Métodos de Previsão

Parte I: Métodos determinísticos

Tópicos e referência básica

1. Métodos determinísticos
   - Previsão e erro de previsão
   - Decomposição de uma série: tendência, sazonalidade, ciclo, …
   - Alisamento exponencial: simples, duplo, Holt e Holt-Winters
2. Métodos estocásticos
   - Função de autocorrelação
   - Estacionaridade
   - Processos estacionários: ruido branco, AR(p), MA(p)
   - Processos não estacionários: ARIMA, SARIMA
   - Modelos ADL, VAR
   - Quebras de estrutura

Referência fundamental: Caiado, J. (2016), Métodos de Previsão em Gestão, Edições Sílabo (2ªed)

Conceito

Métodos de previsão: baseiam-se em séries temporais

Uma **série temporal** consiste num conjunto de observações de uma variável, feitas em períodos sucessivos de tempo, durante um determinado intervalo e representa-se por Y_t, $t = 1, 2, \ldots, T$

Exemplos:

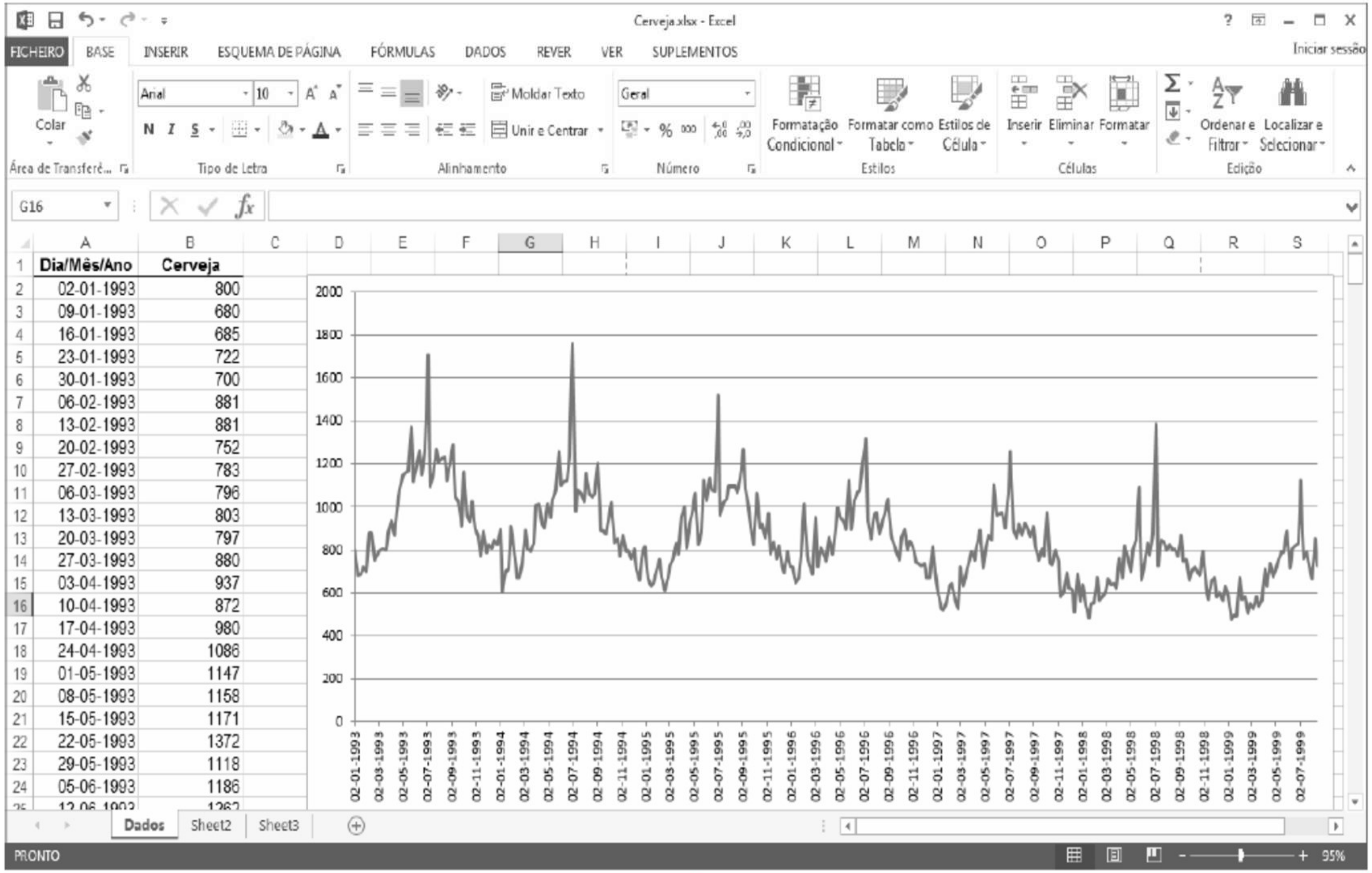
- cotações diárias das acções
- vendas semanais de um dado produto financeiro
- número mensal de dormidas na hotelaria
- despesas públicas trimestrais do país

A representação gráfica de uma série temporal designa-se por **cronograma** e constitui o ponto de partida para a sua análise.

Conceito

Exemplo de séries em excel e respectivos cronogramas (Caiado, 2016)

Figura 2.1. Cronograma da série de vendas de cerveja nos Estados Unidos



Dia/Mês/Ano	Cerveja
02-01-1993	800
09-01-1993	680
16-01-1993	685
23-01-1993	722
30-01-1993	700
06-02-1993	881
13-02-1993	881
20-02-1993	752
27-02-1993	783
06-03-1993	796
13-03-1993	803
20-03-1993	797
27-03-1993	880
03-04-1993	937
10-04-1993	872
17-04-1993	980
24-04-1993	1086
01-05-1993	1147
08-05-1993	1158
15-05-1993	1171
22-05-1993	1372
29-05-1993	1118
05-06-1993	1186

Conceito

Exemplo de séries em excel e respectivos cronogramas (Caiado, 2016)

Figura 2.2. Cronograma do preço de cotação das ações da Microsoft

Objectivo geral da unidade

Examinar procedimentos baseados em valores presentes e passados da série para prever os seus valores futuros

Em geral, utilizando as observações disponíveis $y_1, y_2, \dots y_T$, pretende-se prever os valores futuros desconhecidos $y_{T+1}, y_{T+2}, \dots$:

$$\boxed{y_1 \quad y_2 \quad \dots y_T} \boxed{y_{T+1} = ? \quad y_{T+2} = ? \dots}$$

Existe uma grande variedade de métodos de previsão. A escolha de um método em particular dependerá dos dados disponíveis, da natureza da variável de interesse e dos objectivos do estudo

Além disso, pode-se decidir usar apenas as observações mais recentes para fazer previsão, ou todos os valores passados, dando igual peso a todos os valores. Num caso intermédio, usam-se todos os valores passados, aumentando o peso quanto mais esses valores se aproximam de t

Objectivos especificos

- **Descrição**: análise descritiva da série (média, desvio padrão, mínimo, máximo, ...), construção do cronograma da série e caracterização do seu andamento geral, procurando identificar os pontos de viragem (mudança de estrutura) e eventuais observações anómalas (outliers)
- **Modelação**: construção de modelos que permitam explicar o comportamento da série no período observado
- **Previsão**: prever a evolução futura da série com base exclusivamente no seu passado (modelos univariados ou não-causais) ou com base no comportamento passado de outras variáveis (modelos multivariados)
- **Controlo**: procurar modificar o comportamento futuro do processo através do ajustamento de variáveis controláveis

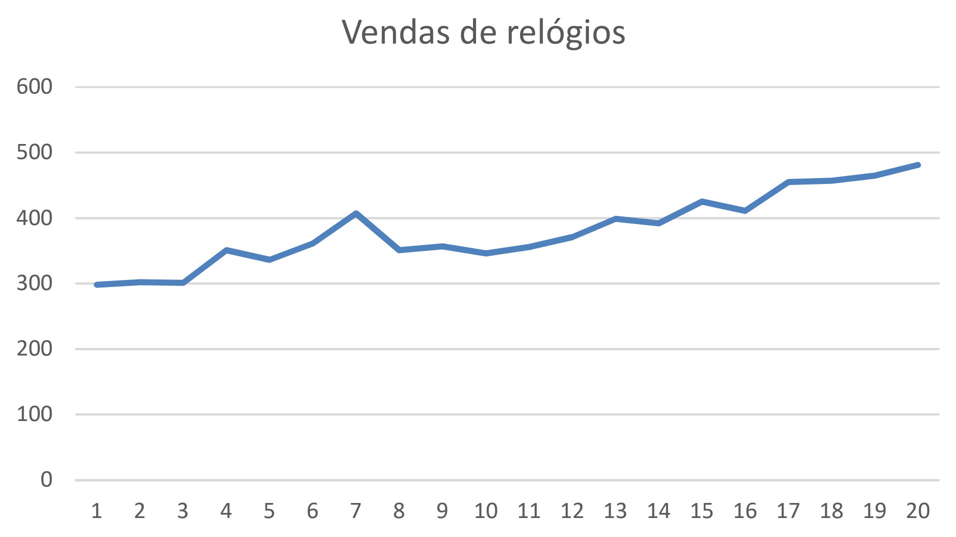
Exemplo: numa linha de fabrico e montagem de automóveis, prever o número de viaturas produzidas com base nas matérias-primas e mão-de-obra utilizadas

Previsão e erro de previsão

Uma fase prévia ao desenvolvimento dos modelos de previsão é a descrição detalhada dos dados.

Segue-se um exemplo respeitante a vendas de relógios, ver Caiado (2016).

Começa-se por apresentar o cronograma e apresentam-se também as principais estatísticas descritivas, calculadas em Excel.



Previsão e erro de previsão

Estatística descritiva

Vendas	
Média	381,1
Erro-padrão	12,41241
Mediana	366
Moda	351
Desvio-padrão	55,50998
Variância da amostra	3081,358
Curtose	-0,85006
Assimetria	0,257442
Intervalo	183
Mínimo	298
Máximo	481
Soma	7622
Contagem	20

Com dados mensais ou trimestrais, poderia considerar-se fazer também a descrição por ano ou por mês ou por trimestre, testar diferenças de médias, etc.

Previsão e erro de previsão

Para uma observação *y* no período t, y_t pode obter-se uma **previsão** $\boldsymbol{\hat{y}_t}$ a partir de um método de previsão. Diferentes métodos de previsão produzem diferentes resultados, havendo necessidade de selecionar o método a usar

Sendo o objectivo de estimação do modelo fazer previsão, um critério razoável para a escolha do modelo é a precisão relativa das suas previsões. Na comparação de modelos de **previsão minimizam-se funções do erro de previsão**

Erro de previsão: diferença entre a observação y_t e a previsão $\hat{y}_t$ obtida pelo modelo com base nas observações passadas

$$e_t = y_t - \hat{y}_t$$

- $e_t \simeq 0$ se o modelo descrever a série, restando apenas ligeiras flutuações

aleatórias com causadas por fenómenos não previsíveis.

Previsão e erro de previsão

Medidas de erro de previsão

Seja $\hat{y}_t$ a previsão um passo a frente feita para t com origem em informação disponível até t-1. Consideram-se as seguintes medidas para os momentos 1, 2, ...m

1) Erro quadrático médio

$$EQM = \frac{1}{m}\sum_{t=1}^{m} e_t^2$$

- A raíz quadrada deste valor, REQM, pode ser utilizada como estimativa do desvio padrão do erro de previsão a um passo, quando representado por e_t

Previsão e erro de previsão

2) Erro absoluto médio

$$EAM = \frac{1}{m}\sum_{t=1}^{m}|e_t|$$

- Mais robusto a erros extremos do que o EQM
- Tal como o EQM, apenas permite comparar métodos de previsão aplicados à mesma série (vendas de automóveis não são comparáveis com vendas de produtos financeiros) e também não é adequado para comparar séries com diferentes periodicidades (vendas diárias, mensais e trimestrais de óculos graduados)

Previsão e erro de previsão

3) Erro percentual absoluto médio

$$EPAM = \frac{1}{m}\sum_{t=1}^{m}\left|\frac{e_t}{Y_t}\right|100$$

- Permite avaliar a precisão de um dado método relativamente a séries diferentes (não é afectado pela unidade de medida)
- Problema: não é definido para zeros e toma valores extremos para Y_t pequeno

4) Erro percentual absoluto médio simétrico

$$EPAMS = \frac{1}{m}\sum_{t=1}^{m}\left|\frac{e_t}{Y_t + \hat{Y}_t}\right|200$$

- Também pode tomar valores extremos: quando Y_t é pequeno, $\hat{Y}_t$ também o será

Previsão e erro de previsão

5) Erro escalado absoluto médio

$$EEAM = \frac{1}{m}\sum_{t=1}^{m}\left|\frac{e_t}{\frac{1}{m-1}\sum_{t=2}^{m}|Y_t - Y_{t-1}|}\right|100$$

- Proposto por Hyndman e Koehler (2006) para evitar os valores extremos das medidas anteriores
- O indicador compara e_t com o erro absoluto médio da previsão *naïve* (na ausência de informação adicional, pode utilizar-se como previsão para o momento *t* o valor observado no momento *t*-1: $\hat{Y}_t = Y_{t-1}$
- Necessita de ajustamento em séries com sazonalidade

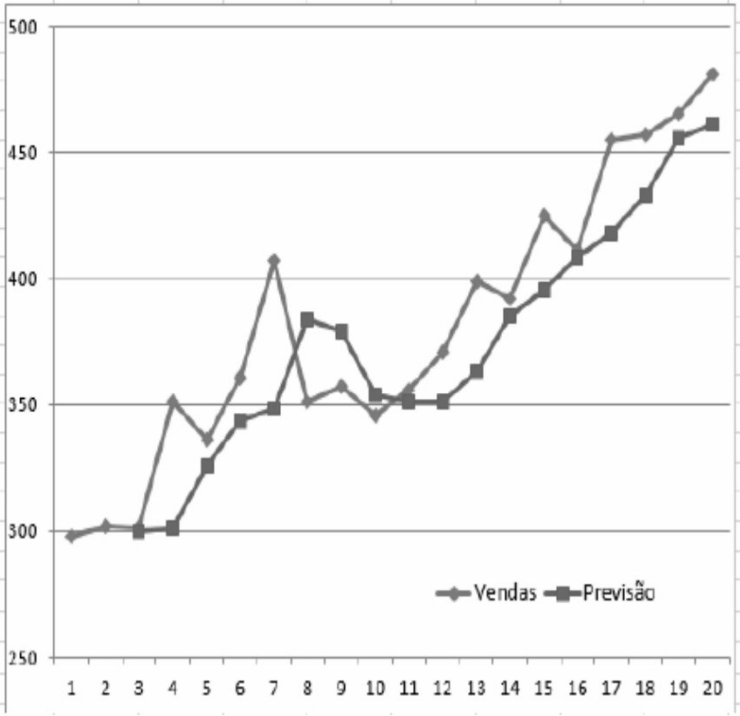
Previsão e erro de previsão

Ilustração: vendas de relógios

Objectivo: previsão 1 passo à frente baseada na média dos dois valores anteriores (Caiado, 2016)

K16 =ABS(D16/MÉDIA($F$5:$F$23))

	A	B	C	D	E	F	G	H	I	J	K
1						Desvio	Erro	Erro	Erro absoluto	Erro absoluto	Erro escalonado
2		Vendas	Previsão	Erro	Desvio	absoluto	quadrático	absoluto	percentual	perc. simétrico	absoluto
3	t	Yt	Pt	Yt - Pt	$Y_t - Y_{t-1}$	$\|Y_t - Y_{t-1}\|$	(Yt - Pt)^2	|Yt - Pt|	|(Yt - Pt)/Yt|*100	200*|Yt-Pt|/(Yt+Pt)	|qt|
4	1	298							=abs(D6/B6)*100	=200*abs(D6)/(B6+C6)	=abs(D6/MÉDIA($F$5:$F$23))
5	2	302			4	4					
6	3	301	300,00	1,00	-1	1	1,00	1,00	0,33	0,33	0,05
7	4	351	301,50	49,50	50	50	2450,25	49,50	14,10	15,17	2,41
8	5	336	326,00	10,00	-15	15	100,00	10,00	2,98	3,02	0,49
9	6	361	343,50	17,50	25	25	306,25	17,50	4,85	4,97	0,85
10	7	407	348,50	58,50	46	46	3422,25	58,50	14,37	15,49	2,84
11	8	351	384,00	-33,00	-56	56	1089,00	33,00	9,40	8,98	1,60
12	9	357	379,00	-22,00	6	6	484,00	22,00	6,16	5,98	1,07
13	10	346	354,00	-8,00	-11	11	64,00	8,00	2,31	2,29	0,39
14	11	356	351,50	4,50	10	10	20,25	4,50	1,26	1,27	0,22
15	12	371	351,00	20,00	15	15	400,00	20,00	5,39	5,54	0,97
16	13	399	363,50	35,50	28	28	1260,25	35,50	8,90	9,31	1,73
17	14	392	385,00	7,00	-7	7	49,00	7,00	1,79	1,80	0,34
18	15	425	395,50	29,50	33	33	870,25	29,50	6,94	7,19	1,43
19	16	411	408,50	2,50	-14	14	6,25	2,50	0,61	0,61	0,12
20	17	455	418,00	37,00	44	44	1369,00	37,00	8,13	8,48	1,80
21	18	457	433,00	24,00	2	2	576,00	24,00	5,25	5,39	1,17
22	19	465	456,00	9,00	8	8	81,00	9,00	1,94	1,95	0,44
23	20	481	461,00	20,00	16	16	400,00	20,00	4,16	4,25	0,97
24					Soma	391,00	12948,75	388,50	98,87	102,02	18,88
25					Média	20,58	719,38	21,58	5,49	5,67	1,05
26							EQM	EAM	EPAM	EPAMS	EEAM



Previsão e erro de previsão

Intervalo de previsão – previsão 1 passo à frente

Admitindo que os erros de previsão têm distribuição aproximadamente normal de média zero, o intervalo de previsão para cada instante de tempo, a $(1-\alpha)100\%$ de confiança, é

$$\left(\hat{Y}_t - z^{\alpha/2}\sqrt{EQM}; \hat{Y}_t + z^{\alpha/2}\sqrt{EQM}\right)$$

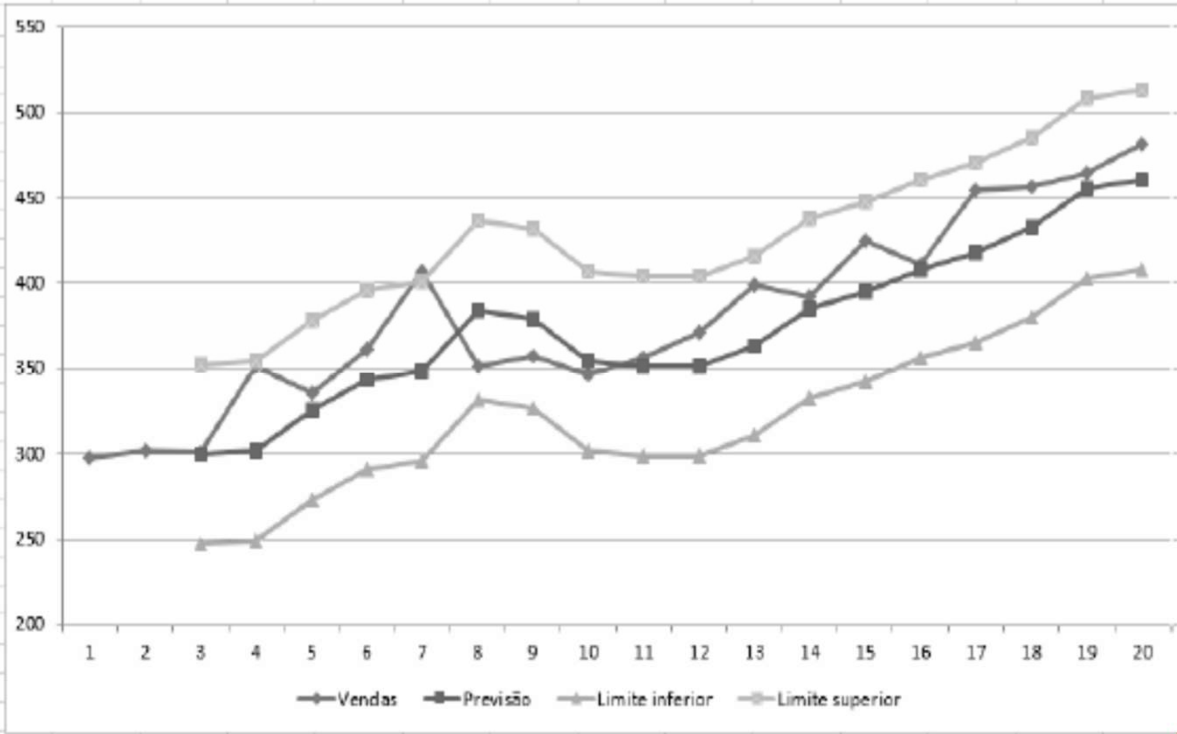
- Usa como estimativa da variância do erro de previsão a 1 passo à frente o EQM
- $z^{\alpha/2}$ para 90%, 95% e 99% é dado por, respectivamente, 1.645, 1.96, 2.576

Previsão e erro de previsão

Intervalo de previsão – previsão 1 passo à frente

- Dentro da amostra (de t=3 a t=20): $\hat{Y}_t \pm 1{,}96\sqrt{719{,}38}$
- Fora da amostra
  - t=21: usa observações t=19 e t=20: $473 \pm 1{,}96\sqrt{719{,}38} \rightarrow (420{,}43; 525{,}57)$
  - previsão passo a passo: t=22,23: usa observações e previsões:
    - $(481 + 473)/2 \pm 1{,}96\sqrt{719{,}38} \rightarrow (424{,}43; 529{,}57)$
    - $(473 + 477)/2 \pm 1{,}96\sqrt{719{,}38} \rightarrow (422{,}43; 527{,}57)$

		Intervalo de previsão a 95% (z=1,96)		
Periodo	Vendas	Previsão	Limite inferior	Limite superior
t	Yt	Pt	Pt-z*(EQM^0,5)	Pt+z*(EQM^0,5)
1	298		+C6-$D$29*($D$28)^0,5	
2	302			
3	301	300,00	247,43	352,57
4	351	301,50	248,93	354,07
5	336	326,00	273,43	378,57
6	361	343,50	290,93	396,07
7	407	348,50	295,93	401,07
8	351	384,00	331,43	436,57
9	357	379,00	326,43	431,57
10	346	354,00	301,43	406,57
11	356	351,50	298,93	404,07
12	371	351,00	298,43	403,57
13	399	363,50	310,93	416,07
14	392	385,00	332,43	437,57
15	425	395,50	342,93	448,07
16	411	408,50	355,93	461,07
17	456	418,00	365,43	470,57
18	457	433,00	380,43	485,57
19	465	456,00	403,43	508,57
20	481	461,00	408,43	513,57
21		473,00	420,43	525,57
22		477,00	424,43	529,57
23		475,00	422,43	527,57



Previsão e erro de previsão

Intervalo de previsão – previsão h passos à frente

Utiliza-se quando o objectivo da previsão é um horizonte temporal de médio/longo prazo. Calculam-se **previsões a h passos à frente utilizando informação disponível até ao momento *t***, $\hat{Y}_{t+h}$: $\hat{Y}_{t+1}$, $\hat{Y}_{t+2}$,…
Em termos práticos, a série de previsões "desliza" (h-1) posições para h>1 relativamente à previsão 1 passo à frente

Neste caso, o EQM baseia-se nos erros de previsão a h passos de tempo:

$$EQM_{(h)} = \frac{1}{m-h}\sum_{t=h-1}^{m}\left[Y_t - \hat{Y}_{t+h}\right]^2$$

O intervalo de confiança correspondente é

$$\left(\hat{Y}_{t+h} - z^{\alpha/2}\sqrt{EQM_{(h)}}; \hat{Y}_{t+h} + z^{\alpha/2}\sqrt{EQM_{(h)}}\right)$$

$\hat{Y}_{t+h}$ corresponde em valor à previsão 1 passo à frente, mesmo para h>1

Previsão e erro de previsão

Intervalo de previsão – previsão h passos à frente

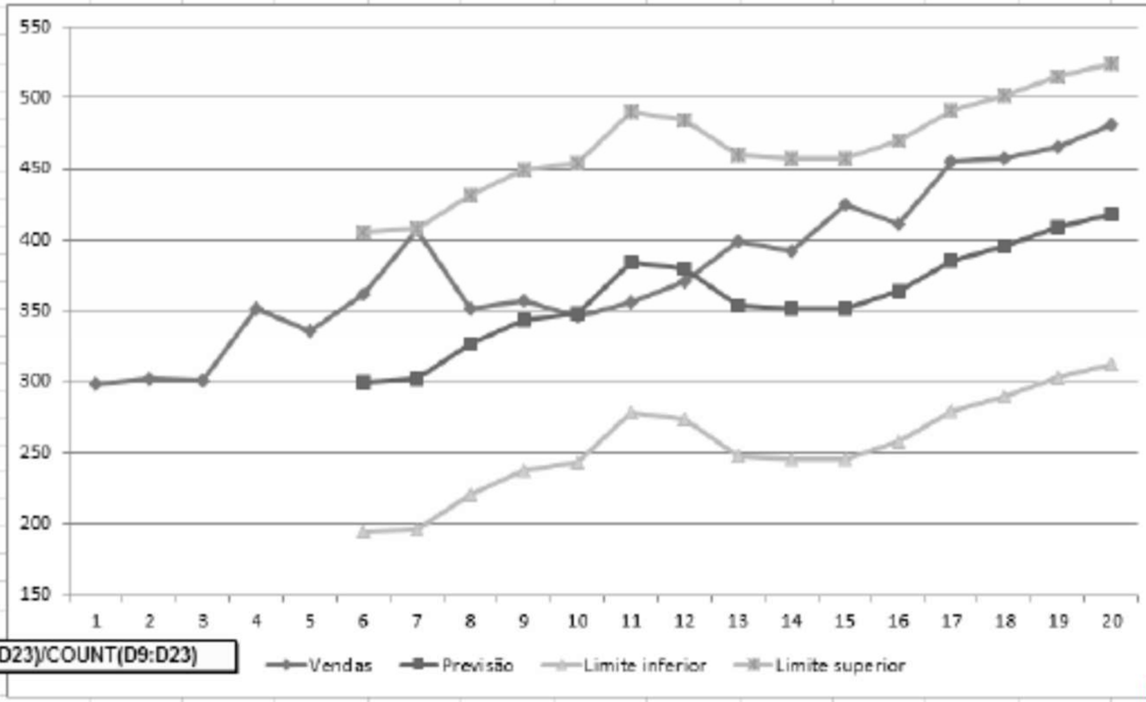
- Exemplo h=4
  - Dentro da amostra (de t=6 a t=20)
    - Na prática corresponde a avançar com a previsão um passo à frente $\hat{Y}_t$ por (h-1) períodos (a primeira previsão $\hat{Y}_3$ corresponde agora a $\hat{Y}_6$)
    - Calcula-se novo EQM=2916.917
    - $\hat{Y}_{t+h} \pm 1{,}96\sqrt{2916.917}$

		Intervalo de previsão a 95% (z=1,96, h=4)			
Período	Vendas	Previsão	Erro	Limite inferior	Limite superior
t	Yt	Pt	Et	P(t+h)-z*(EQMh^0,5)	P(t+h)+z*(EQMh^0,5)
1	298				
2	302				
3	301				
4	351				
5	336				
6	361	300,00	61,00	194,14	405,86
7	407	301,50	105,50	195,64	407,36
8	351	326,00	25,00	220,14	431,86
9	357	343,50	13,50	237,64	449,36
10	346	348,50	-2,50	242,64	454,36
11	356	384,00	-28,00	278,14	489,86
12	371	379,00	-8,00	273,14	484,86
13	399	354,00	45,00	248,14	459,86
14	392	351,50	40,50	245,64	457,36
15	425	351,00	74,00	245,14	456,86
16	411	363,50	47,50	257,64	469,36
17	455	385,00	70,00	279,14	490,86
18	457	395,50	61,50	289,64	501,36
19	465	408,50	56,50	302,64	514,36
20	481	418,00	63,00	312,14	523,86
		EQMh =	2916,92		
		z =	1,96		

=+C9+$E$26*($E$25)^0,5

=+C9-$E$26*($E$25)^0,5

=SUMSQ(D9:D23)/COUNT(D9:D23)



Previsão e erro de previsão

Intervalo de previsão – previsão h passos à frente

- Exemplo h=4
  - Fora da amostra (t=21)
    - Previsão =473, que é a previsão 1 passo à frente
    - Usa-se o EQM=2916.917 de h=4
    - $473 \pm 1{,}96\sqrt{2916.917} \rightarrow (367.1434;\, 578.8566)$

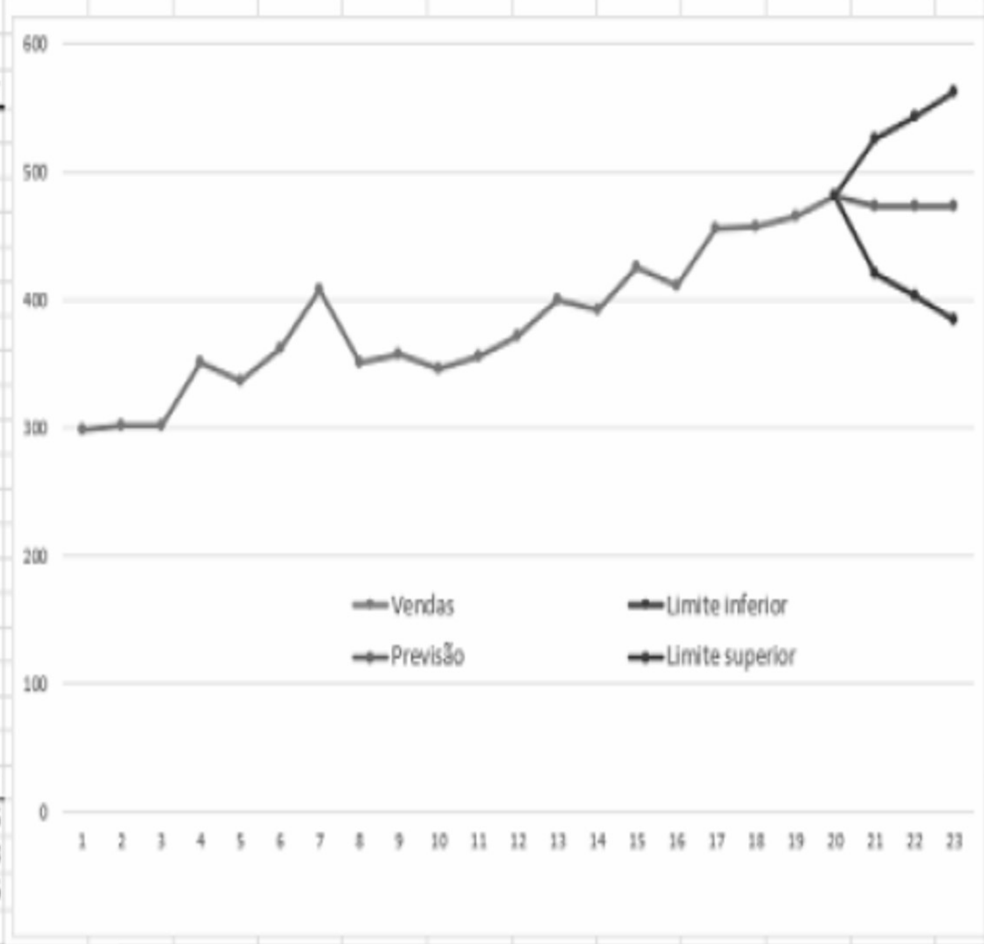
- Fora da amostra: h=2 e h=3
- Exemplo h=2 (fora da amostra: t=21)
    - Previsão =473, que é a previsão 1 passo à frente e EQM=1267.4 de h=2
    - $473 \pm 1{,}96\sqrt{1267{,}4} \rightarrow (403{,}2;\, 542{,}8)$
- Exemplo h=3 (fora da amostra: t=21)
    - Previsão =473, que é a previsão 1 passo à frente e EQM=2036.8 de h=3
    - $473 \pm 1{,}96\sqrt{2036.8} \rightarrow (384.5;\, 561.5)$

Previsão e erro de previsão

Intervalo de previsão – previsão h passos à frente

Previsão e Intervalos de previsão a 95% (h=1,2 e 3)

t	Vendas Yt	Previsão Pt (h=1)	Erro Et	Previsão Pt (h=2)	Erro Et	Previsão Pt (h=3)	Erro Et	Limite inferior	Previsão h=1,2 e 3	Limite superior
1	298									
2	302									
3	301	300	1							
4	351	302	50	300	51,00					
5	336	326	10	302	34,50	300,00	36,00			
6	361	344	18	326	35,00	301,50	59,50			
7	407	349	59	344	63,50	326,00	81,00			
8	351	384	-33	349	2,50	343,50	7,50			
9	357	379	-22	384	-27,00	348,50	8,50			
10	346	354	-8	379	-33,00	384,00	-38,00			
11	356	352	5	354	2,00	379,00	-23,00			
12	371	351	20	352	19,50	354,00	17,00			
13	399	364	36	351	48,00	351,50	47,50			
14	392	385	7	364	28,50	351,00	41,00			
15	425	396	30	385	40,00	363,50	61,50			
16	411	409	3	396	15,50	385,00	26,00			
17	455	418	37	409	46,50	395,50	59,50			
18	457	433	24	418	39,00	408,50	48,50			
19	465	456	9	433	32,00	418,00	47,00			
20	481	461	20	456	25,00	433,00	48,00	481	481	481
21								420,4	473	525,6
22								403,2	473	542,8
23								384,5	473	561,5
		EQM1 =	719,4	EQM2 =	1267,4	EQM3 =	2036,8			



- Intervalo em torno a 473

Previsão e erro de previsão

Amostra de treino e amostra de teste

Amostra de treino (*training set*) destina-se a estimar o modelo

Amostra de teste (*test set*) destina-se a testar a sua qualidade preditiva. Um bom modelo de previsão, não é necessariamente o que melhor descreve a amostra de treino. A avaliação da sua qualidade preditiva faz-se numa amostra de observações que não tenha sido utilizada na sua estimação (amostra de teste)

Previsão e erro de previsão

Amostra de treino e amostra de teste

Dimensão da amostra de teste: igual ou superior ao horizonte temporal pretendido para a previsão futura

- Por exemplo, tem-se uma série de n = 120 observações mensais e pretende-se prever os valores futuros até 24 meses, $\hat{Y}_{t+h}$, h = 1, 2, ..., 24. Usam-se as últimas 24 observações disponíveis, para t = 97, 98, ..., 120, como amostra de teste e as restantes observações, para t = 1, 2, ..., 96, como amostra de treino para estimar o modelo
- No caso de escolher uma amostra de teste de dimensão superior ao número de previsões requeridas (neste caso, 24), aconselha-se a fazê-lo com um número de observações múltiplo do ciclo sazonal (por exemplo, 36 ou 48 observações).

Previsão e erro de previsão

Previsão passo a passo e a múltiplos passos

Previsão passo-a-passo: para a amostra de treino de dimensão T, calculam-se recursivamente as previsões a um passo à frente para o momento $T + 1$, $T + 2$, ... , aumentando sucessivamente a amostra de treino em uma observação

Previsão a *h* passos à frente: utilizam-se as observações até ao instante T para calcular a previsão para o momento $T + h$. Em seguida, utilizam-se observações até $T + 1$ para prever a série no instante $T + h + 1$, e por aí adiante

Em ambos os casos, calculam-se as habituais medidas de avaliação dos erros de previsão (EQM, EAM, EPAM, EPAMS ou EEAM) com base nos erros obtidos na amostra de teste

Previsão e erro de previsão: Exercício

2.1. No quadro seguinte encontram-se os dados da taxa de inflação homóloga (INFL) e da taxa de juro de determinada operação bancária (TJURO) em Portugal (*Fonte*: Banco de Portugal).

Data	INFL	TJURO	Data	INFL	TJURO	Data	INFL	TJURO
2005M05	1,822802	4,89	2007M09	2,081372	6,26	2010M01	0,121556	4,42
2005M06	1,62037	4,74	2007M10	2,57107	6,34	2010M02	0,202696	4,11
2005M07	2,120375	4,75	2007M11	2,797495	6,35	2010M03	0,573152	4,24
2005M08	2,593372	4,92	2007M12	2,676526	6,28	2010M04	0,722166	4,61
2005M09	2,872671	4,92	2008M01	2,840731	6,5	2010M05	1,05517	4,45
2005M10	2,700022	4,87	2008M02	2,861918	6,05	2010M06	1,174935	4,64
2005M11	2,564665	4,99	2008M03	3,095017	6,27	2010M07	1,887935	4,26
2005M12	2,584602	4,76	2008M04	2,484155	6,61	2010M08	1,974284	4,59
2006M01	2,69527	4,95	2008M05	2,786852	6,76	2010M09	1,970493	4,98
2006M02	2,949593	5,16	2008M06	3,350015	6,56	2010M10	2,353298	4,64
2006M03	3,904793	5,07	2008M07	3,042098	6,89	2010M11	2,277078	5,07
2006M04	3,725896	5,12	2008M08	3,015334	6,99	2010M12	2,51737	5,21
2006M05	3,743083	5,25	2008M09	3,104508	7,13	2011M01	3,601781	5,13
2006M06	3,677188	5,09	2008M10	2,384349	7,11	2011M02	3,560231	5,31
2006M07	3,125338	5,44	2008M11	1,380991	6,96	2011M03	4,019196	5,75
2006M08	2,851896	5,37	2008M12	0,781012	6,32	2011M04	4,043019	5,61
2006M09	3,083558	5,5	2009M01	0,253884	6,11	2011M05	3,838504	5,84
2006M10	2,671961	5,26	2009M02	0,192933	5,5	2011M06	3,404467	6,06
2006M11	2,372302	5,65	2009M03	-0,480336	5,26	2011M07	3,151011	6,15
2006M12	2,508808	5,67	2009M04	-0,548628	4,91	2011M08	2,909055	6,51
2007M01	2,570969	5,92	2009M05	-1,171914	4,99	2011M09	3,547716	6,69
2007M02	2,351935	5,88	2009M06	-1,591066	4,7	2011M10	4,193803	6,73
2007M03	2,32239	6,05	2009M07	-1,540755	4,51	2011M11	3,940498	6,85
2007M04	2,741309	5,95	2009M08	-1,328671	4,18	2011M12	3,614576	6,5
2007M05	2,447187	6,01	2009M09	-1,659545	4,37	2012M01	3,505859	6,56
2007M06	2,437748	5,84	2009M10	-1,462978	4,45	2012M02	3,603868	6,66
2007M07	2,380453	5,88	2009M11	-0,590946	4,21	2012M03	3,152634	6,39

Previsão e erro de previsão: Exercício

a) Construa um gráfico incluindo ambas as séries

b) Apresente a estatística descritiva das séries

c) Construa e interprete o coeficiente de correlação entre as séries

d) Escolha o período 2010m05-2012m04 como amostra de teste e calcule previsões a 1, 2 e 3 passos à frente da variável TJURO utilizando a média aritmética das últimas 2 observações disponíveis. Com bases nestas previsões:

- i) Calcule o EQM
- ii) Construa intervalos de previsão a 95% para as previsões obtidas.

e) Calcule previsões e intervalos de previsão a 95% da variável TJURO para os instantes futuros 2012m05, 2012m06 e 2012m07. Comente a amplitude dos intervalos obtidos.

Componentes de uma Série Temporal

Considere-se a decomposição de uma série nas suas componentes principais. A decomposição, não sendo um método de previsão, permite explicar o padrão de comportamento da série, facilitando a previsão

Considera-se que o comportamento de uma série resulta de quatro componentes que, sendo não observáveis, se poderão estimar:

- Tendência
- Ciclo
- Sazonalidade
- Movimentos irregulares/aleatórios

Componentes de uma Série Temporal

1) Tendência

Alterações suaves, resultantes de efeitos permanentes do comportamento por um longo período de tempo. Uma série pode exibir uma tendência linear crescente ou decrescente, ou não linear, ou simplesmente não exibir tendência

2) Ciclo

Flutuações repetidas por conjunto de anos em torno à tendência, estando associados às fases de expansão e de recessão dos sistemas económicos. Têm periodicidade pouco definida, sendo de difícil modelação formal. Nomeadamente, nos ciclos longos, poderão ser difíceis de separar da tendência, pelo que se poderá considerar a componente conjunta de tendência-ciclo

Componentes de uma Série Temporal

3) Sazonalidade

Oscilações de subida e de queda nas séries, relativamente à tendência, com periodicidade anual ou infra-anual: repetem-se ao ano, mês, semana, dia ou horário. Decorrem de causas naturais (clima: estações do ano no consumo de água, turismo), medidas administrativas (inicio do ano escolar), ou tradições (vendas no período natalício)

4) Movimentos irregulares

Oscilações de subida e descida que ocorrem sem padrão definido, para além do efeito de tendência, ciclo e sazonalidade. Referem-se a flutuações de período curto, causadas, entre outros motivos, por eventos políticos e oscilações climáticas imprevisíveis. É o comportamento não explicado pelas três componentes anteriores

Componentes de uma Série Temporal

Modelo de decomposição

$$Y_t = f(T_t, C_t, S_t, E_t)$$

onde T_t é a tendência, C_t é o ciclo, S_t é a sazonalidade e E_t é a componente irregular. Em casos onde a tendência não é separada do ciclo tem-se $Y_t = f(TC_t, S_t, E_t)$, onde TC_t representa a componente tendência-ciclo

Tipos de modelos:

- Modelo aditivo
- Modelo multiplicativo

Componentes de uma Série Temporal

Modelo aditivo

$$Y_t = T_t + C_t + S_t + E_t$$

- Assume a independência dos efeitos, considerando a sua soma. Todos os efeitos estão expressos na mesma unidade

Modelo multiplicativo

$$Y_t = T_t * C_t * S_t * E_t$$
$$ln(Y_t) = ln(T_t) + ln(C_t) + ln(S_t) + ln(E_t)$$

- Permite a dependência dos efeitos, por exemplo a amplitude da sazonalidade aumenta/diminui ao longo do tempo
- A tendência está expressa na mesma unidade da série e as restantes componentes vêm expressas em percentagem.
- Aplicável apenas a séries com valores não negativos

Componentes de uma Série Temporal

Tendência

A tendência de uma série analisa-se para:

- Descrição do passado
- Previsão do futuro, assumindo a manutenção do mesmo comportamento
- Eliminação desta componente, de forma a que possa estudar-se a sazonalidade e o ciclo

Tipos de tendência:

- linear (crescente ou decrescente)
- não linear
- ausente

A tendência analisa-se por médias móveis (MM) ou por modelos de regressão (nesta fase, estudar-se-á por MM e só mais tarde por modelos de regressão)

Componentes de uma Série Temporal
Tendência: método das médias móveis

Objectivo: produzir uma nova série temporal, alisada ou suavizada, a partir da série original, a qual resulta de médias sucessivas dos valores observados

Média móvel simples de ordem K, MM(k):

- k impar

$$MM(k) = \frac{1}{k}\left(Y_{t-(k-1)/2} + \cdots Y_t + Y_{t+(k-1)/2}\right)$$

  - Exemplo: $MM(3) = \frac{Y_{t-1}+Y_t+Y_{t+1}}{3}$

- k par

$$MM(k) = \frac{1}{2}\left(\frac{1}{k}\left(Y_{t-(k-2)/2} + \cdots Y_t + Y_{t+k/2}\right) + \frac{1}{k}\left(Y_{t-k/2} + \cdots Y_t + Y_{t+(k-2)/2}\right)\right)$$

$$= \frac{1}{k}\left(0.5Y_{t-(k-2)/2} + \cdots Y_t + 0.5Y_{t+(k-2)/2}\right)$$ (usa k+1 termos)

  - Exemplo: $MM(4) = \frac{0.5Y_{t-2}+Y_{t-1}+Y_t+Y_{t+1}+0.5Y_{t+2}}{4}$
  - Também designada de média dupla

Componentes de uma Série Temporal
Tendência: método das médias móveis

- A escolha de k é subjectiva, sendo que quanto maior k, maior o alisamento. Uma solução pode ser a minimização do EQM, EAM, … (se a série for estável, a melhor escolha para a período amostral, será a escolha razoável para previsão)

- Recomenda-se a escolha de k elevado para séries razoavelmente constantes e k pequeno em séries mais instáveis ou quando se pretende captar as alterações mais recentes da série

Previsão usando médias móveis

- Baseia-se na média aritmética das *k* observações presentes e passadas mais recentes:

$$Pk_{t+h} = \frac{Y_t + Y_{t-1} + \cdots + Y_{t-k+1}}{k}, h = 1,2\ \ldots$$

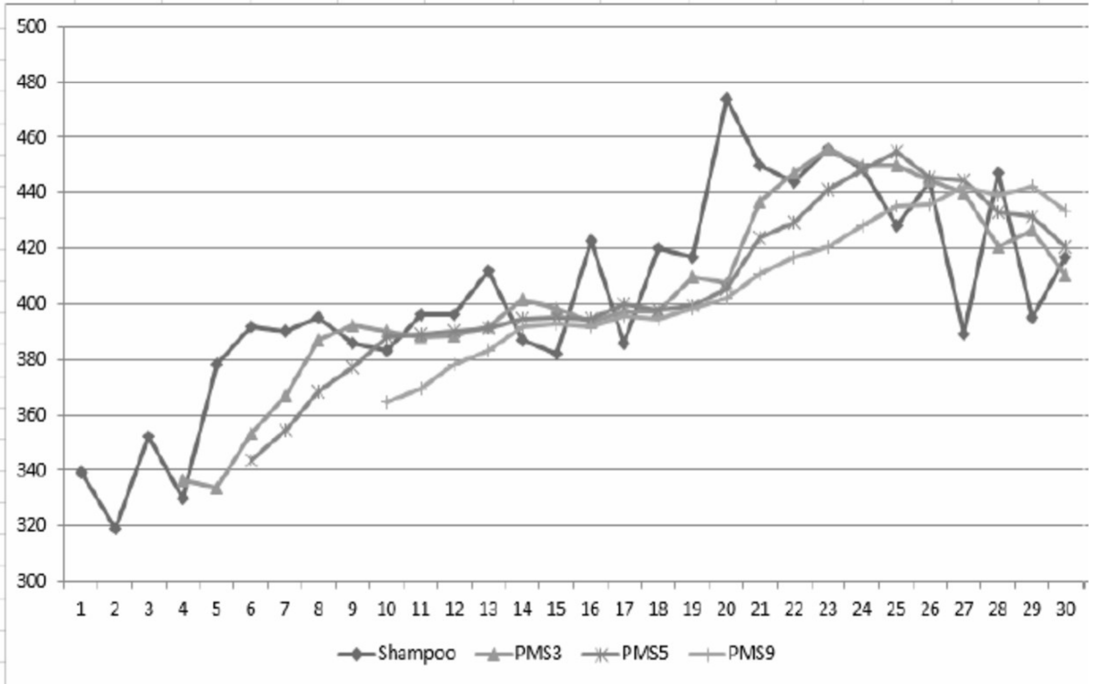
  - Na prática, corresponde a “deslisar” a série de MM(k), de forma a iniciar-se na posição k+1. Isto corresponde a fazer uma média com a observação em t e com os k-1 valores anteriores
  - Exemplo 1 passo à frente: $P3_{t+1} = \frac{Y_t+Y_{t-1}+Y_{t-2}}{3}$, com $P3_4 = \frac{Y_3+Y_2+Y_1}{3}$, $P3_5 = \frac{Y_4+Y_3+Y_2}{3}$
  - Exemplo 2 passos à frente: $P3_{t+2} = \frac{Y_t+Y_{t-1}+Y_{t-2}}{3}$, com $P3_4 = \frac{Y_3+Y_2+Y_1}{3}$, $P3_6 = \frac{Y_4+Y_3+Y_2}{3}$
- Pode-se considerar ponderar (WMM), de forma a dar mais peso a observações recentes (escolha dos pesos, em concreto, subjectiva):

$$P3_{t+1} = \frac{0.6Y_t+0.4Y_{t-1}+0.1Y_{t-2}}{3}$$

Ilustração: vendas de shampoo (Caiado, 2016)

H11 =MÉDIA(B2:B10)

t	Shampoo	MM3	PMS3	MM5	PMS5	MM9	PMS9
1	339						
2	319	336,7					
3	352	333,7		343,6			
4	330	353,3	336,7	354,2			
5	378	366,7	333,7	368,4		364,6	
6	392	386,7	353,3	377,0	343,6	369,4	
7	390	392,3	366,7	388,2	354,2	378,0	
8	395	390,3	386,7	389,2	368,4	382,9	
9	386	388,0	392,3	390,0	377,0	392,0	
10	383	388,3	390,3	391,2	388,2	393,0	364,6
11	396	391,7	388,0	394,6	389,2	391,9	369,4
12	396	401,3	388,3	394,8	390,0	395,6	378,0
13	412	398,3	391,7	394,6	391,2	394,6	382,9
14	387	393,7	401,3	400,0	394,6	398,3	392,0
15	382	397,3	398,3	398,0	394,8	402,1	393,0
16	423	397,0	393,7	399,6	394,6	410,8	391,9
17	386	409,7	397,3	405,6	400,0	416,8	395,6
18	420	407,7	397,0	424,0	398,0	420,3	394,6
19	417	437,0	409,7	429,4	399,6	428,0	398,3
20	474	447,0	407,7	441,0	405,6	435,4	402,1
21	450	456,0	437,0	448,2	424,0	436,0	410,8



Médias simples | MM3 | MM9 | Previsões | Médias duplas

Componentes de uma Série Temporal
Tendência: método das médias móveis

Ilustração: vendas de shampoo

- Confirma-se um aumento do alisamento quando k aumenta

Componentes de uma Série Temporal
Tendência método das médias móveis

Ilustração: vendas de shampoo

Q7 =SOMARQUAD(D5:D31)/CONTAR(D5:D31)

t	Shampoo	PMS3	Erro	EA	EPA	PMS5	Erro	EA	EPA	PMS9	Erro	EA	EPA
1	339												
2	319												
3	352												
4	330	336,7	-6,7	6,7	2,0								
5	378	333,7	44,3	44,3	11,7								
6	392	353,3	38,7	38,7	9,9	343,6	48,4	48,4	12,3				
7	390	366,7	23,3	23,3	6,0	354,2	35,8	35,8	9,2				
8	395	386,7	8,3	8,3	2,1	368,4	26,6	26,6	6,7				
9	386	392,3	-6,3	6,3	1,6	377,0	9,0	9,0	2,3				
10	383	390,3	-7,3	7,3	1,9	388,2	-5,2	5,2	1,4	364,6	18,4	18,4	4,8
11	396	388,0	8,0	8,0	2,0	389,2	6,8	6,8	1,7	369,4	26,6	26,6	6,7
12	396	388,3	7,7	7,7	1,9	390,0	6,0	6,0	1,5	378,0	18,0	18,0	4,5
13	412	391,7	20,3	20,3	4,9	391,2	20,8	20,8	5,0	382,9	29,1	29,1	7,1
14	387	401,3	-14,3	14,3	3,7	394,6	-7,6	7,6	2,0	392,0	-5,0	5,0	1,3
15	382	398,3	-16,3	16,3	4,3	394,8	-12,8	12,8	3,4	393,0	-11,0	11,0	2,9
16	423	393,7	29,3	29,3	6,9	394,6	28,4	28,4	6,7	391,9	31,1	31,1	7,4
17	386	397,3	-11,3	11,3	2,9	400,0	-14,0	14,0	3,6	395,6	-9,6	9,6	2,5
18	420	397,0	23,0	23,0	5,5	398,0	22,0	22,0	5,2	394,6	25,4	25,4	6,1
19	417	409,7	7,3	7,3	1,8	399,6	17,4	17,4	4,2	398,3	18,7	18,7	4,5
20	474	407,7	66,3	66,3	14,0	405,6	68,4	68,4	14,4	402,1	71,9	71,9	15,2
21	450	437,0	13,0	13,0	2,9	424,0	26,0	26,0	5,8	410,8	39,2	39,2	8,7

Medidas dos erros de previsão

	PMS3	PMS5	PMS9
EQM	601,30	721,76	911,08
EAM	18,31	20,90	25,17
EPAM	4,48	5,05	5,96

Legenda:
Erro = Erro de previsão
EA = Erro absoluto
EPA = Erro percentual absoluto
EQM = Erro quadrático médio
EAM = Erro absoluto médio
EPAM = Erro percentual absoluto médio
PMS3 = Previsões com médias simples de ordem 3
PMS5 = Previsões com médias simples de ordem 5
PMS9 = Previsões com médias simples de ordem 9

Médias simples | MM3 | MM9 | Previsões | Médias duplas

- A previsão com k=3 tem menores medidas de erro de previsão

Componentes de uma Série Temporal
Tendência: método das médias móveis

Pode-se obter a série sem tendência (ou sem tendência ciclo, caso estas duas componentes se tratem agregadas), subtraindo ou dividindo Y_t por MM(k), conforme o modelo assumido seja o aditivo ou o multiplicativo, respectivamente:

- $Y_t = T_t + C_t + S_t + E_t \rightarrow Y_t - MM_t = C_t + S_t + E_t$

  $\rightarrow Y_t - MM_t = S_t + E_t$ (versão tendência-ciclo)

- $Y_t = T_t * C_t * S_t * E_t \rightarrow \frac{Y_t}{MM_t} = C_t * S_t * E_t$

  $\rightarrow \frac{Y_t}{MM_t} = S_t * E_t$(versão tendência-ciclo)

Componentes de uma Série Temporal
Sazonalidade

Objectivo

Estima-se sobretudo para se proceder à dessazonalização (remoção da sazonalidade) da série observada de forma a que as suas outras características se tornem evidentes.

Há vários métodos de estimação da sazonalidade dos quais se estuda de seguida o método clássico de decomposição, na versão aditiva e multiplicativa.

Componentes de uma Série Temporal
Sazonalidade: método tradicional

Passo 1

- Remover da série a componente tendência-cíclica, para isolar as componentes sazonal e irregular, da forma já definida:

$$R_t = Y_t - MM_t = S_t + E_t$$

$$R_t = \frac{Y_t}{MM_t} = S_t * E_t$$

onde MM_t terá ordem k de 12, 4, ou 2, para dados mensais, trimestrais ou semestrais

Passo 2

- Estimar os índices sazonais (I_t) através das médias de R_t para cada mês, trimestre ou semestre, assumido que a componente sazonal é constante de ano para ano. Em seguida, calcular os fatores sazonais (S_t) com base nos quocientes entre os índices sazonais e a sua média aritmética ou geométrica, conforme o modelo seja aditivo ou multiplicativo. Caso de dados mensais

$$S_t = I_t - \left(\sum_{t=1}^{12} I_t \,/12\right) \text{ ou } S_t = I_t / \sqrt[12]{\prod_{i=1}^{12} I_t}$$

Componentes de uma Série Temporal
Sazonalidade: método tradicional

Passo 3

- Construir a série dessazonalizada, subtraindo ou dividindo os valores da série original pelos respetivos fatores sazonais (a série resultante exibe tendência e ciclo)

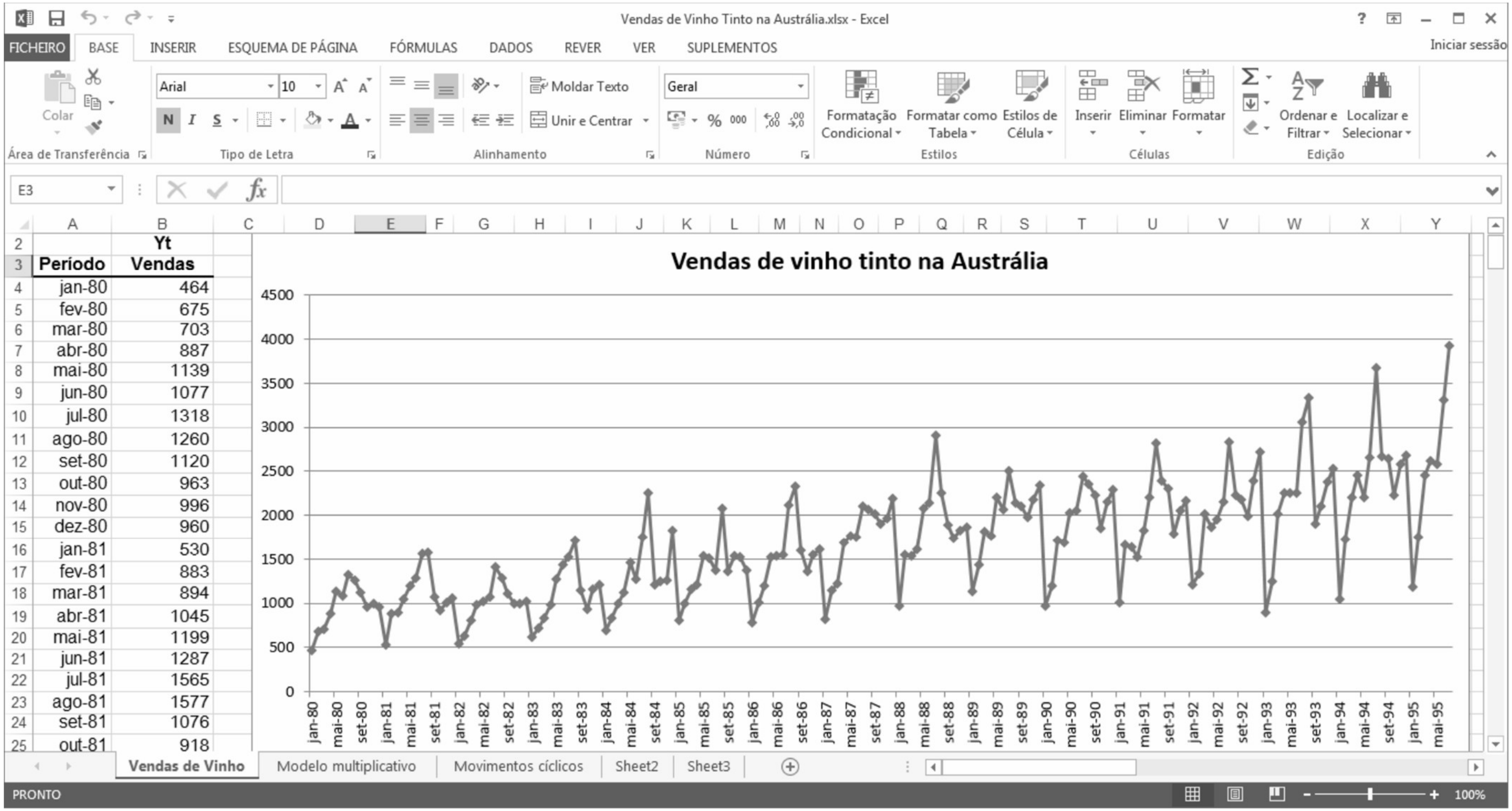
$$Y_t^D = Y_t - S_t$$
$$Y_t^D = \frac{Y_t}{S_t}$$

- E/ou obter as componentes irregular e de ciclo (a série resultante não exibe tendência nem sazonalidade)

$$E_t = Y_t - MM_t - S_t$$
$$E_t = \frac{Y_t}{MM_t * S_t}$$

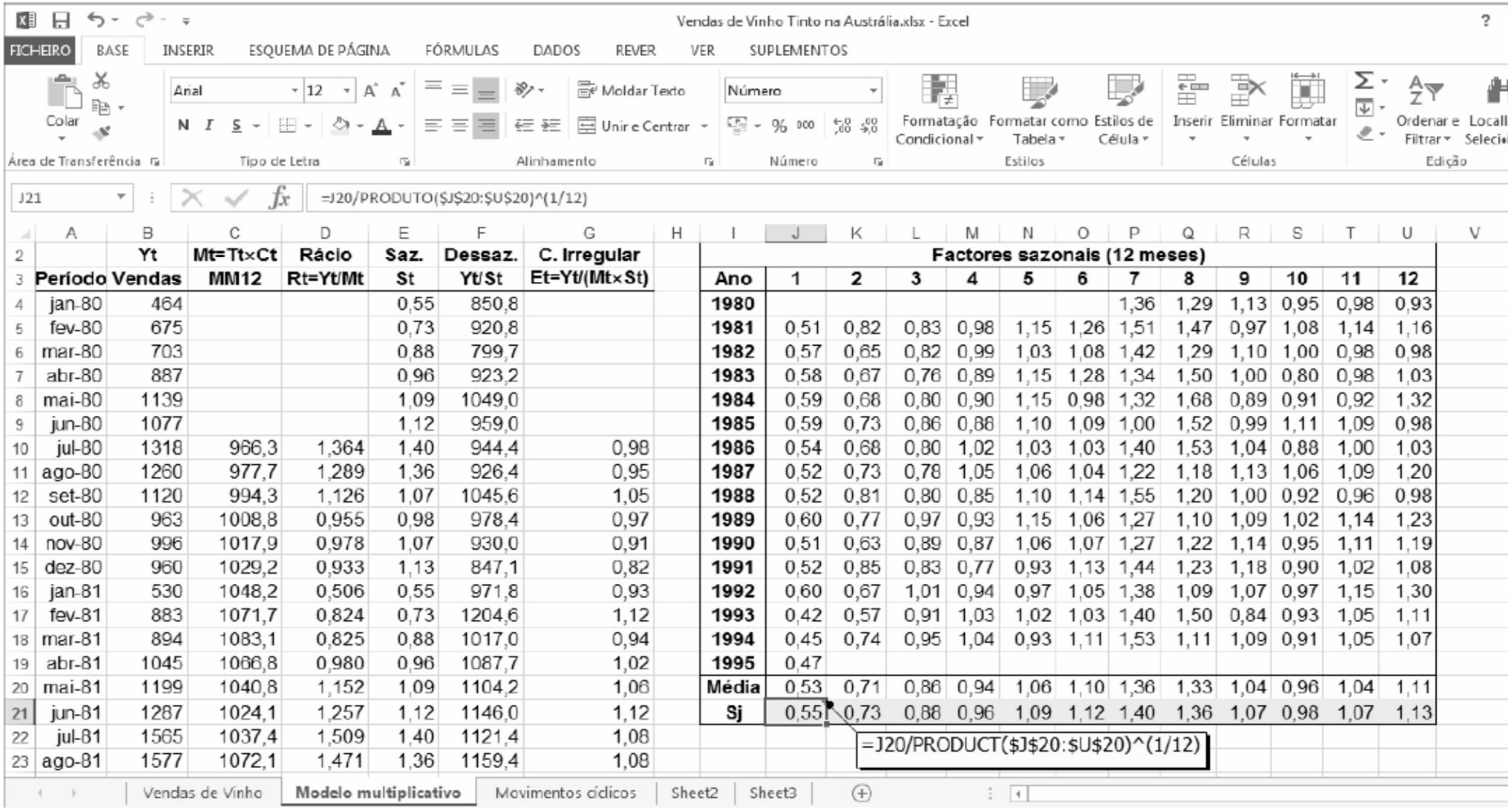
Componentes de uma Série Temporal
Sazonalidade: método tradicional

Ilustração: vendas de vinho tinto, Australia (Caiado, 2016)



Período	Yt Vendas
jan-80	464
fev-80	675
mar-80	703
abr-80	887
mai-80	1139
jun-80	1077
jul-80	1318
ago-80	1260
set-80	1120
out-80	963
nov-80	996
dez-80	960
jan-81	530
fev-81	883
mar-81	894
abr-81	1045
mai-81	1199
jun-81	1287
jul-81	1565
ago-81	1577
set-81	1076
out-81	918

Ilustração: vendas de vinho tinto, Australia (Caiado, 2016)



Vendas de Vinho Tinto na Austrália.xlsx - Excel

J21 =J20/PRODUTO($J$20:$U$20)^(1/12)

Período	Yt Vendas	Mt=Tt×Ct MM12	Rácio Rt=Yt/Mt	Saz. St	Dessaz. Yt/St	C. Irregular Et=Yt/(Mt×St)
jan-80	464			0,55	850,8	
fev-80	675			0,73	920,8	
mar-80	703			0,88	799,7	
abr-80	887			0,96	923,2	
mai-80	1139			1,09	1049,0	
jun-80	1077			1,12	959,0	
jul-80	1318	966,3	1,364	1,40	944,4	0,98
ago-80	1260	977,7	1,289	1,36	926,4	0,95
set-80	1120	994,3	1,126	1,07	1045,6	1,05
out-80	963	1008,8	0,955	0,98	978,4	0,97
nov-80	996	1017,9	0,978	1,07	930,0	0,91
dez-80	960	1029,2	0,933	1,13	847,1	0,82
jan-81	530	1048,2	0,506	0,55	971,8	0,93
fev-81	883	1071,7	0,824	0,73	1204,6	1,12
mar-81	894	1083,1	0,825	0,88	1017,0	0,94
abr-81	1045	1066,8	0,980	0,96	1087,7	1,02
mai-81	1199	1040,8	1,152	1,09	1104,2	1,06
jun-81	1287	1024,1	1,257	1,12	1146,0	1,12
jul-81	1565	1037,4	1,509	1,40	1121,4	1,08
ago-81	1577	1072,1	1,471	1,36	1159,4	1,08

Factores sazonais (12 meses)

Ano	1	2	3	4	5	6	7	8	9	10	11	12
1980							1,36	1,29	1,13	0,95	0,98	0,93
1981	0,51	0,82	0,83	0,98	1,15	1,26	1,51	1,47	0,97	1,08	1,14	1,16
1982	0,57	0,65	0,82	0,99	1,03	1,08	1,42	1,29	1,10	1,00	0,98	0,98
1983	0,58	0,67	0,76	0,89	1,15	1,28	1,34	1,50	1,00	0,80	0,98	1,03
1984	0,59	0,68	0,80	0,90	1,15	0,98	1,32	1,68	0,89	0,91	0,92	1,32
1985	0,59	0,73	0,86	0,88	1,10	1,09	1,00	1,52	0,99	1,11	1,09	0,98
1986	0,54	0,68	0,80	1,02	1,03	1,03	1,40	1,53	1,04	0,88	1,00	1,03
1987	0,52	0,73	0,78	1,05	1,06	1,04	1,22	1,18	1,13	1,06	1,09	1,20
1988	0,52	0,81	0,80	0,85	1,10	1,14	1,55	1,20	1,00	0,92	0,96	0,98
1989	0,60	0,77	0,97	0,93	1,15	1,06	1,27	1,10	1,09	1,02	1,14	1,23
1990	0,51	0,63	0,89	0,87	1,06	1,07	1,27	1,22	1,14	0,95	1,11	1,19
1991	0,52	0,85	0,83	0,77	0,93	1,13	1,44	1,23	1,18	0,90	1,02	1,08
1992	0,60	0,67	1,01	0,94	0,97	1,05	1,38	1,09	1,07	0,97	1,15	1,30
1993	0,42	0,57	0,91	1,03	1,02	1,03	1,40	1,50	0,84	0,93	1,05	1,11
1994	0,45	0,74	0,95	1,04	0,93	1,11	1,53	1,11	1,09	0,91	1,05	1,07
1995	0,47											
Média	0,53	0,71	0,86	0,94	1,06	1,10	1,36	1,33	1,04	0,96	1,04	1,11
Sj	0,55	0,73	0,88	0,96	1,09	1,12	1,40	1,36	1,07	0,98	1,07	1,13

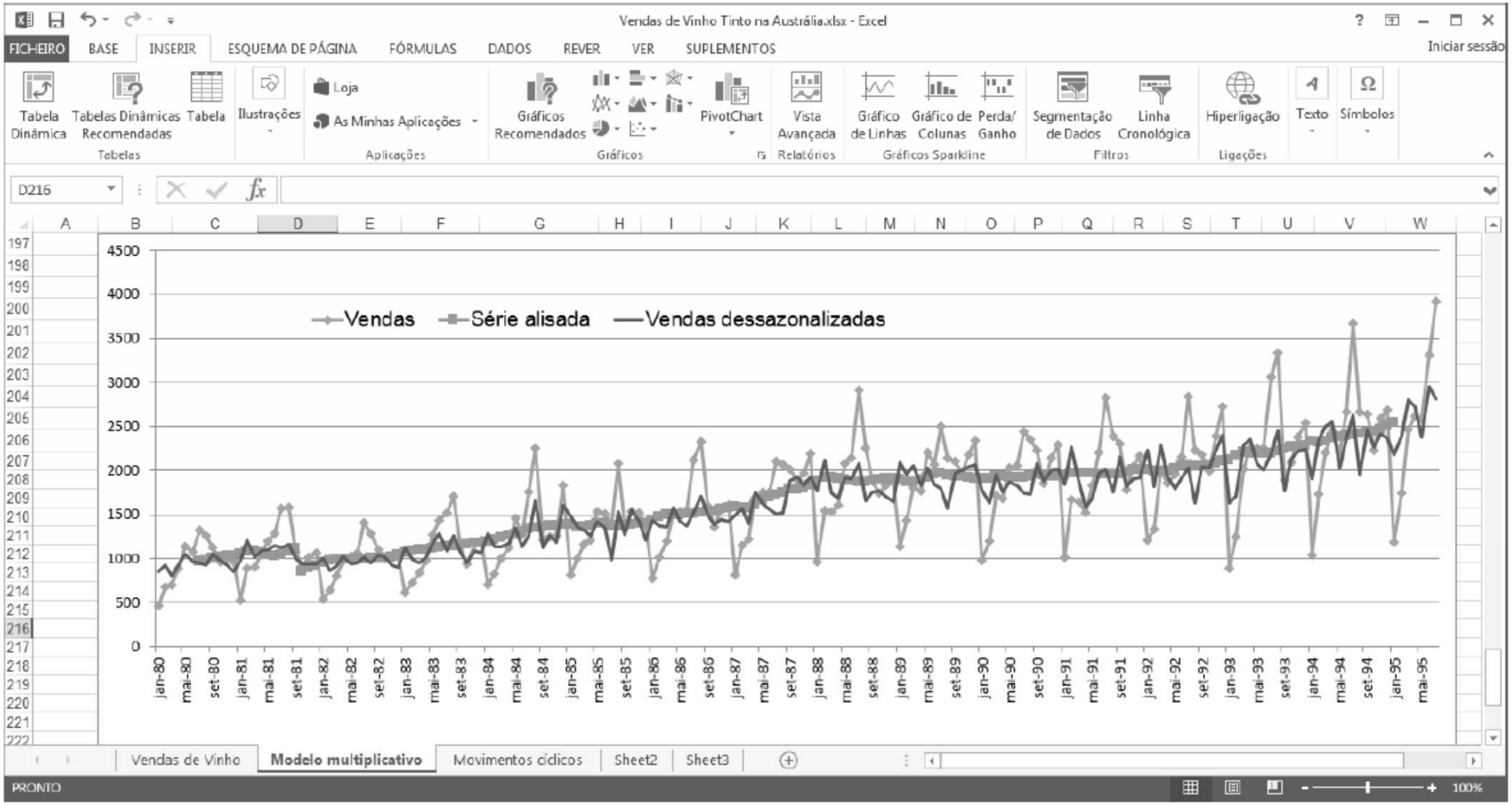
=J20/PRODUCT($J$20:$U$20)^(1/12)

Vendas de Vinho | Modelo multiplicativo | Movimentos cíclicos | Sheet2 | Sheet3

Componentes de uma Série Temporal
Sazonalidade: método tradicional

Ilustração: vendas de vinho tinto, Australia (Caiado, 2016)

Série original, série alisada (sem tendência) e série alisada e dessazonalizada



Limitações

- Perde observações (por exemplo, ao calcular a média móvel de ordem 12, perdem-se as primeiras e as últimas 6 observações)

- Assume que o padrão sazonal não se altera de ano para ano, ao reproduzir os mesmos 4 ou 12 fatores sazonais (dados trimestrais ou mensais) ao longo de todo o período observado. Muitas vezes, sobretudo em séries longas, as oscilações de caráter sazonal não são constantes ao longo do tempo (por exemplo, os consumos de água e de energia podem apresentar diferentes padrões de sazonalidade ao longo do tempo em resultado das alterações climáticas e tecnológicas)

- É muito sensível a outliers nos dados, o que lhe confere pouca robustez.

Componentes de uma Série Temporal
Sazonalidade: método de decomposição census X-12 ARIMA

Assume-se um modelo multiplicativo aplicado a dados mensais.
Primeiros passos:

Passo 1

- Igual ao passo 1 anterior

Passo 2

- Calcular MM(3) triplas para cada um dos meses a partir de R_t, para obter estimativas iniciais para os factores sazonais $\hat{S}_t$

Passo 3

- Igual ao passo 3 anterior para obter E_t

Passo 4

- P 83 Os outliers podem ser identificados e removidos da componente irregular para obter uma estimativa revista da componente irregular

(…)

Componentes de uma Série Temporal

Ciclo

Muitas séries exibem movimentos de expansão e recessão associados aos ciclos económicos

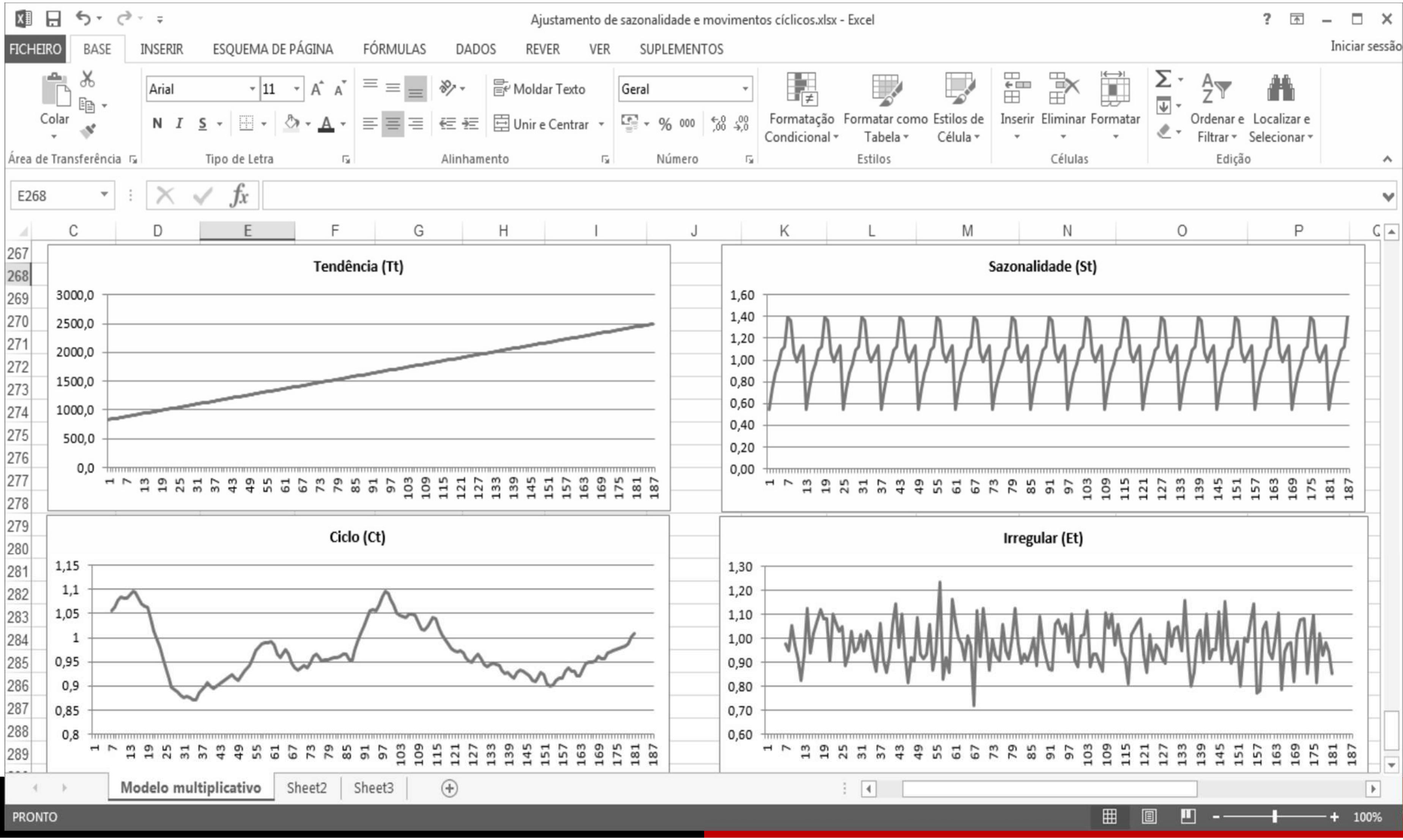
Procedimento para extrair a componente cíclica: eliminar as componentes de tendência e sazonal. No modelo multiplicativo fazer $\frac{Y_t}{MM_t * S_t}$

Componentes de uma Série Temporal

Ciclo

Ilustração: Componentes das vendas de vinho tinto, Australia

(Caiado, 2016)



Componentes de uma Série Temporal
Efeitos de calendário

A série Y_t pode depender do número de dias úteis de cada semana ou mês, ou do número de fins de semana ou feriados de cada mês

Por exemplo, o volume de vendas de um hipermercado tende a apresentar uma forte componente sazonal, com picos de vendas nos dias de sábado o que exige um ajustamento das previsões mensais

Tipos de ajustamento mais importantes:

- comprimento do mês (length-of-month adjustment)
- número de fins de semana do mês (4 vs. 5 week periods adjustment)
- número de dias úteis do mês ou semana (trading day adjustment).

Ajustamento length-of-month

Transformam-se os dados da série mensal de acordo com o # de dias do respetivo mês:

$$Y_t^{(m)} = Y_t \frac{\overline{nd}}{nd_t}$$

onde $\overline{nd}$ é # médio de dias por mês ($\overline{nd} = \frac{365.25\, dias}{12\, meses} = 30.4375$ e nd_t é o # de dias do mês t

Ajustamento 4 / 5 week periods

Transforma-se os dados da série mensal de acordo com o # de fins de semana (sábados e/ou domingos) de cada mês:

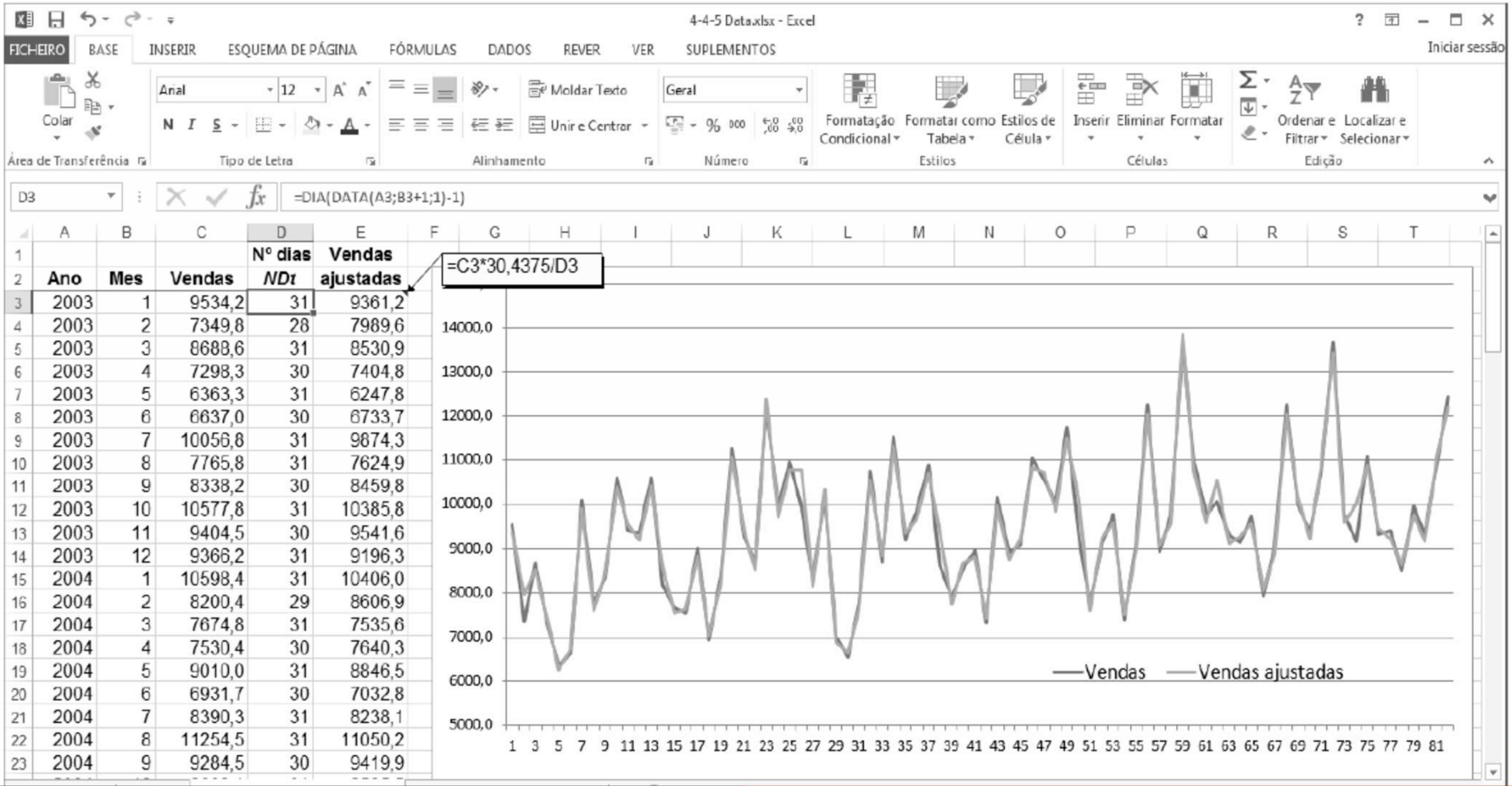
$$Y_t^{(s)} = Y_t \frac{\overline{ns}}{ns_t}$$

onde $\overline{ns}$ é # médio de fins de semana por mês ($\overline{ns} = \frac{52.18\, semanas}{12\, meses} = 4.348$ e ns_t é o # de fins de semana do mês t

Componentes de uma Série Temporal
Efeitos de calendário

Ilustração: vendas sazonais (Caiado, 2016)

Figura 3.20. Vendas ajustadas de efeitos de calendário (*month adjustment*)



Ano	Mes	Vendas	Nº dias ND_t	Vendas ajustadas
2003	1	9534,2	31	9361,2
2003	2	7349,8	28	7989,6
2003	3	8688,6	31	8530,9
2003	4	7298,3	30	7404,8
2003	5	6363,3	31	6247,8
2003	6	6637,0	30	6733,7
2003	7	10056,8	31	9874,3
2003	8	7765,8	31	7624,9
2003	9	8338,2	30	8459,8
2003	10	10577,8	31	10385,8
2003	11	9404,5	30	9541,6
2003	12	9366,2	31	9196,3
2004	1	10598,4	31	10406,0
2004	2	8200,4	29	8606,9
2004	3	7674,8	31	7535,6
2004	4	7530,4	30	7640,3
2004	5	9010,0	31	8846,5
2004	6	6931,7	30	7032,8
2004	7	8390,3	31	8238,1
2004	8	11254,5	31	11050,2
2004	9	9284,5	30	9419,9

Ilustração: vendas sazonais (Caiado, 2016)

Figura 3.22: Vendas ajustadas pelo n.º de fins de semana completos

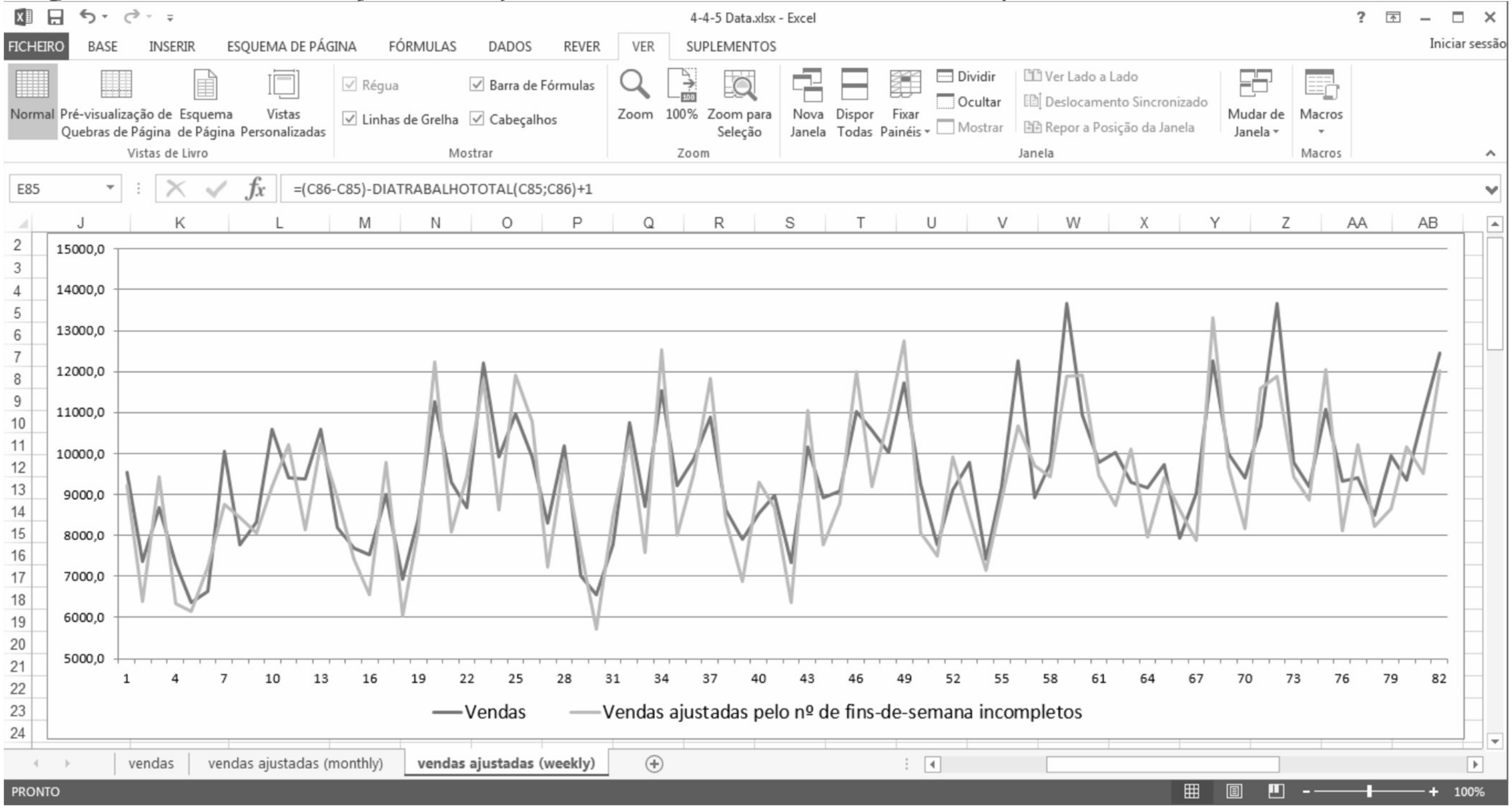
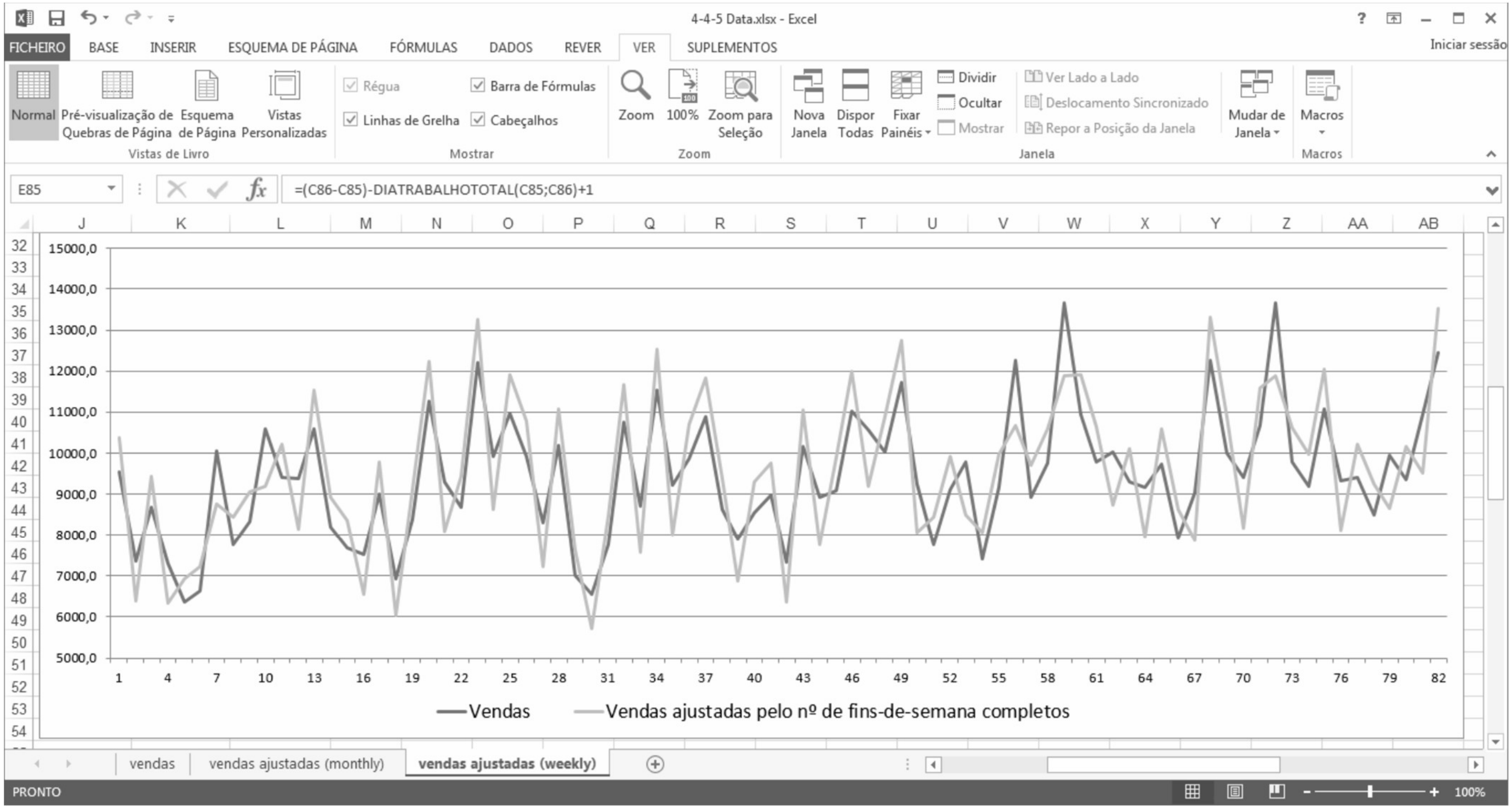


Ilustração: vendas sazonais (Caiado, 2016)

Figura 3.22: Vendas ajustadas pelo n.º de fins de semana incompletos



Alisamento exponencial

Este método serve de base à previsão, sendo também a base de métodos de previsão mais sofisticados.

Objectivo: promover o alisamento, tendo em conta o facto de se esperar que a correlação entre observações de períodos adjacentes ou próximos ser mais forte do que entre observações distantes no tempo

Variantes de alisamento:

- Exponencial simples: séries sem tendência e sem sazonalidade
- Exponencial duplo: séries com tendência mas sem sazonalidade
- Holt: séries com tendência mas sem sazonalidade
- Holt-Winters: séries com tendência e com sazonalidade

Alisamento exponencial

Considera-se uma série de dados observados até ao instante t. Em t tem-se a observação Y_t e a sua estimativa $\hat{Y}_t$. *Em t+1* obtém-se Y_{t+1}, sendo a estimativa correspondente $\hat{Y}_{t+1}$ obtida de acordo com a forma de alisamento considerada

Posteriormente, obtém-se as previsões h passos à frente $\hat{Y}_{t+h}$

Alisamento exponencial
Exponencial simples

Aplica-se a séries sem sazonalidade e sem tendência

1. $\hat{Y}_{t+1}$ é a média ponderada por uma constante de alisamento α, $0 < \alpha < 1$, de Y_t com $\hat{Y}_t$:

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t \qquad (1)$$
$$= \hat{Y}_t + \alpha\left(Y_t - \hat{Y}_t\right), \text{ t=1, ...T}$$

- Constante de alisamento α:
  - quando maior, maior o peso do presente (Y_t) relativamente ao passado ($\hat{Y}_t$), sendo menor o alisamento
  - A sua escolha pode ser feita de modo a minimizar uma função dos erros de previsão (EQM, por exemplo)

Alisamento exponencial
Exponencial simples

- Expressão recursiva para $\hat{Y}_{t+1}$ (versão com médias móveis ponderadas, com ponderação decrescente de tipo exponencial/geométrico com antiguidade):

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \cdots (1-\alpha)^t \hat{Y}_1$$

- Inicialização

A primeira previsão a calcular é para t=1, $\hat{Y}_2$, onde se necessita de $\hat{Y}_1$, que se obtém alternativamente por

  - Igualização à primeira observação, $\hat{Y}_1 = Y_1$, e que implica $\hat{Y}_2 = Y_1$
  - média das primeiras 4 ou 5 observações

Quanto menor α, maior a influencia da escolha da inicialização

Alisamento exponencial
Exponencial simples

- Expressão alternativa

Em lugar de (1), há literatura (por exemplo, Newbold, Carlson & Thorne, 2013) que escreve

$$\hat{Y}_{t+1} = \alpha Y_{t+1} + (1-\alpha)\hat{Y}_t \qquad (2)$$

Neste caso, inicializando por $\hat{Y}_1 = Y_1$, $\hat{Y}_2 = \alpha Y_2 + (1-\alpha)Y_1 \neq Y_1$. De facto, neste contexto, $\hat{Y}_2$ é dado por o $\hat{Y}_3$ da eq. (1) ($\hat{Y}_3 = \alpha Y_2 + (1-\alpha)\hat{Y}_2 = \alpha Y_2 + (1-\alpha)Y_1$) e assim sucessivamente.

Procedimentos automáticos de excel e stata seguem a eq. (1)

- Erro de previsão: para (1) é dado por $e_t = Y_t - \hat{Y}_t$ e para (2) por $e_t = Y_t - \hat{Y}_{t-1}$, sendo numericamente iguais

Alisamento exponencial
Exponencial simples

2. Previsão a h passos

As previsões para horizontes maiores do que um passo são constantes para todo o horizonte temporal:

$$\hat{Y}_{t+h} = \hat{Y}_t, h = 1, 2, \dots$$

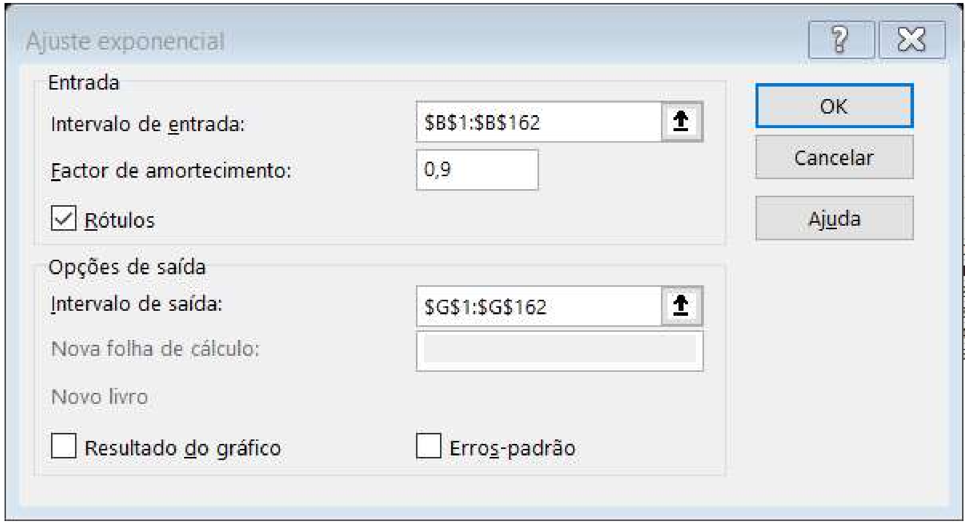
Por este motivo, este método só deve ser utilizado para séries sem tendência e sem sazonalidade

Alisamento exponencial
Exponencial simples

Exemplo: vendas de livros em hardcore

Em excel

- Facilmente se introduz a formula manualmente ou



A	B	C	D	
Data	Vendas	Vhat(manu	Vhat(excel)	Et
1	45,9	45,9	#N/D	
2	45,4	45,9	45,9	
3	42,8	45,85	45,85	
4	34,4	45,545	45,545	
5	31,9	44,4305	44,4305	-
6	36,6	43,17745	43,17745	-
7	39,2	42,51971	42,519705	-
8	41,4	42,18773	42,1877345	-
9	40,3	42,10896	42,10896105	-
10	43,1	41,92806	41,92806495	1,
11	43,2	42,04526	42,04525845	1,
12	41,2	42,16073	42,16073261	-
13	38,4	42,06466	42,06465934	-
14	38,3	41,69819	41,69819341	-

- Note-se que o factor de amortecimento a colocar no excel é $(1-\alpha)$
- Para determinar α optimo, usar o Solver

Alisamento exponencial

Exponencial simples

Em STATA

```
. tsset data
        time variable:  data, 1 to 161
                delta:  1 unit
```

- Escolhendo $\alpha = 0.1$

```
. tssmooth exponential vendashat_01=vendas, parms(.1) s0(45.9)
exponential coefficient  =       0.1000
sum-of-squared residuals =       7031.8
root mean squared error  =       6.6087
```

* resultado igual ao manual do excel

- Minimizando o EQM

```
. tssmooth exponential vendashat_opt=vendas, s0(45.9)
computing optimal exponential  coefficient (0,1)
optimal exponential coefficient =        0.9998
sum-of-squared residuals        =     1275.0797
root mean squared error         =      2.814205
```

- O erro quadrático médio é bem menor com o valor optimo de α

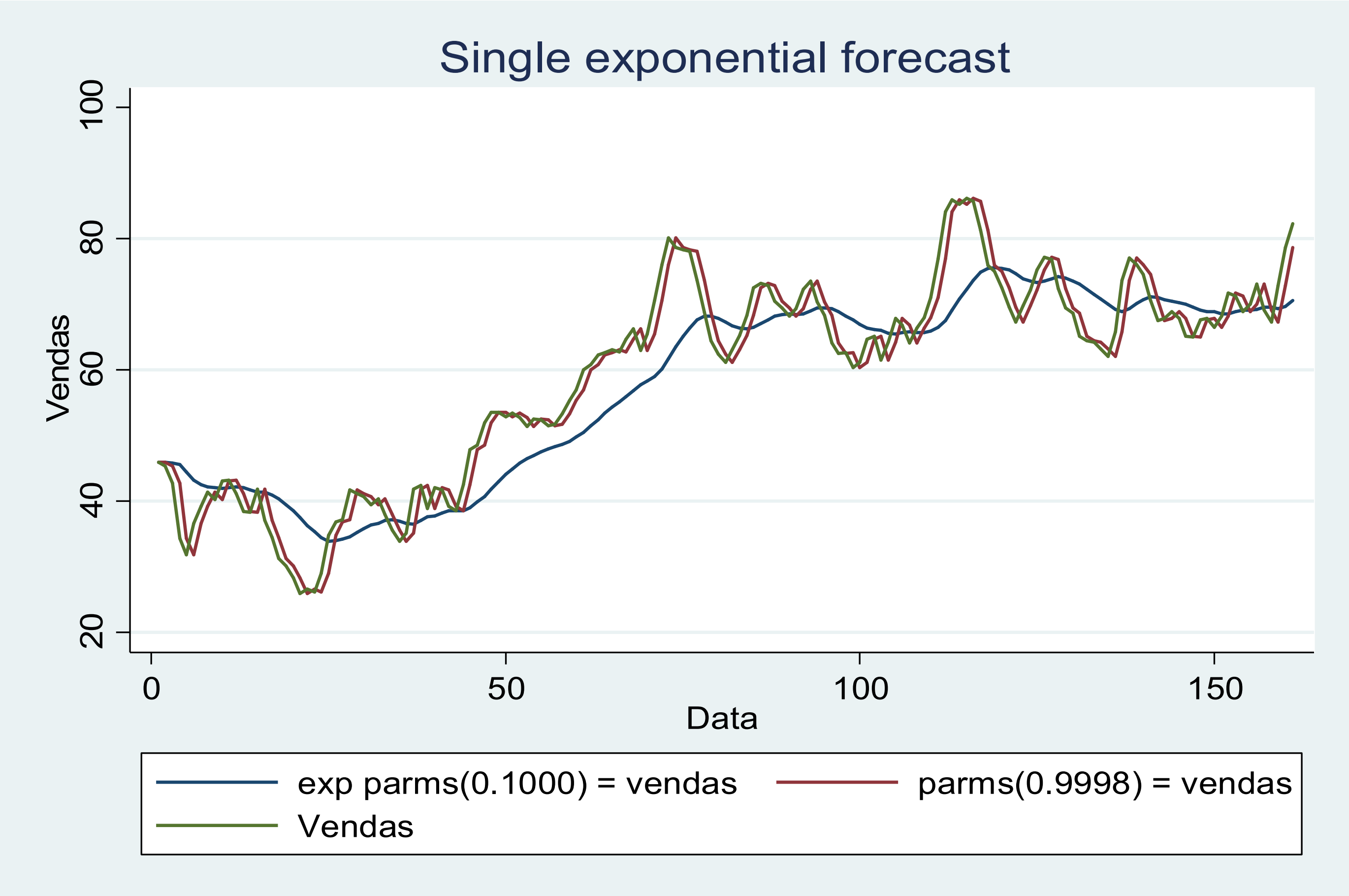
Alisamento exponencial
Exponencial simples

- Calculo do EQM

```
. gen sqe_01=(vendas-vendashat_01)^2
. gen sqe_opt=(vendas-vendashat_opt)^2

. sum sqe_01 sqe_opt

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      sqe_01 |       161    43.67554    60.19993          0   336.6491
     sqe_opt |       161     7.91975    11.30058          0   70.56701

. line vendashat_01 vendashat_opt vendas data, title("Single exponential
forecast") ytitle(Vendas) xtitle(Data)
```

Alisamento exponencial

Exponencial simples



- Note-se a proximidade da previsão baseada na escolha optima relativamente à série original

O modelo de alisamento exponencial duplo aplica-se a séries com tendência linear (mas não sazonalidade)

1) Aplicação do método de alisamento exponencial simples duas vezes, utilizando a mesma constante de alisamento:

$$\hat{Y}_t^{[1]} = \alpha Y_t + (1-\alpha)\hat{Y}_{t-1}^{[1]}$$
$$\hat{Y}_t^{[2]} = \alpha \hat{Y}_t^{[1]} + (1-\alpha)\hat{Y}_{t-1}^{[2]}$$

onde $\hat{Y}_t^{[1]}$ e $\hat{Y}_t^{[2]}$ são as séries resultantes do alisamento simples e duplo, respectivamente

- Inicialização

Para obter $\hat{Y}_1^{[1]}$ e $\hat{Y}_1^{[2]}$, faz-se

$$\hat{Y}_1^{[1]} = \hat{a}_1 - \hat{b}_1 \frac{(1-\alpha)}{\alpha}$$
$$\hat{Y}_1^{[2]} = \hat{a}_1 - 2\hat{b}_1 \frac{(1-\alpha)}{\alpha}$$

onde a_1 e b_1 podem resultar de uma destas alternativas:

- regressão da série numa tendência linear

$$Y_t = a_1 + b_1 t + u_t$$

* por defeito, o STATA, utiliza apenas metade das observações nesta regressão

- derivados de médias

$$a_1 = \bar{y}_1 - b_1 \frac{k+1}{2} \text{ e } b_1 = \frac{\bar{y}_2 - \bar{y}_1}{k}$$

onde $\bar{y}_1$ é a média das primeiras *k* observações da série e $\bar{y}_2$ a média das *k* observações seguintes

Alisamento exponencial
Exponencial duplo

- Previsão 1 passo à frente

$$\hat{Y}_{t+1} = a_t + b_t$$

onde $a_t = 2\hat{Y}_t^{[1]} - \hat{Y}_t^{[1]}$, $b_t = \left(\hat{Y}_t^{[1]} - \hat{Y}_t^{[2]}\right)\frac{\alpha}{1-\alpha}$ e $a_1\ e\ b_1$ foram antes obtidos

*Não há comando automático no excel que possa ser utilizado, pois mesmo o exponencial simples repetido, inicia com a primeira observação

2. Previsão para h passos à frente:

$$\hat{Y}_{t+h} = a_t + b_t h\ h = 1, 2, \dots$$

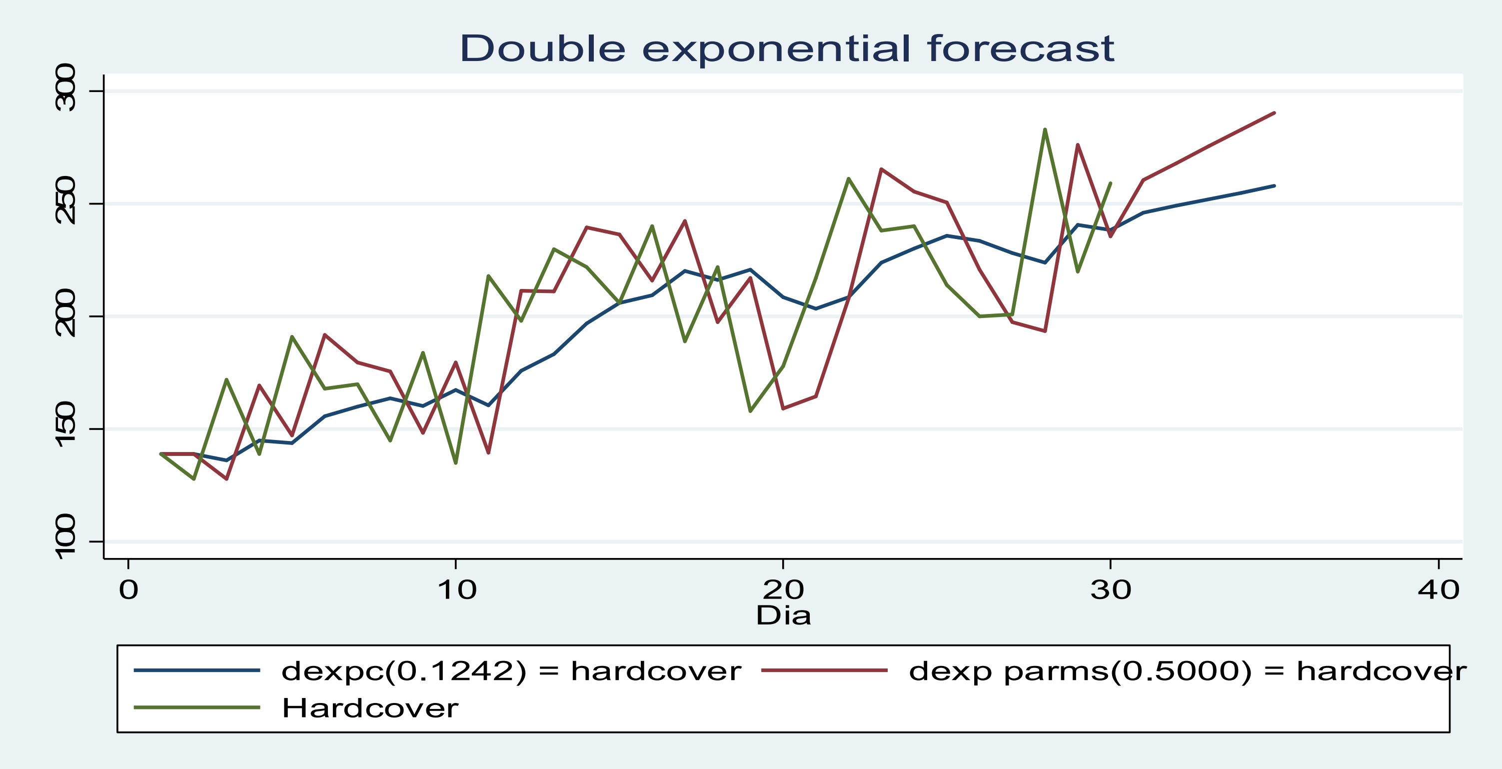
Alisamento exponencial
Exponencial duplo

Exemplo: vendas de livros em hardcore

STATA

```
.tsset dia

. tssmooth dexponential double D=hardcover, p(.5) s0(139,139)

double-exponential coefficient  =       0.5000
sum-of-squared residuals        =        17337
root mean squared error         =        24.04

. tssmooth dexponential double f1=hardcover , p(.5) s0(139,139) forecast(5)

double-exponential coefficient  =       0.5000
sum-of-squared residuals        =        45616
root mean squared error         =       38.994

. tssmooth dexponential double fopt=hardcover , s0(139,139) forecast(5)
computing optimal double-exponential coefficient (0,1)

optimal double-exponential coefficient =        0.1242
sum-of-squared residuals               =     29194.749
root mean squared error                =     31.195485

. line fopt f1 hardcover dia, title("Double exponential forecast")
```

Alisamento exponencial
Exponencial duplo



- Note-se o menor alisamento para α=0.5 e o mínimo EQM na previsão optima (mesmo menor do que o do excel em Caiado (2016), p. 114

Comparação com os resultados excel (Caiado, 2016)

Figura 4.12. Previsões de vendas de livros (AED, α = 0,05 e α ótimo)

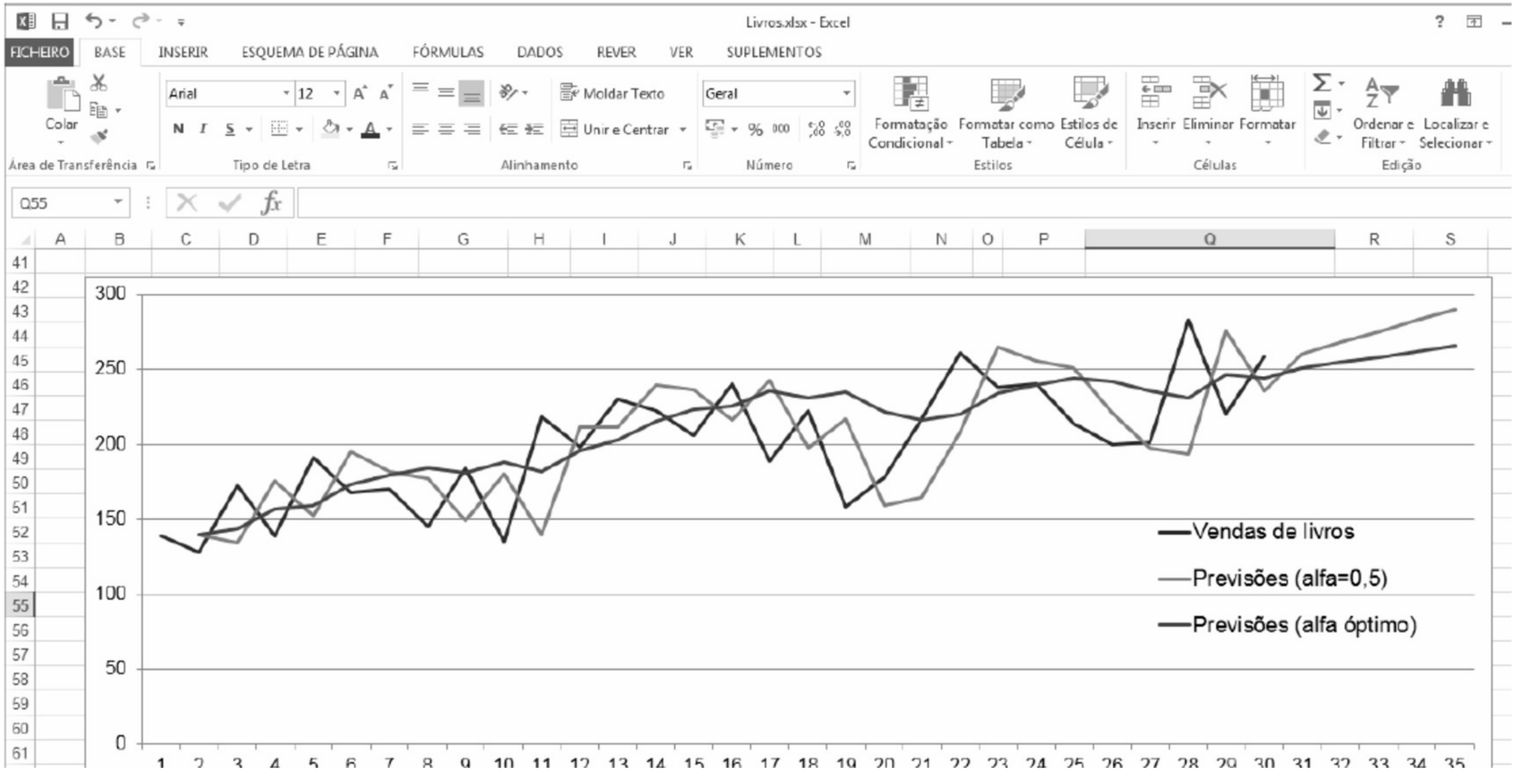
		Alisamento duplo com α = 0,5						Alisamento duplo com alfa óptimo					
Dia	Vendas	Mt	Dt	a(t)	b(t)	Previsão	Erro	Mt	Dt	a(t)	b(t)	Previsão	Erro
25	214	225,54	227,90	223,17	-2,37	250,67	-36,67	206,89	176,48	237,30	4,19	244,17	-30,17
26	200	212,77	220,34	205,20	-7,57	220,80	-20,80	206,05	180,06	232,04	3,58	241,49	-41,49
27	201	206,88	213,61	200,16	-6,73	197,63	3,37	205,44	183,14	227,74	3,08	235,63	-34,63
28	283	244,94	229,28	260,61	15,67	193,43	89,57	214,84	186,98	242,70	3,84	230,82	52,18
29	220	232,47	230,87	234,07	1,60	276,27	-56,27	215,47	190,43	240,50	3,45	246,54	-26,54
30	259	245,74	238,30	253,17	7,43	235,67	23,33	220,74	194,11	247,38	3,67	243,95	15,05
31	P31					260,60						251,05	
32	P32					268,03						254,72	
33	P33					275,46						258,40	
34	P34					282,89						262,07	
35	P35					290,32						265,74	

	α (alfa)	EQM
Escolha	0,5	1542,65
Óptimo	0,121186367	1022,62

- α=0.5: apesar de nos valores iniciais a previsão do STATA diferir da do excel (a inicialização é diferente), nos últimos períodos e fora da amostra os resultados são muito similares
- α optimo ligeiramente diferente, com o do STATA a gerar um menor EQM

Alisamento exponencial
Exponencial duplo

- A representação do excel é muito parecida com a de STATA:



Alisamento exponencial
Método de Holt

É uma generalização do método de alisamento exponencial duplo, que se aplica nas mesmas condições (séries com tendência linear e sem sazonalidade), mas que usa duas constantes de alisamento, α e β, $0<\alpha,\beta<1$.

1. Usa

$$a_t = \alpha Y_t + (1-\alpha)(a_{t-1} + b_{t-1})$$
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$

com os valores iniciais para a_1 e b_1 obtidos como no método anterior

2. Previsão h passos à frente

$$\hat{Y}_{t+h} = a_t + b_t h, \qquad h = 1, 2, \ldots$$

Exemplo: vendas de livros em hardcore

α=0.25 e β=0.1

STATA

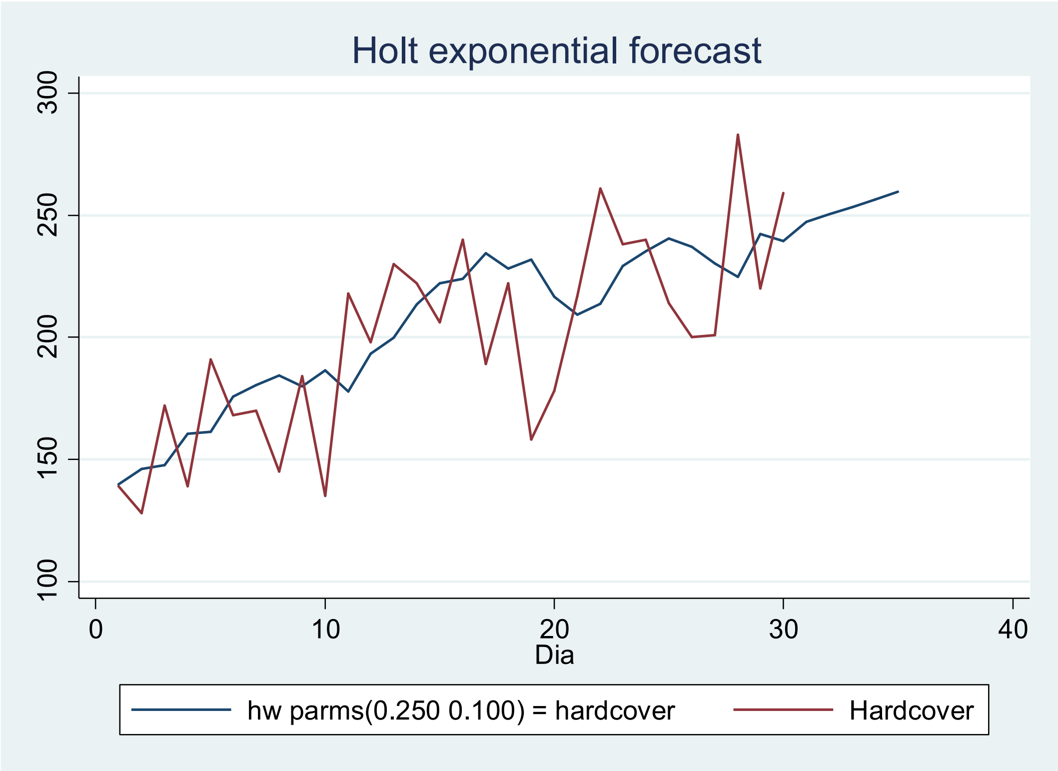
```
. tssmooth hwinters fH=hardcover , parms (0.25 0.1) s0(133.22 6.56)
forecast(5)

Specified weights:
                   alpha = 0.2500
                    beta = 0.1000
sum-of-squared residuals = 28430.13
 root mean squared error = 30.78427
```

- Não convergiu na obtenção de valores para α e β

- `line fH hardcover dia, title("Holt exponential forecast")`

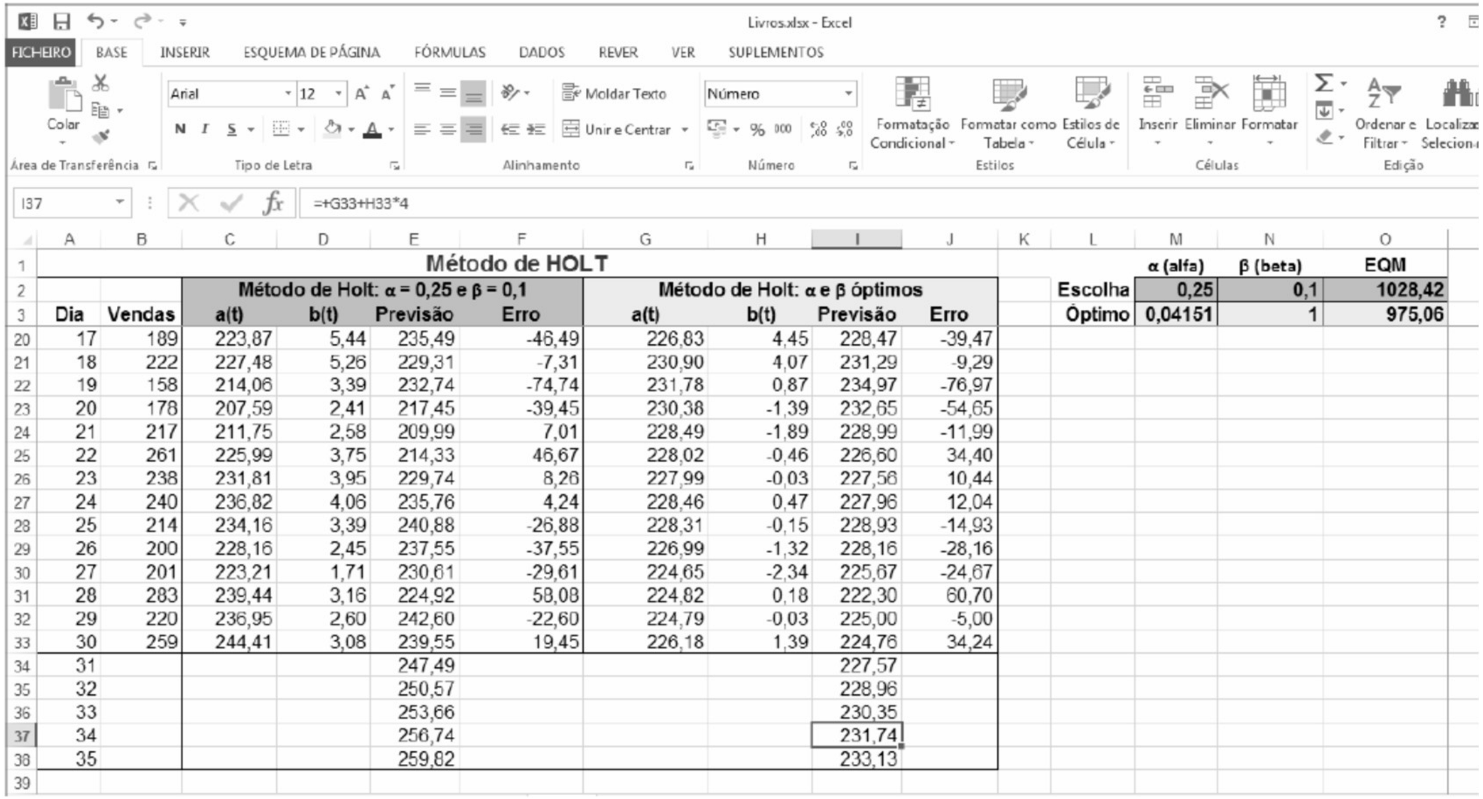
Alisamento exponencial

Método de Holt



Alisamento exponencial
Método de Holt

Comparação com excel (Caiado, 2016)



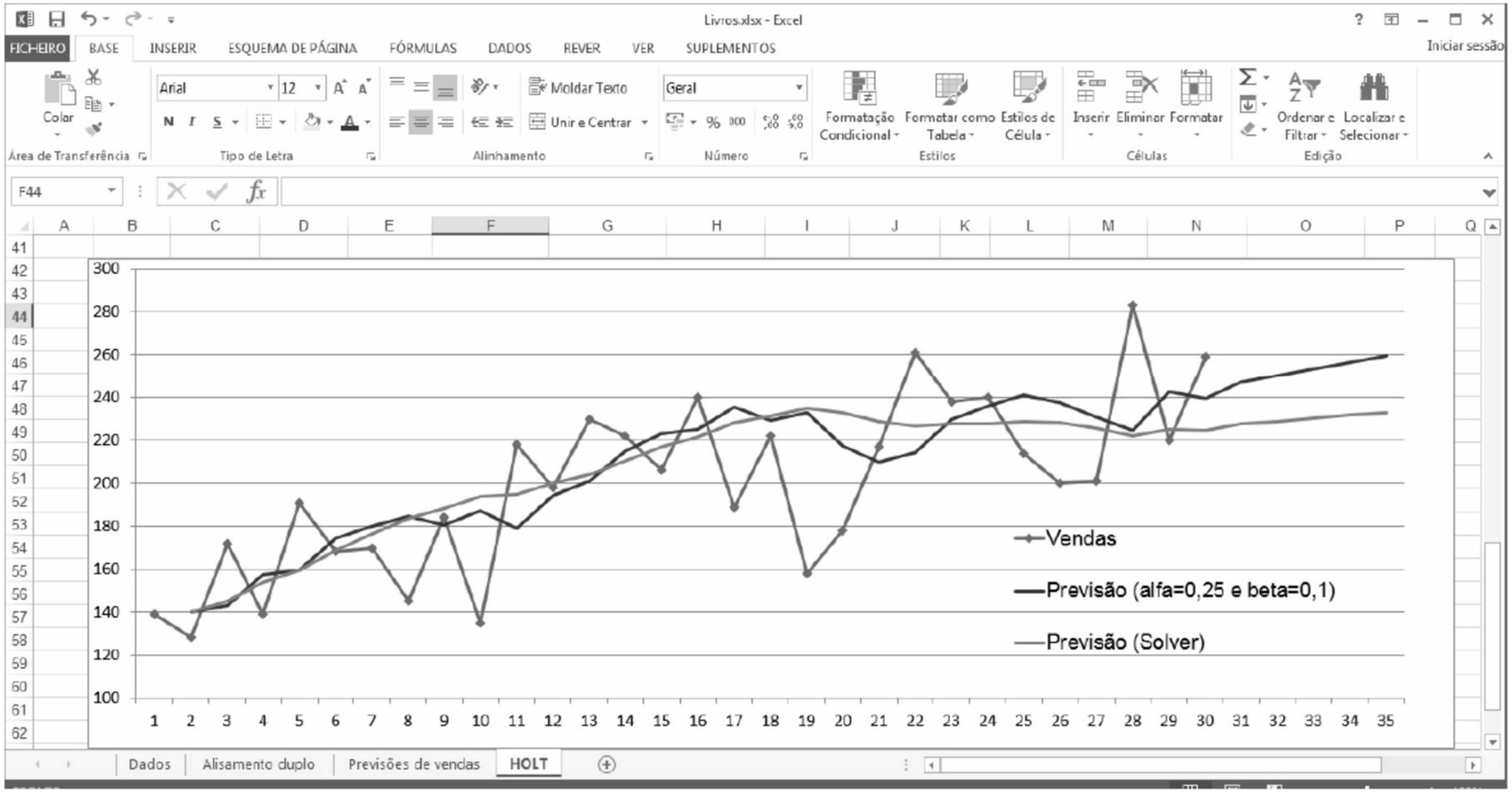
Método de HOLT

		Método de Holt: α = 0,25 e β = 0,1				Método de Holt: α e β óptimos			
Dia	Vendas	a(t)	b(t)	Previsão	Erro	a(t)	b(t)	Previsão	Erro
17	189	223,87	5,44	235,49	-46,49	226,83	4,45	228,47	-39,47
18	222	227,48	5,26	229,31	-7,31	230,90	4,07	231,29	-9,29
19	158	214,06	3,39	232,74	-74,74	231,78	0,87	234,97	-76,97
20	178	207,59	2,41	217,45	-39,45	230,38	-1,39	232,65	-54,65
21	217	211,75	2,58	209,99	7,01	228,49	-1,89	228,99	-11,99
22	261	225,99	3,75	214,33	46,67	228,02	-0,46	226,60	34,40
23	238	231,81	3,95	229,74	8,26	227,99	-0,03	227,56	10,44
24	240	236,82	4,06	235,76	4,24	228,46	0,47	227,96	12,04
25	214	234,16	3,39	240,88	-26,88	228,31	-0,15	228,93	-14,93
26	200	228,16	2,45	237,55	-37,55	226,99	-1,32	228,16	-28,16
27	201	223,21	1,71	230,61	-29,61	224,65	-2,34	225,67	-24,67
28	283	239,44	3,16	224,92	58,08	224,82	0,18	222,30	60,70
29	220	236,95	2,60	242,60	-22,60	224,79	-0,03	225,00	-5,00
30	259	244,41	3,08	239,55	19,45	226,18	1,39	224,76	34,24
31				247,49				227,57	
32				250,57				228,96	
33				253,66				230,35	
34				256,74				231,74	
35				259,82				233,13	

	α (alfa)	β (beta)	EQM
Escolha	0,25	0,1	1028,42
Óptimo	0,04151	1	975,06

- Para α=0.25 e β=0.1 novamente as ultimas previsões e as previsões fora da amostra são muito semelhantes às do STATA

Alisamento exponencial
Método de Holt



Alisamento exponencial
Método de Holt-Winters

Aplica-se a séries com tendência linear e sazonalidade

Forma multiplicativa e aditiva

Na forma multiplicativa:

1. Expressões

$$a_t = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(a_{t-1} + b_{t-1})$$
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$
$$S_t = \gamma \frac{Y_t}{a_t} + (1-\gamma)S_{t-s}$$

onde S_t é o índice sazonal, s é o comprimento de sazonalidade (12 se mensal, 4 se trimestral, 2 se semestral) e α, β e γ $0< \alpha,\beta,\gamma <1$ são constantes de alisamento

Alisamento exponencial
Método de Holt-Winters

- Inicialização: obtém-se a_s, b_s e $S_1 \ldots S_s$
  - $a_s = \frac{\sum_{t=1}^{s} Y_t}{s}$ (média das observações da primeira estação completa)
  - $b_s = \frac{1}{s^2}\left(\sum_{t=s+1}^{2s} Y_t - \sum_{t=1}^{s} Y_t\right) = \frac{1}{s}\left(\frac{\sum_{t=s+1}^{2s} Y_t}{s} - a_s\right)$
    $(\frac{1}{s}(diferença\ de\ médias\ da\ segunda\ e\ primeira\ estação))$
  - $S_1 = \frac{Y_1}{a_s}, S_2 = \frac{Y_2}{a_s} \ldots S_s = \frac{Y_s}{a_s}$ (divide cada uma das primeiras s observações por a_s

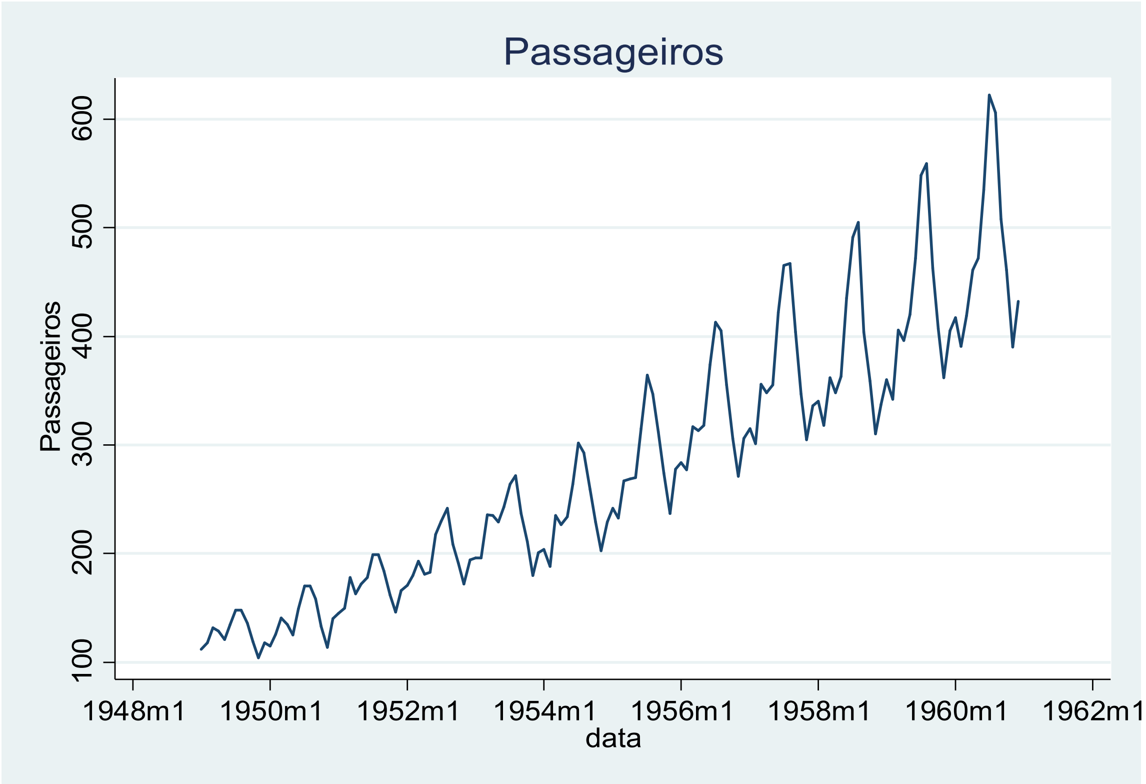
2. Previsão a h passos à frente:

$$\hat{Y}_{t+h} = [a_t + b_t h] S_{t+h-s},\ \ h = 1, 2, \ldots$$

Usa os últimos valores de a e b e o ultimo valor de S correspondente ao mês, trimestre, …

Alisamento exponencial
Método de Holt-Winters

Exemplo: passageiros de um tipo de viagem (Caiado, 2016)



- A série tem tendência e movimentos sazonais

Alisamento exponencial
Método de Holt-Winters

α =0.4 β=0.25 γ=0.15

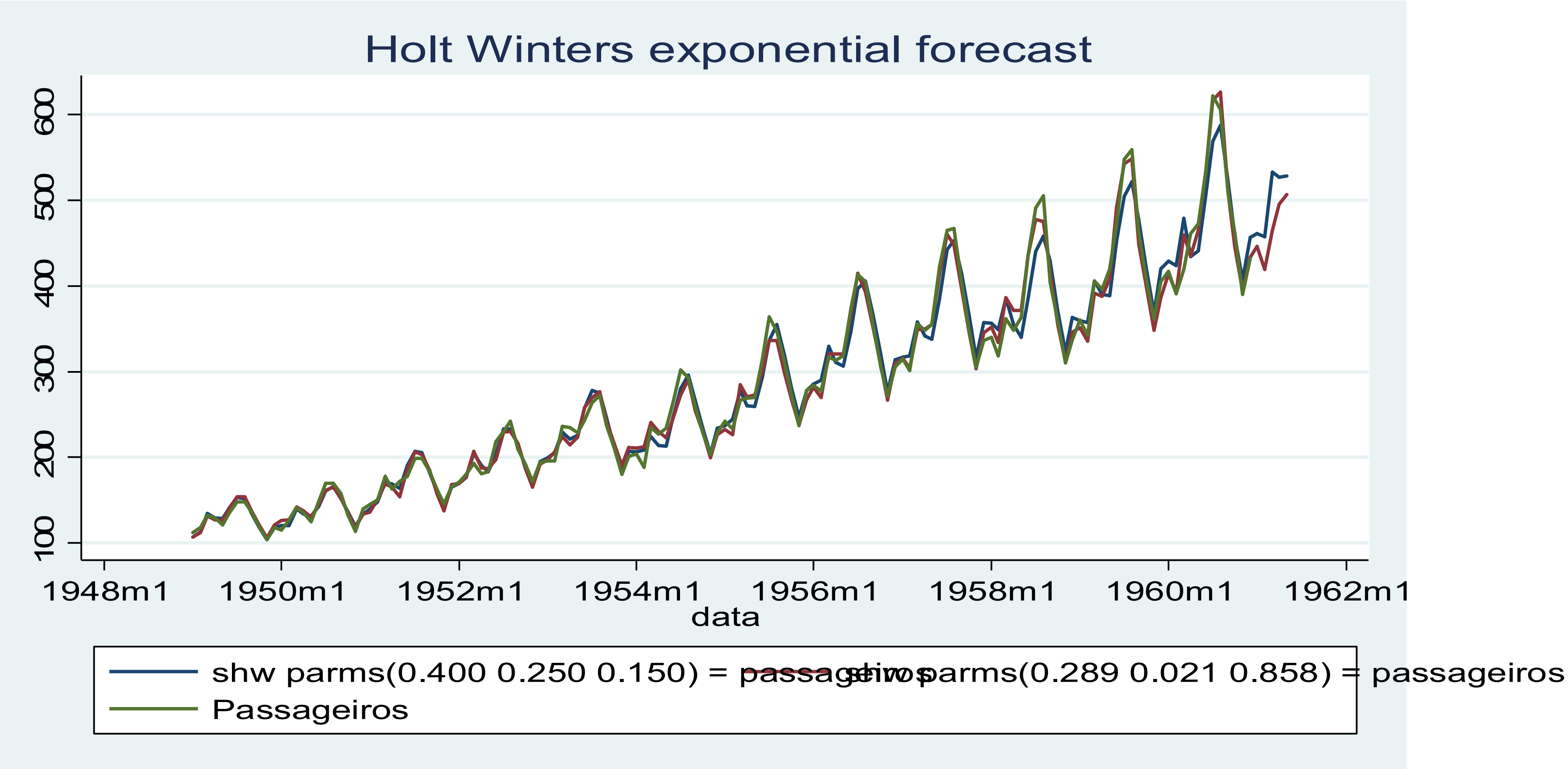
STATA

```
. tsset data

. tssmooth shwinters passhat = passageiros, parms(0.4 0.25 0.15)
forecast(5)

Specified weights:
                    alpha = 0.4000
                     beta = 0.2500
                    gamma = 0.1500
sum-of-squared residuals = 40662.14
 root mean squared error = 16.80405
```

Alisamento exponencial
Método de Holt-Winters

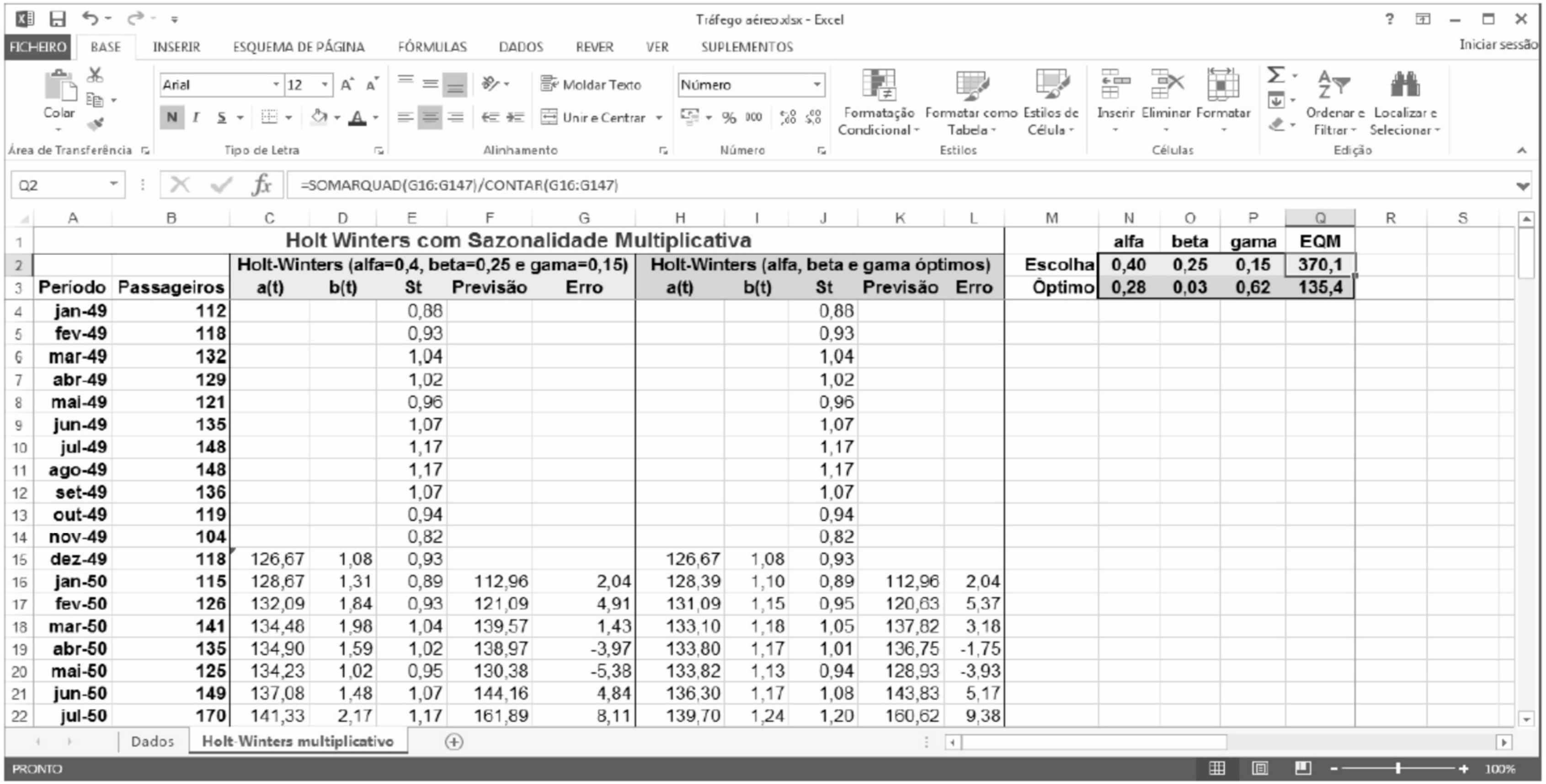
```
. tssmooth shwinters passhatopt = passageiros, forecast(5)
computing optimal weights
Iteration 0:   penalized RSS = -56587.051  (not concave)
…
Optimal weights:
                              alpha = 0.2891
                               beta = 0.0212
                              gamma = 0.8579
penalized sum-of-squared residuals = 16514.49
          sum-of-squared residuals = 16514.49
          root mean squared error  = 10.70906

. line passhat passhatopt passageiros data, title("Holt Winters
exponential forecast")
```

Alisamento exponencial
Método de Holt-Winters



- Note-se a semelhança das previsões

Comparação com excel (Caiado, 2016)



Tráfego aéreo.xlsx - Excel

Q2 =SOMARQUAD(G16:G147)/CONTAR(G16:G147)

Holt Winters com Sazonalidade Multiplicativa

		Holt-Winters (alfa=0,4, beta=0,25 e gama=0,15)					Holt-Winters (alfa, beta e gama óptimos)				
Período	Passageiros	a(t)	b(t)	St	Previsão	Erro	a(t)	b(t)	St	Previsão	Erro
jan-49	112			0,88					0,88		
fev-49	118			0,93					0,93		
mar-49	132			1,04					1,04		
abr-49	129			1,02					1,02		
mai-49	121			0,96					0,96		
jun-49	135			1,07					1,07		
jul-49	148			1,17					1,17		
ago-49	148			1,17					1,17		
set-49	136			1,07					1,07		
out-49	119			0,94					0,94		
nov-49	104			0,82					0,82		
dez-49	118	126,67	1,08	0,93			126,67	1,08	0,93		
jan-50	115	128,67	1,31	0,89	112,96	2,04	128,39	1,10	0,89	112,96	2,04
fev-50	126	132,09	1,84	0,93	121,09	4,91	131,09	1,15	0,95	120,63	5,37
mar-50	141	134,48	1,98	1,04	139,57	1,43	133,10	1,18	1,05	137,82	3,18
abr-50	135	134,90	1,59	1,02	138,97	-3,97	133,80	1,17	1,01	136,75	-1,75
mai-50	125	134,23	1,02	0,95	130,38	-5,38	133,82	1,13	0,94	128,93	-3,93
jun-50	149	137,08	1,48	1,07	144,16	4,84	136,30	1,17	1,08	143,83	5,17
jul-50	170	141,33	2,17	1,17	161,89	8,11	139,70	1,24	1,20	160,62	9,38

	alfa	beta	gama	EQM
Escolha	0,40	0,25	0,15	370,1
Óptimo	0,28	0,03	0,62	135,4

Dados | Holt-Winters multiplicativo

- Para α=0.25 e β=0.1 novamente as ultimas previsões e as previsões fora da amostra são muito semelhantes às do STATA

Alisamento exponencial
Método de Holt-Winters

K151 f_x =($H$147+$I$147*4)*J139

Holt Winters com Sazonalidade Multiplicativa

		Holt-Winters (alfa=0,4, beta=0,25 e gama=0,15)					Holt-Winters (alfa, beta e gama ótimos)				
Periodo	Passageiros	a(t)	b(t)	St	Previsão	Erro	a(t)	b(t)	St	Previsão	Erro
jul-60	622	485,17	8,97	1,22	564,94	57,06	459,30	3,08	1,35	612,36	9,64
ago-60	606	500,96	10,67	1,19	585,76	20,24	459,85	3,00	1,33	618,16	-12,16
set-60	508	504,11	8,79	1,03	527,38	-19,38	463,22	3,01	1,10	506,57	1,43
out-60	461	511,44	8,43	0,90	464,33	-3,33	471,41	3,18	0,97	443,27	17,73
nov-60	390	507,82	5,41	0,79	413,99	-23,99	473,70	3,15	0,82	392,63	-2,63
dez-60	432	499,05	1,87	0,90	464,04	-32,04	476,36	3,13	0,91	433,60	-1,60
jan-61					456,04					448,11	
fev-61					452,90					422,77	
mar-61					522,98					476,86	
abr-61					518,29					496,91	
mai-61					513,20					511,74	
jun-61					572,62					584,46	
jul-61					625,87					671,92	
ago-61					611,20					664,42	
set-61					529,98					552,90	
out-61					468,41					491,23	
nov-61					411,59					421,34	
dez-61					468,51					466,56	

	alfa	beta	gama	EQM
Escolha	0,40	0,25	0,15	370,1
Óptimo	0,28	0,03	0,62	135,4

P(Abr61)=[a(Dez60)+b(Dez60)*4]*S(Abr60)
=(476,36+3,13*4)*1,02
= 498,7 (com arrendondamentos)

- Tal como no STATA, as previsões fora da amostra são menores quando os parâmetros optimos são usados

Na forma aditiva:

1.

$$a_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1})$$
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$
$$S_t = \gamma(Y_t - a_t) + (1-\gamma)S_{t-s}$$

A inicialização é igual ao método multiplicativo, excetuando os índices sazonais

$$S_1 = Y_1 - a_s, S_2 = Y_2 - a_s \ldots S_s = Y_s - a_s$$

2. Previsão a h passos à frente:

$$\hat{Y}_{t+h} = [a_t + b_t h] + S_{t+h-s},\ h = 1, 2, \ldots$$

Forma multiplicativa ou aditiva?