dados existentes?
  - Que período deverá ser previsto?
  - Qual a precisão desejada?
  - Qual o custo da previsão?

Escolha de uma Técnica de Previsão

- Envolve:
  - Definir a natureza do problema.
  - Explicar a natureza dos dados.
  - Descrever os benefícios e as limitações das possíveis técnicas de previsão.
  - Estabelecer critérios para selecionar uma determinada técnica de previsão.

Técnica de Previsão para Dados Estacionários

Séries estacionárias são aquelas em que o valor médio não muda ao longo do tempo.
As condições que influenciam uma série de dados são estáveis. Ex: vendas de um produto que já alcançou um estado de maturação; produção de uma linha de montagem onde a taxa de falha das máquinas é conhecida e controlada.
Os dados são estáveis pois passaram por fatores de correção ou ajustes, como crescimento populacional ou inflação.
A série foi transformada em uma série estável aplicando logaritmo, raiz quadrada, etc.

Técnicas aplicáveis: métodos ingênuos, média móvel, suavizamento exponencial, método de Box-Jenkins.

Técnica de Previsão para Dados com Tendência

Série com componente de longo prazo que representa o crescimento ou o decréscimo de uma série temporal sobre um período estendido de tempo.
Aumento de produtividade, ou era tecnológica introduzindo mudanças no estilo de vida. Ex: aumento de demanda por componentes eletrônicos com a evolução de computadores de smart-phones.
Aumento da população, que resulta no aumento de demanda por matéria prima e produtos acabados.
Aumento progressivo da fatia de mercado de um determinado produto.

Técnicas aplicáveis: média móvel, suavizamento exponencial linear de Holt e de Brown, suavizamento exponencial quadrático de Brown, modlo de Gompertz, curvas de crescimento.

Técnica de Previsão para Dados com Sazonalidade

Componente sazonal indica um padrão de mudança que se repete a cada ano, e requer a aplicação de modelos aditivos ou multiplicativos, que permitam estimar os *índices de sazonalidade*.

Questões meteorológicas influenciando as variáveis de interesse. Ex: consumo de bebida e sorvete no verão.

Calendário anual influenciando as variáveis de interesse. Ex: vendas no período que antecede a volta às aulas; vendas no Natal, dia das Mães, etc.

Técnicas aplicáveis: suavizamento exponencial de Winter, regressão multivariada, método de Box-Jenkins.

Técnica de Previsão para Dados Cíclicos

A componente cíclica é uma flutuação 'em forma de onda' em torno de uma tendência, e usualmente é afetada por questões econômicas (ciclos econômicos)
Ciclos econômicos.
Mudanças no comportamento / preferências / hábitos da população.
Variação da população (Ex: reduções causadas por migração, mortes, desastres naturais, etc.)

Técnicas aplicáveis: técnicas de decomposição, indicadores econômicos, modelos econométricos, regressão multi-variada, método de Box-Jenkins.

Técnica de Previsão para Dados com Sazonalidade

Método	Forma dos Dados	Horizonte de Planejamento	Tipo de Modelo	Requisito Mínimo de Dados	
				Não Sazonal	Sazonal
Ingênuo	ST, T, S	S	TS	1	
Média simples	ST	S	TS	30	
Média móvel	ST	S	TS	4-20	
Suavizamento exponencial	ST	S	TS	2	
Suavizamento exponencial linear	T	S	TS	3	
Suavizamento exponencial quadrático	T	S	TS	4	
Suavizamento exponencial sazonal	S	S	TS		2xL
Filtro adaptativo	S	S	TS		5xL
Regressão simples	T	I	C	10	
Regressão multivariada	C, S	I	C	10xV	

Forma dos Dados ST: Estacionário; T: Tendência; S: Sazonal; C:Cíclico
Horizonte de Planejamento S: Curto Prazo (menos de 3 meses); I: Intermediário; L: Longo prazo
Tipo de Modelo TS: Série temporal; C: Causal
Sazonal L: Extensão da sazonalidade
Variável V: Número de variáveis

Técnica de Previsão para Dados com Sazonalidade

Método	Forma dos dados	Horizonte de Planejamento	Tipo de Modelo	Requisito Mínimo de Dados Não Sazonal	Sazonal
Decomposição clássica	S	S	TS		5xL
Modelos de tendência exponencial	T	I, L	TS	10	
Ajuste de curva-S	T	I, L	TS	10	
Modelos de Grompertz	T	I, L	TS	10	
Curvas de crescimento	T	I, L	TS	10	
Census II	S	S	TS		6xL
Box-Jenkins	ST, T, C, S	S	TS	24	3xL
Indicadores principais	C	S	C	24	
Modelos econométricos	C	S	C	30	
Regressão de série temporal múltipla	T, S	I, L	C		6xL

Forma dos Dados ST: Estacionário; T: Tendência; S: Sazonal; C:Cíclico

Horizonte de Planejamento S: Curto Prazo (menos de 3 meses); I: Intermediário; L: Longo prazo

Tipo de Modelo TS: Série temporal; C: Causal

Sazonal L: Extensão da sazonalidade

Variável V: Número de variáveis

Medindo o Erro de Previsão

Medindo o Erro de Previsão

- Notação matemática:

  Y_t valor da série temporal no instante t

  $\hat{Y}_t$ valor previsto para o instante t

  $e_t = Y_t - \hat{Y}_t$ erro de previsão ou resíduo
- Um modelo de previsão consiste, para uma série histórica de dados, fazer previsões para os valores desta série.

MAD - Mean Absolute Deviation
Desvio Absoluto Médio

- $\text{MAD} = \frac{\sum_{t=1}^{n} |Y_t - \hat{Y}_t|}{n}$
- Esta medida de erro refere-se à média das magnitudes dos erros de previsão, e tem como vantagem ter a mesma unidade de medida da série original.

MSE - Mean Square Error
Erro Quadrático Médio

- $\text{MSE} = \frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}{n}$
- Esta medida de erro penaliza os maiores desvios, já que computa o erro quadrático. Pode ser que o modelo desejado de previsão seja aquele que erre ‘pouco’, e o MSE dará subsídios para comparar diferentes técnicas.

MAPE - Mean Absolute Percentage Error
Erro Absoluto Percentual Médio

- $$\text{MAPE} = \frac{\sum_{t=1}^{n} \frac{|Y_t - \widehat{Y}_t|}{Y_t}}{n}$$
- Esta medida de erro expressa o valor do erro (desvio), em relação ao valor da série de dados, indicando, assim, o erro percentual médio.

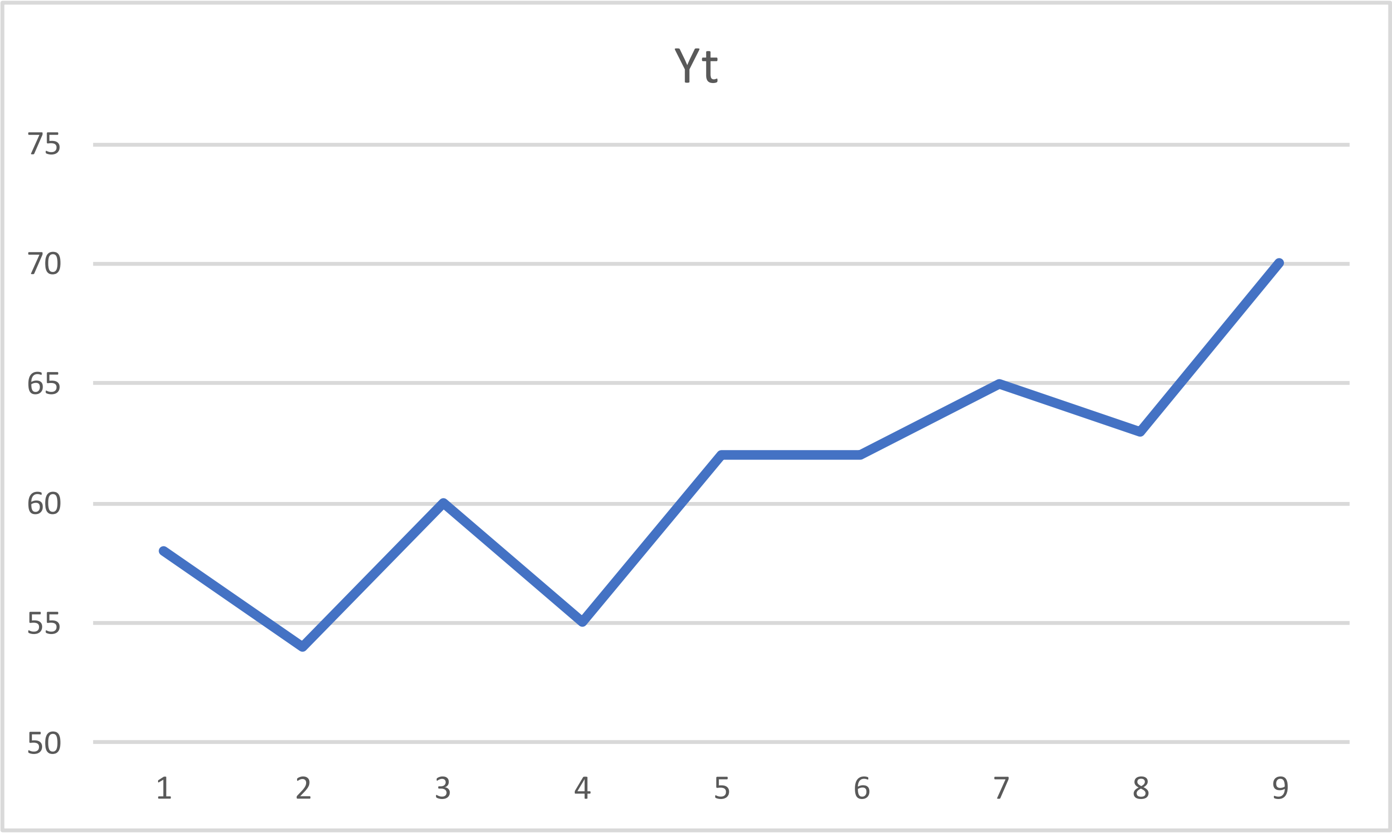
MPE - Mean Percentage Error
Erro Percentual Médio

- $$\text{MPE} = \frac{\sum_{t=1}^{n} \frac{(Y_t - \widehat{Y}_t)}{Y_t}}{n}$$
- Esta medida de erro permite avaliar se há viés no modelo de previsão. Se o modelo não for enviesado, então o MPE será próximo de zero. Caso MPE seja 'muito' positivo, conclui-se que o modelo, em média, subestima o valor da série temporal. Em caso contrário, o modelo de previsão, em média, superestima o valor da série temporal.

Exemplo

- Seja o seguinte conjunto de dados: 58, 54, 60, 55, 62, 62, 65, 63, 70. Estes dados indicam o número de clientes que procuram um determinado serviço especializado.
- Considerando que o valor da série histórica de um período é uma boa forma de prever o que vai ocorrer no período seguinte, calcule a previsão para o instante 10, e os erros de previsão associados.

Exemplo



Exemplo

t	Y_t	$\hat{Y}_t$	e_t	$\lvert e_t \rvert$	${e_t}^2$	$\lvert e_t \rvert / Y_t$	e_t / Y_t
1	58						
2	54	58	-4	4	16	7,4%	-7,4%
3	60	54	6	6	36	10,0%	10,0%
4	55	60	-5	5	25	9,1%	-9,1%
5	62	55	7	7	49	11,3%	11,3%
6	62	62	0	0	0	0,0%	0,0%
7	65	62	3	3	9	4,6%	4,6%
8	63	65	-2	2	4	3,2%	-3,2%
9	70	63	7	7	49	10,0%	10,0%
				MAD	**MSE**	**MAPE**	**MPE**
			1,5	4,25	23,5	6,9%	2,0%

- Os erros indicam que o desvio médio é 4,25 clientes (6,9%); a previsão não apresenta viés, já que o MPE ~ 2,0%. A previsão para o período 10 é 70.

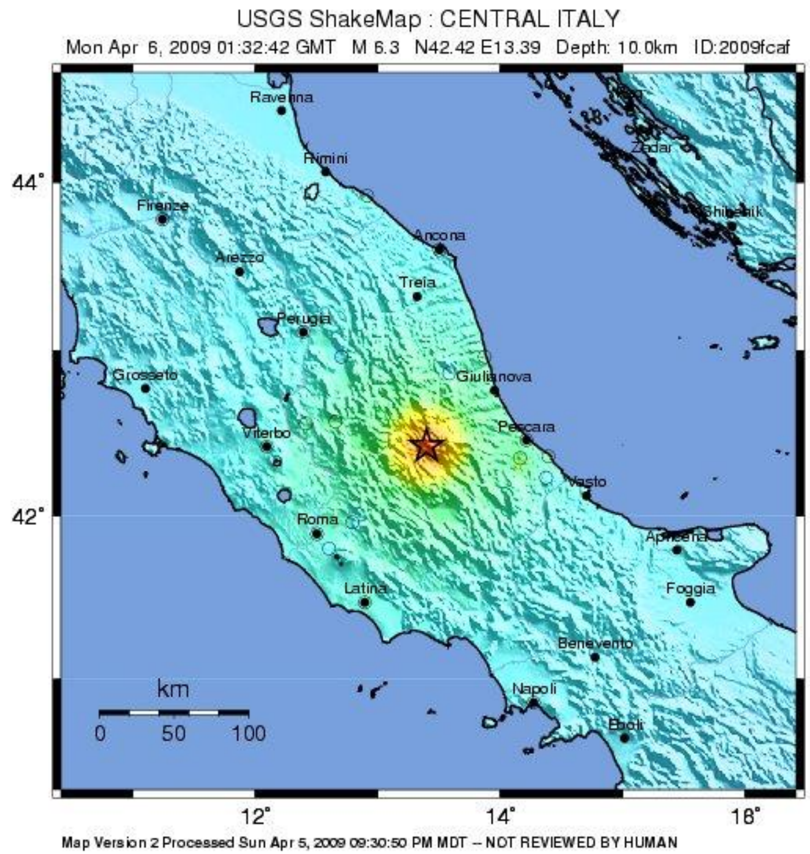
Critério de Aceitação do Modelo de Previsão

- Os resíduos podem ser considerados como uma série aleatória? Para isso, aplicar uma análise de autocorrelação aos resíduos.
- Os resíduos são distribuídos aproximadamente por uma curva normal?
- A técnica é simples de ser usada, e de fácil compreensão pelos tomadores de decisão?

Terremoto em L'Aquila, Itália (2009)

Terremoto em L'Aquila, Itália (2009)

O sismo de Áquila de 2009 foi um sismo de 6,3 graus na escala de magnitude de momento sísmico, segundo o United States Geological Survey (6.7 graus na escala de Richter) registado em 6 de abril de 2009 na zona central da península Itálica. O epicentro foi sob a cidade de Áquila, região de Abruzos. Em Roma, a sua magnitude foi de 4,6 graus Richter. O sismo deixou 309 mortos, cerca de 1000 feridos, 15 desaparecidos e centenas de edificações total ou parcialmente destruídas, sobretudo na cidade de Áquila, mas também em outras localidades próximas, como Onna.



PERCEIVED SHAKING	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
POTENTIAL DAMAGE	none	none	none	Very light	Light	Moderate	Moderate/Heavy	Heavy	Very Heavy
PEAK ACC.(%g)	<.17	.17-1.4	1.4-3.9	3.9-9.2	9.2-18	18-34	34-65	65-124	>124
PEAK VEL.(cm/s)	<0.1	0.1-1.1	1.1-3.4	3.4-8.1	8.1-16	16-31	31-60	60-116	>116
INSTRUMENTAL INTENSITY	I	II-III	IV	V	VI	VII	VIII	IX	X+

Fonte:https://pt.wikipedia.org/wiki/Sismo_de_%C3%81quila_de_2009

Terremoto em L'Aquila, Itália (2009)

- Por não prevenirem a população local a respeito do risco do terremoto, apesar dos quase 200 tremores de menor intensidade, os geólogos responsáveis pela análise sísmica foram considerados culpados pela morte de 309 pessoas, e presos.
- Italy scientists on trial over L'Aquila earthquake - http://www.bbc.com/news/world-europe-14981921
- Can we predict when and where quakes will strike? - http://www.bbc.com/news/science-environment-14991654
- Jailing of Italian seismologists leaves scientific community in shock - https://www.theguardian.com/world/2012/oct/23/jailing-italian-seismologists-scientific-community

Italy scientists on trial over L'Aquila earthquake

The 6.3 magnitude quake devastated the city and killed 309 people. Prosecutors allege the defendants gave a falsely reassuring statement before the quake after studying hundreds of tremors that had shaken the city. *The defence argues that there is no way to predict major earthquakes even in a seismically active area.* The prosecutors accuse the seven of "*negligence and imprudence... of having provided an approximate, generic and ineffective assessment of seismic activity risks as well as incomplete, imprecise and contradictory information*". As the trial opened, L'Aquila prosecutor Alfredo Rossini told reporters: "We simply want justice”. The defendants face up to 15 years in jail. Lawyers for civil plaintiffs - who include the local council - are seeking damages of 50m euros (£45m). The civil portion of the case will be heard alongside the criminal case. Only one of the seven defendants - who include some of Italy's most distinguished geophysicists and members of the country's civil protection agency - was present on the opening day of the trial, which has now been adjourned until 1 October. "I thought it was important to be here because this is my land, and I also wanted to underline the professionalism and the quality of the other public officials," said Bernardo De Bernardinis, former vice-president of the Civil Protection Agency's technical department. "I am from Abruzzo and I owe it to the people of this area”.
http://www.bbc.com/news/world-europe-14981921